Selection of tobacco etch virus protease variants with enhanced oxidative stability for tag-removal in refolding of two disulfide-rich proteins

Bayar Enkhtuya  
Anhui Agricultural University

Yuanyuan Ren  
Anhui Agricultural University

Yafang Hu  
Anhui Agricultural University

Yinhua Chen  
Anhui Agricultural University

Jiong Hu  
Anhui Agricultural University

Xuelian Yu  
Anhui Agricultural University

Weiyu Wang  
Anhui Agricultural University

Jun Fan (fanjun@ahau.edu.cn)  
Anhui Agricultural University

Research Article

Keywords: TEVp, mutation, cysteine residues, tag removal, Escherichia coli

Posted Date: February 18th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-174410/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
**Abstract**

Tobacco etch virus protease (TEVp) is a powerful enzymatic reagent for removing fusion tag. In this work, we constructed nine TEVp variants with introducing one to three mutations of C19S, C110S and C130S into the soluble TEVp variant, TEVp$^{5M}$. Using the C-terminal green fluorescent protein (GFP) variant reporter, all constructs showed different solubility levels among four *E. coli* strains. The TEVp$^{5M}$ containing the C110S and/or C130S mutations in the hyperoxic strain showed the enhanced the cleavage activity. Addition of dithiothreitol to the cultural medium increased the activity of certain constructs produced in the BL21(DE3), contrary to the added hydrogen peroxide, due to cytoplasmic redox change measured by the redox sensitive GFP construct. The more cysteine residues in the purified TEVp$^{5M}$ were modified specifically than those in the other variants. All purified constructs showed similar specific activities in the presence of 5 mM dithiothreitol. In the buffer containing the compounds to aid disulfide bond formation of the refolded protein, the double mutant TEVp$^{5M}$C110S/C130S exhibited the highest cleavage efficiency. This variant was efficient for removing the fusion tag after refolding of cellulose-binding module tagged disulfide-rich proteins including bovine enteropeptidase and maize peroxidase absorbed on the regenerated amorphous cellulose.

**Introduction**

Owing to high specificity, the mutated TEVp variant with the decreased auto-inactivation via self-cleavage, and the improved protein solubility is a commonly-used tool in various biotechnological applications [1]. Other mutated amino acid residues are contributed to enhance the thermostability [2], inhibit the translational modification in eukaryote cells [3], change of the thiol protease to serine protease [4], alter substrate specificity [5], and increase the cleavage activity [6].

Because of its easier culture, short life cycle, well-known genetics, and easy genetic manipulation, *Escherichia coli* is the preferred choice for production of recombinant protein [7]. However, many proteins are formed into insoluble aggregates, also called as inclusion bodies (IBs). Since IBs are produced at higher level, easily prepared and purified, target protein in IBs is refolded for studying protein structure and function [8]. Several fusion tags improve the refolding efficiency, and the specific protease is used for removing the fused tag [9]. On-column refolding dependent on the fused affinity tag saves the refolding period and increases yields of refolded protein [10]. The cellulose-binding module (CBM) tag fused with the single-chain antibodies is efficient for renaturation of the fusion protein bound to the cellulose [11]. For renaturing the disulfide-rich proteins, the reagents for enabling disulfide bond formation were required, such as the reduced and oxidized glutathione (GSH and GSSG), cysteine and cystine [9]. Little information of these reagents on affecting the TEVp activity is available.

Previously, we combined missense mutations in the TEVp to generate the TEVp$^{5M}$ variant with inhibiting the auto-cleavage, and improving protein folding and solubility [12]. The TEVp activity analysis based on the coupled assay is developed [13]. Once the designed fusion protein is cleaved by the sequence-specific protease, the released tag-free *E. coli* diaminopropionate ammonia-lyase (eDAL) from the glutathione S-
transferase (GST) tagged eDAL increases activity significantly. In this study, we mutated and combined three cysteine residues in the TEVp\textsuperscript{SM} and identified the change of protein solubility and cleavage activity in different \textit{E. coli} strains. Among nine variants, only C110S and/or C130S mutations showed little impact on protein production, and afforded the TEVp\textsuperscript{SM} with the improved oxidative stability in vivo and in vitro. The double mutant was efficient for removing the CBM tag as the fusion partner for oriented immobilization and matrix-assisted refolding of the disulfide-bonded bovine enterokinase (also called enteropeptidase) light chain (bEK) and maize peroxidase (mPex) as the isoenzyme C homolog of horseradish peroxidase (HRP). Both disulfide-rich enzymes are widely applied in biotechnology.

**Materials And Methods**

**Bacterial strains, plasmids and reagents**

\textit{E. coli} strains DH5\textalpha{}, BL21(DE3), BL21(DE3)pLys, Rosetta(DE3) and Origami(DE3), the plasmids pET-22b and pET-28b are products of Novagen (Madison, WI). The plasmid pET32a-pEK containing the bEK coding sequence was kindly provided by Professor Zhao Zhongbao [14]. The plasmids pET28-TEVp\textsuperscript{SM} for expressing the double His6-tagged TEVp\textsuperscript{SM} containing the S219V for inhibiting auto-cleavage, and T17S, N68D, and I77V for enhancing protein solubility, L56V and S135G for improving protein folding, the pET28-GFP for expressing the C-terminal EmGFP to detect N-terminal target protein solubility, the pGST-tevS-eDAL and pCBM-tevS-GFP for expressing the fusion proteins as the TEVp substrates were constructed in our laboratory [12,15]. Since the eDAL is cleaved by the EK, the GST fused to the DAL from \textit{Salmonella typhimurium} (sDAL) was constructed [13]. MutanBEST Kit for site-directed mutation and reagents used in plasmid construction and protein overexpression were purchased from Takara (Dalian, China). Nickel-nitrilotriacetic acid (Ni-NTA) agarose was made by Qiagen (Chatsworth, CA). Ultra-15 centrifugal filter tube equipped with Ultracel-10 membrane was obtained from Amicon (USA). 4-acetoamide-4′-maleimidyl-stilbene-2,2′-disulfonate (AMS) was purchased from Invitrogen (USA). The compounds including pyridoxal 5′-phosphate (PLP), DL-α, β-diaminopropionate (L-DAP), ω-phenylenediamine (OPA), 2,4-dinitrophneylhydrazine (2,4-DNP) were bought from Sigma (USA).

**Plasmids construction**

The C19S, C110S and C130S mutations or their mixed mutations were introduced into the TEVp\textsuperscript{SM} gene in the plasmid pET28-TEVp\textsuperscript{SM} independently and successively using each of the primer pairs C19S1 and C19S2, C110S1 and C110S2, C130S1 and C130S2, and pET28-TEVp\textsuperscript{SM} as the template (Fig. S1). After amplification, the products were phosphorylated, ligated, and sequenced. Then, each of the sequence was excised with \textit{Nco I}/\textit{Xho I} and subcloned into \textit{Nco I}/\textit{Sal I} sites of pET28-GFP for expressing the constructed TEVp tagged emerald GFP (EmGFP) to quantitatively analyze protein solubility. Due to the Origami (DE3) strain bearing kanamycin resistance, each of the coding sequence through \textit{Xba I} and \textit{Xho I} excision was inserted into the \textit{Xba I}/\textit{Xho I} sites of the pET-22b vector. In this strain, the genes encoding thioredoxin reductase and glutathione reductase are disrupted for generating the oxidative cytoplasm [16].
Based on comparison of the mature HRP amino acid sequence (Fig. S1), the sequence encoding the mature mPex Q45-S350 with deletion of the hydrophobic N-terminal leader sequence responsible for directing the protein to the endoplasmic reticulum (ER) was amplified by RT-PCR using the total RNAs extracted from maize leaves as the template, and primers mPex1 and mPex2. The PCR amplicon was incubated with *Bam*H I and *Xho* I, and placed between the *Bam*H I-*Xho* I site of pCBM-tevS-GFP to replace the GFP sequence. The similar substitution using the *Bam*H I-*Xho* I excised bEK coding sequence from the pEK plasmid was conducted for expressing the CBM tagged bEK. The linker between the CBM and the tevS was introduced by amplification of the CBM tag using the primers CBM1 and CBM2, digestion with *Nde* I and *Bam*H I and insertion into *Nde* I/*Bgl* II sites of the pCBM-tevS-GFP plasmids. The flexible linker between the tevS and the bEK was introduced as the similar procedure by using the primers beK1 and beK2 for amplification of the bEK coding sequence to prolong sequence upstream of the original *Bam*H I site encoding the extra six-amino-acid, and excising with *Bgl* II and *Xho* I for introduction of the *Bam*H I/*Xho* I sites of the former constructed plasmids. The *Bam*H I-*Xho* I excised fragment encoding bEK was then substituted with that encoding the mPex.

For constructing the redox sensitive GFP reporter roGFP [17], the sequence encoding the GFP variant mGFP5 with improved stability in *E. coli* [18], was amplified using the plant expression vector pCAMBIA1302 as the template and primers mGFP5-1 and mGFP5. The purified PCR amplicon with the *Nco* I and *Hind* III was subcloned into the *Nco* I-*Hind* III site of the pET-28b plasmid. The mutations C48S, S147C and Q204C were introduced into the mGFP5 step by step by using the primers for PCR. All inserts were sequenced to identify correction. The primers used in this study were listed in the table S1.

**Solubility determination of the overexpressed TEVp constructs in different *E. coli* strains**

Except where noted, induction and extraction of recombinant proteins in this study were conducted as follows. The constructed plasmids were transformed independently into the different *E. coli* strains. The recombinant cells were cultured overnight at 37 °C in 5 ml of Luria-Bertani (LB) broth, diluted to 50-fold and grown at 37°C. Induction of the target protein was performed by addition of isopropylthio-β-D-galactoside (IPTG) at final concentration 0.5 mM, when OD₆₀₀ value reached about 0.5. After cultured at 28 °C for 12 h in 10 ml liquid culture of a 50-ml shake flask, cells were collected by centrifugation, washed with buffer A (20 mM Tris-HCl, pH 8.0, 100 mM NaCl), sonicated and centrifuged, and analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The protein amounts were measured by Bradford method, using bovine serum albumin as the reference. After electrophoresis, proteins on the gel were transferred to polyvinylidene fluoride membrane, immunoblotted with anti-His6 monoclonal antibodies, and treated with HRP conjugated anti-rabbit IgG. The band representing the target protein was visualized by addition of 4-chloro-1-naphthol solution dissolved in 20% methanol and 0.08% H₂O₂.

The C-terminally fused EmGFP reporter is used for quantitative analysis of the soluble TEVp amounts, based on the fluorescence emitted from soluble fractions on the F-4500 fluorescence spectrometer (Hitachi, Japan) with excitation and emission maximum 488 and 515 nm [13].
Purification of the constructed TEVp with C110S and/or C130S mutations

The expression plasmids were transformed into the Rosetta (DE3) strain with supply rare tRNAs for augmenting soluble expression level of the TEVp construct. After induction at 28 °C for 12 h, cells were collected by centrifugation and washed with buffer B (50 mM sodium phosphate, 300 mM NaCl, and 10 mM imidazole, pH 8.0), and disrupted by sonication. The supernatants were loaded on a Ni-NTA spin, centrifuged twice with buffer B, washed twice with 40 mM imidazole in buffer B (pH 8.0), and eluted with 250 mM imidazole in buffer B (pH 8.0). Purified protein was concentrated by an Ultra-15 centrifugal filter tube equipped with an Ultracel-10 membrane, and exchanged with buffer A. The freshly prepared protein was used for cysteine labeling and activity assay.

Modification of the free cysteine residues in the purified TEVp constructs

The free cysteine residues of the TEVp constructs were labeled with AMS as the described method [19]. Purified TEVp variants was incubated with 150 µM CuCl₂ as the oxidative agent or 1 mM dithiothreitol (DTT) for 1 h at 25 °C, precipitated by trichloroacetic acid. After it was washed and re-suspended with buffer A twice, the precipitated protein was labeled with AMS, a maleimidyl reagent specifically alkylating free SH-group of cysteine to increases the molecular weight in about 0.5 kDa. When the reaction was finished, the mixture was centrifuged and washed with buffer A. The labeled TEVp variants were incubated with loading buffer in absence of DTT, and separated by 12% SDS-PAGE under non-reducing condition.

Quantitative analysis of in vitro activity

In vitro cleavage activity was assayed using the GST-eDAL purified by Ni-NTA [13]. The mass ratio was 30:1 for the purified protein substrate and soluble extracts containing the recombinant TEVp construct, and 50:1 for purified protein substrate and the protease in the buffer A. The cleavage was reacted at 30 °C for 1 h, and the activity was determined by coupled assay of the eDAL activity [11]. As a PLP dependent enzyme, eDAL catalyzes L-DAP to pyruvate and ammonia. The amount of pyruvate was measured with 2,4-DNP. The reaction mixture for testing DAL activity contained 50 µM PLP and 10 mM L-DAP, and the crude extracts in a final volume of 1 ml. After incubated at 37 °C for 5 min, 1 ml of 2 mM HCl plus 0.03% 2,4-DNP was added to stop the DAL catalytic reaction. Following incubation at 4 °C for 5 min, 2 ml of 2 M NaOH was supplemented. After centrifugation, absorbance at 520 nm representing the amounts of pyruvate in the mixture was measured on a U-2001 spectrometer (Hitachi, Japan).

Recovery of inclusion bodies

The induced cells were re-suspended in buffer A and disrupted by sonication. After centrifugation at 4000 g for 10 min, insoluble fraction was collected, washed with buffer A twice and re-suspended with buffer B [30 mM Tris/HCl, 150 mM NaCl, 10% (v/v) glycerol, 0.5% (v/v) Triton X-100, pH 7.5]. The mixture was centrifuged and IBs were re-suspended with buffer B in the absence of Triton X-100.
The washed IBs containing the bEK was solubilized in buffer A containing 8 M urea, 5 mM DTT (V/V) and 10 mM EDTA [20]. The mixture was left for 2 h at room temperature to ensure sufficient amounts of IBs for solubilization and reduction of the mismatched disulfide bonds, and centrifuged at 18000 g for 30 min to remove the pellet. IBs containing the mPex dissolved in the solution (4.5 M urea, 40 mM Tris-HCl, pH 9.0, 5 mM DTT and 0.2 mM hemin) to a protein concentration of 0.3 mg/mL, with slight modification [21]. The resultant suspension was centrifuged at 12000 g for 20 min and the debris was discarded. The solubilized protein was used for refolding.

Matrix-assisted refolding of target proteins

For increasing the cellulose loading yields of target proteins, the regenerated amorphous cellulose (RAC) was prepared based on the published paper [15]. The microcrystalline cellulose was suspended with double-distilled water at ratio of 1:3 (W/V), and slowly added 20 fold volume of 10 ml ice-cold H₃PO₄ with vigorous stirring, than supplemented 2 fold volume of ice-cold water. After centrifugation, the pellet was suspended with ice-cold water several times for removing phosphoric acid in soluble fraction, and centrifuged. With adding 2 M Na₂CO₃ to the pellet, the mixture was centrifuged and discarded. After the resin was washed with ice-cold water several times until the pH value in mixture containing RAC was reached about 7.0. The solubilized proteins from IBs were mixed with RAC at mass ratio of about 1:20 to make the urea concentration decreased from 8 M to 6 M. The mixture was stirred for 1 h at room temperature to allow the target protein binding RAC.

For the bEK refolding [20], the protein was diluted with buffer C (100 mM Tris–HCl, 6 M urea, 10 mM cystine, pH 8.0) to facilitate correct formation of disulfide bonds. Then, the resin was put in, mixed in buffer C containing 3 M urea 5% glycerol (V/V) and 10 mM cysteine, and diluted with buffer C containing 0.7 M urea, 15% (v/v) glycerol, 0.5 mM cysteine-HCl, 5 mM cysteine and 2mM CaCl₂. For the mPex refolding with slight modification [21], the denatured proteins were diluted with buffer D [20 mM Tris-HCl, pH 8.5, containing 0.5 M urea, 5% glycerol (V/V), 2 µM hemin, 2 mM CaCl₂, 0.5 mM GSH and 5 mM GSSG] in the presence of RAC.

Purification and activity assay of the refolded tag-free proteins by the TEVp variant digestion

The purified TEVp variant was incubated with the refolded protein bound to the RAC with mass ratio of 1:10 at 10 °C for 24 h. After reaction was finished, the Ni-NTA-resin was added and incubated for 2 h at room temperature. The supernatant after centrifugation was subjected to SDS-PAGE analysis. The bEK cleaving the GST tagged sDAL with incorporation of D4K as the bEK recognition sequence was analyzed, based on the coupled assay of sDAL activity. The mPex catalyzes the degradation of H₂O₂ using OPA as a hydrogen donor, which turns yellow upon oxidation [22]. The freshly prepared mPex was incubated in the mixture (20 mM Tris/HCl, pH 7.5, 50 µg/mL OPA, 10 mM or 30 mM H₂O₂) at 37 °C for 30 min, and absorption at 411 nm was measured.

Results
Site-directed mutations and plasmids construction

To improve the TEVp variant to prohibit obvious loss of activity under the oxidative conditions, we mutated one to three cysteine residues, including C19, C110 and C130 in the TEVp<sup>5M</sup> variant were mutated to serine ones. Three variants were constructed including M1-M3 with one cysteine residue replaced with serine one. The variants M4–M6 contained the combined two mutations in M1-M3, and M7 contain all three mutations (Tab. 1).

Production of the TEVp variants in different <i>E. coli</i> strains

The constructed seven variants in soluble and insoluble fractions from <i>E. coli</i> BL21 (DE3) cells were varied, as shown on the SDS-PAGE gel and by Western blotting (Fig. 1a). Four variants including the M1, M4, M5 and M7 containing the C19S mutation displayed poor solubility. In contrast, other three variants in soluble fractions were detected. Inhibition of the background expression in the BL21(DE3) pLysS strain decreased production of all variants (Fig. 1b). With augment of rare tRNA levels in the Rossetta (DE3) strain, soluble production of all variants were not significantly increased (Fig. 1c). Since C19S mutation decreased protein solubility, we expressed other four TEVp variants in the Origami (DE3). Except for M7, three certain variants displayed the enhanced soluble production (Fig. 1d). The results indicated that the C110S and/or C130S mutations conferred the TEVp<sup>5M</sup> with augment of protein solubility, different from the C19 mutation, and suggested that both mutations failed to recover the C19 mutation on impairing the TEVp<sup>5M</sup> folding.

Quantitatively analyzing the soluble amounts of the TEVp constructs

Using the C-terminal EmGFP reporter, soluble production of the TEVp variants were quantitatively analyzed. Introducing C110S and C130S mutations to the TEVp<sup>5M</sup> rendered the protease variant with different soluble productivity in various strains, but higher soluble amounts in the Origami (DE3) strain than the TEVp<sup>5M</sup> (Figs. 2a-2d). Soluble production of the fusion constructs were also detected by SDS-PAGE (Figs. S2, a-d). The current data reflected that soluble productivity of the various TEVp variants were dependent with the <i>E. coli</i> strains used in this study. Combination of the C110S and C130S further enhanced soluble production of the TEVp<sup>5M</sup> in the Origami (DE3).

Cleavage activity of the TEVp variants

The purified GST-tagged eDAL was cleaved into two parts by TEVp<sup>5M</sup>, M2, M3 and M6 variants in the crude extracts, whereas the other variants had less cleavage efficiency, as detected by SDS-PAGE (Figs. S3, a-d). The changed cleavage activity among the TEVp variants was probably attributed to soluble amounts. By the coupled assay, the M2 and M3 variants produced in the BL21(DE3) cells showed higher cleavage activity than the TEVp<sup>5M</sup>, but two variants produced in the recombinant BL21(DE3)pLysS and Rosetta (DE3) cells had less efficiency on cleaving the protein substrate (Figs. 3a-3d). The M6 variant from the Origami (DE3) exhibited higher activity, but the M3 variant displayed almost equal activity to the
TEVp\(^{5M}\). The results suggested that mutations of the C110 and C130 increased soluble production and cleavage activity of the TEVp variants in the oxidative environment, probably resulted from the avoidance of the mis-matched disulfide bond(s) formation.

**Addition of DTT or H\(_2\)O\(_2\) to the culture affecting the enzymatic activity**

When DTT was added to the culture at final concentration of 2 mM or 10 mM, the TEVp\(^{5M}\) or M2, M3 or M6 produced in the BL21(DE3) displayed the increased specific activity (Fig. 4a). In contrast, addition of 2 mM or 10 mM (final concentration) H\(_2\)O\(_2\) to the LB medium decreased the TEVp activity (Fig. 4b). Supply of reducing or oxidative agent changed the cellular redox in the BL21(DE3) strain, based on the changed excitation at 485 emitted by the produced roGFP1 (Fig. S4). Upon reduction with DTT, the ratio of cell fluorescence excited at 485 and 400 nm (F485/F400) presenting the redox potential in the cells was also increased, different from the cells treated with H\(_2\)O\(_2\) (Fig. 4c). The reducing environment affected the enzymatic activity for TEVp\(^{5M}\) more efficiently than the M2 and M6 variants. Similarly, M2, M3 and M6 mutants displayed less decrease in cleavage activity than the TEVp\(^{5M}\) with supplementary H\(_2\)O\(_2\) to the medium, resulted from probability that the cellular redox affected the TEVp folding, and the mis-matched disulfide bond formation.

**Free cysteine residues labelling in the purified TEVp constructs**

The TEVp variants were independently expressed in the Rosetta (DE3) with boosting rare tRNAs levels for enhancing soluble production. The main bands were observed on the SDS-PAGE gel for purified TEVp constructs under reducing condition (Fig. 5a). The AMS-labelled TEVp\(^{5M}\) exhibited several bands in non-reducing SDS-PAGE (Fig. 5b), resulted from incubation with CuCl\(_2\) facilitating disulfide bond formation, and prohibiting AMS labeling. In contrast, the M3 and M6 variant showed less bands than the TEVp\(^{5M}\), due to one and/or two cysteine residues mutation. Modification of the M2 variant was not efficient, and the faint bands were observed. The reason is not known. Under reducing condition, the TEVp\(^{5M}\) displayed the similar bands to the other mutants (Fig. 5b). Presence of DTT caused slower mobility of the labeled TEVp, suggesting that more AMS molecules binding the TEVp upon disruption of the disulfide bonds.

**Effect of reducing reagent or oxidative agents on the activity of purified TEVp constructs**

The purified TEVp construct cleaved the fusion protein substrate (Fig. 6a). The specific activity of the constructed TEVp was comparable to the TEVp\(^{5M}\), except for the M3 showing the slightly decreased activity (Fig. 6b). With addition of 2 mM DTT, activity of the construct increase was increased trivially (Fig. 6c). On the contrary, addition of compound pairs including 5 mM cystine and 0.5 mM cysteine, or 5 mM GSSG and 0.5 mM GSH for facilitating disulfide bridge formation during the protein renaturation slightly inhibited the cleavage activity of purified TEVp constructs. The activity of the TEVp\(^{5M}\)C110S/C130S variant was increased by about 20%, in contrast to that of the TEVp\(^{5M}\) (Fig. 6c). The data suggested that the engineered TEVp variant was more insensitive to the reagents for changing the redox in the reaction mixture.
Removal of fusion tags in two refolded proteins containing multiple disulfide bonds

The CBM attached bEK or Mper was produced mainly as IBs (Fig. 7a). Initial attempt for releasing the target proteins from RAC was failed. So, we introduced the flexible linker in the fusion protein for enhancing the cleavage efficiency. After refolded, the target protein was detached from the resin using the purified M6 variant (Fig. 7b). The refolded bEK cleaved the GST tagged sDAL into two parts, and cleavage activity was assayed based on the released sDAL catalysis (Fig. 7c). With addition the refolded mPex, the H_2O_2 at 10 and 30 mM concentrations was transformed into yellow compounds in the presence of OPA, in contrast to the heat-inactive enzyme (Fig. 7d). It is noted that auto-oxidation was reacted spontaneously in the mixture containing 30 mM H_2O_2 as the control, based on the color change. The results indicated that renaturation of the fusion protein for tag-removal to release the tag-free active bEK and Mper was efficient.

Discussion

In this study, we engineered one to three cysteine residues in the TEVp^5M to select the variant with the improved activity in E. coli oxidative cytoplasm, and identified the various production levels of the recombinant constructs. So far, different GFP variants have been used for monitoring protein folding and solubility in E. coli. A linear correlation between in vivo fluorescence and soluble amounts of the engineered target protein variants is yielded by use of the fused enhanced GFP (eGFP) variant [23–25]. However, the correlation changes significantly for various GFP variants [26]. Certain target proteins affect the GFP folding [27], thus, a robust folding variant of the GFP, called superfolder GFP (SfGFP), is generated [28]. In addition, the metabolic burden of the expressed GFP impacts the reporter efficiency [29].

By use of the GFP reporter, the fluorescence intensities were changed irregularly for the different TEVp constructs among the selected E. coli strains. The expressed EmGFP affords the E. coli cell fluorescence, color intensity and temperature tolerance more effectively than the eGFP [30]. So, irregular alternation of fluorescence from soluble fractions in different E. coli strains is probably caused by the GFP toxicity. The GFP fluorophore biosynthesis is accompanied by the generation of H_2O_2 [31]. High level production of the TEVp^5M fused to the EmGFP at 37°C induction for over 4 h inhibits E. coli cell growth [12]. The P_T7 promoter regulating the GFP variant expression at high level decreased cell fluorescence after induction at 37°C induction for over 10 h [32]. Besides, the GFP produced as insoluble aggregates emits fluorescence [29]. So, fluorescence signal emitted from soluble fractions, not from cells, showed improved correlation with the expressed soluble protein amounts. Our current data provided the suggestion that desirable expression level and cellular environment are required for yielding the linear correlation between fluorescence and yields of the recombinant target protein using the GFP reporter.

Based on the crystal structure, the C19, C110 and C130 residues are located in the first, eighth, and tenth β-strands, respectively, and the C130 forms a disulfide bridge between TEVp molecules or reacts with β-mercaptoethanol [33, 34]. Different from the inactive TEVp construct in the eukaryotic cells [3],
introduction of the C110S had little impact on protein solubility and activity. The C19 and C110 are buried in the TEVp molecule, the C19A mutation on the protein solubility is needed to be further assessed. Mutations of N23Q, C130S, and T173G introduced in the wild type TEVp for inhibiting N-glycosylation the improving resistance oxidative stress in the ER of mammalian cells afford the active protease, but addition of the C110S inactivates the protease [3]. In this work, we used the TEVp variant improved protein folding for mutation of the cysteine residues. It is proposed that the naturally occurring disulfide bridges formation is guided by inherent protein folding [35, 36], but little information is available on inherent protein folding resisting formation of the error disulfide bridge(s). Fluorescence change of the GFP variants via secretion to the E. coli periplasm provides the indirect evidence. The eGFP forms non-fluorescent oligomer via the mixed disulfides in the ER, leading to loss of fluorescence in E. coli periplasm by the SecYEG transportation pathway. In contrast, the SfGFP is brightly fluorescent in the ER and E. coli periplasm [37].

The artificial two cysteine residues in the roGFP1 are generated for forming a disulfide bond, and this reporter is effective for monitoring the intracellular redox in E. coli [38]. Using the SfGFP as a model, the roSfGFP is created to increase cell fluorescence, but certain characters of roSfGFP are faintly different from those of the roGFP1 [39, 40]. Considering that protein folding contributes the redox sensitive reporter function, here, we used the mGFP5, not the EmGFP for construction of the roGFP. The error disulfide bond formation in the NpuDnaE intein produced in E. coli is also detected by non-reducing SDS-PAGE, resulting in decline in the cleavage under non-reducing conditions [41]. In this study, we identified that the added DTT to the culture augmented the TEVp activity. This method is simple and cost-effective for increasing soluble production and activity of the other target proteins in E. coli [42, 43].

Oriented immobilization of the CBM tag minimizes the contact between target protein molecules, thus preventing their aggregation. The advantage of the CBM tag used for protein refolding is that, even in the presence of 6 M urea, the CBM retains the binding capability [11]. Additionally, this method is economical and time-saving. The target proteins were liberated from the resin after the TEVp variant cleavage. Moreover, the refolded CBM tagged two enzymes bound to the RAC were used as immobilizates for improving use efficiency and easy extraction from the reaction mixture.

In summary, the multiple mutations afforded the engineered TEVp variant with the enhanced soluble yield, and the decreased self-inactivation, and the improved oxidative stability. This variant is used for removing the fusion partner in the oxidative hyperoxic E. coli strain and in vitro. Removal of the CBM tag with the TEVp variant mediated release of the refolded disulfide-rich proteins from the RAC resin in the refolding buffer.

**Declarations**

**Funding**

This study was financially supported by the Scientific and Technological Major Project of Anhui Province (1803071180).
Author information

Affiliations

School of Life Science, Anhui Agricultural University, Hefei, Anhui, 230036, China.

Enkhtuya Bayar, Yuanyuan Ren, Yafang Hu, Weiyu Wang, Yinghua Chen, Jong Hu, Xuelian Yu, Jingfeng Li, and Jun Fan

Corresponding author

Correspondence to Jun Fan.

Author contributions

Conceived, designed the experiments and wrote the paper: Jun Fan. Performed the experiments: Enkhtuya Bayar, Yuanyuan Ren, Yafang Hu, Weiyu Wang, Yinghua Chen, Jong Hu, Xuelian Yu, and Jingfeng Li. Analyzed the data: Yuanyuan Ren, Yafang Hu, Yinghua Chen, Jong Hu and Xuelian Yu.

Ethics declarations

Conflicts of Interest

All authors declare that they have no conflict of interest.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors. Hence, no formal consent is required.

Supplementary data

Supplementary material related to this article can be found, in the online version.

References

1. Cesaratto, F., Burrone, O. R., & Petris, G. (2016). Tobacco etch virus protease: A shortcut across biotechnologies. Journal of Biotechnology, 231, 239–249.
2. Asial, I., Cheng, Y. X., Engman, H., Dollhopf, M., Wu, B., Nordlund, P., & Cornvik, T. (2013). Engineering protein thermostability using a generic activity-independent biophysical screen inside the cell. Nature Communication, 4, 2901.
3. Cesaratto, F., López-Requena, A., Burrone, O. R., & Petris, G. (2015). Engineered tobacco etch virus (TEV) protease active in the secretory pathway of mammalian cells. Journal of Biotechnology, 212, 159–166.
4. Shafee, T., Gatti-Lafranconi, P., Minter, R., & Hollfelder, F. (2015). Handicap-recover evolution leads to a chemically versatile, nucleophile-permissive protease. *Chembiochem: a European journal of chemical biology, 16*, 1866–1869.

5. Packer, M. S., Rees, H. A., & Liu, D. R. (2017). Phage-assisted continuous evolution of proteases with altered substrate specificity. *Nature Communication, 8*, 956.

6. Sanchez, M. I., & Ting, A. Y. (2020). Directed evolution improves the catalytic efficiency of TEV protease. *Nature Methods, 17*, 167–174.

7. Gopal, G. J., & Kumar, A. (2013). Strategies for the production of recombinant protein in *Escherichia coli*. *Protein Journal, 32*, 419–425.

8. Yamaguchi, H., & Miyazaki, M. (2014). Refolding techniques for recovering biologically active recombinant proteins from inclusion bodies. *Biomolecules, 4*, 235–251.

9. Hwang, P. M., Pan, J. S., & Sykes, B. D. (2014). Targeted expression, purification, and cleavage of fusion proteins from inclusion bodies in *Escherichia coli*. *FEBS Letters, 588*, 247–252.

10. Jungbauer, A., Kaar, W., & Schlegl, R. (2004). Folding and refolding of proteins in chromatographic beds. *Current opinion in biotechnology, 15*, 487–494.

11. Berdichevsky, Y., Lamed, R., Frenkel, D., Gophna, U., Bayer, E. A., Yaron, S., Shoham, Y., & Benhar, I. (1999). Matrix-assisted refolding of single-chain Fv-cellulose binding domain fusion proteins. *Protein Expression and Purification, 17*, 249–259.

12. Wei, L., Cai, X., Qi, Z., Rong, L., Cheng, B., & Fan, J. (2012). In vivo and in vitro characterization of TEV protease mutants. *Protein Expression and Purification, 83*, 157–163.

13. Zhou, C., Yan, Y., Fang, J., Cheng, B., & Fan, J. (2014). A new fusion protein platform for quantitatively measuring activity of multiple proteases. *Microbial Cell Factories, 13*, 44.

14. Tan, H., Wang, J., & Zhao, Z. K. (2007). Purification and refolding optimization of recombinant bovine enterokinase light chain overexpressed in *Escherichia coli*. *Protein Expression and Purification, 56*, 40–47.

15. Yu, X., Sun, J., Wang, W., Jiang, L., Wang, R., Xiao, W., Cheng, B., & Fan, J. (2017). Tobacco etch virus protease mediating cleavage of the cellulose-binding module tagged colored proteins immobilized on the regenerated amorphous cellulose. *Bioprocess and Biosystems Engineering, 40*, 1101–1110.

16. Bessette, P. H., Aslund, F., Beckwith, J., & Georgiou, G. (1999). Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. *Proceedings of the National Academy of Sciences of the United States of America, 96*, 13703–13708.

17. Hanson, G. T., Aggeler, R., Oglesbee, D., Cannon, M., Capaldi, R. A., Tsien, R. Y., & Remington, S. J. (2004). Investigating mitochondrial redox potential with redox-sensitive green fluorescent protein indicators. *Journal of Biological Chemistry, 279*, 13044–13053.

18. Siemering, K. R., Golbik, R., Sever, R., & Haseloff, J. (1996). Mutations that suppress the thermosensitivity of green fluorescent protein. *Current Biology, 6*, 1653–1663.
19. Ikegami, A., Yoshimura, N., Motohashi, K., Takahashi, S., Romano, P. G., Hisabori, T., Takamiya, K., & Masuda, T. (2007). The CHLI1 subunit of Arabidopsis thaliana magnesium chelatase is a target protein of the chloroplast thioredoxin. Journal of Biological Chemistry, 282, 19282–19291.

20. Skala, W., Goettig, P., & Brandstetter, H. (2013). Do-it-yourself histidine-tagged bovine enterokinase: a handy member of the protein engineer's toolbox. Journal of Biotechnology, 168, 421–425.

21. Levy, I., Ward, G., Hadar, Y., Shoseyov, O., & Dosoretz, C. G. (2003). Oxidation of 4-bromophenol by the recombinant fused protein cellulose-binding domain-horseradish peroxidase immobilized on cellulose. Biotechnology and Bioengineering, 82, 223–231.

22. Tao, Y. M., Wang, S., Luo, H. L., & Yan, W. W. (2018). Peroxidase from jackfruit: Purification, characterization and thermal inactivation. International Journal of Biological Macromolecules, 114, 898–905.

23. Seitz, T., Thoma, R., Schoch, G. A., Stihle, M., Benz, J., D'Arcy, B., Wiget, A., Ruf, A., Hennig, M., & Sterner, R. (2010). Enhancing the stability and solubility of the glucocorticoid receptor ligand-binding domain by high-throughput library screening. Journal of Molecular Biology, 403, 562–577.

24. Oelschlaeger, P., Srikan-t-Iyer, S., Lange, S., Schmitt, J., & Schmid, R. D. (2002). Fluorophor-linked immunosorbent assay: a time- and cost-saving method for the characterization of antibody fragments using a fusion protein of a single-chain antibody fragment and enhanced green fluorescent protein. Analytical biochemistry, 309, 27–34.

25. Hedhammar, M., Stenvall, M., Lönneborg, R., Nord, O., Sjölin, O., Brismar, H., Uhlén, M., Ottosson, J., & Hofer, S. (2005). A novel flow cytometry-based method for analysis of expression levels in Escherichia coli, giving information about precipitated and soluble protein. Journal of Biotechnology, 119, 133–146.

26. Gupta, R. D., & Tawk, D. S. (2008). Directed enzyme evolution via small and effective neutral drift libraries. Nature Methods, 5, 939–942.

27. Waldo, G. S., Standish, B. M., Berendzen, J., & Terwilliger, T. C. (1999). Rapid protein-folding assay using green fluorescent protein. Nature Biotechnology, 17, 691–695.

28. Pédelacq, J. D., Cabantous, S., Tran, T., Terwilliger, T. C., & Waldo, G. S. (2006). Engineering and characterization of a superfolder green fluorescent protein. Nature Biotechnology, 24, 79–88.

29. Cha, H. J., Wu, C. F., Valdes, J. J., Rao, G., & Bentley, W. E. (2000). Observations of green fluorescent protein as a fusion partner in genetically engineered Escherichia coli: monitoring protein expression and solubility. Biotechnology and Bioengineering, 67, 565–674.

30. Teerawanichpan, P., Hoffman, T., Ashe, P., Datla, R., & Selvaraj, G. (2007). Investigations of combinations of mutations in the jellyfish green fluorescent protein (GFP) that afford brighter fluorescence, and use of a version (VisGreen) in plant, bacterial, and animal cells. Biochimica et Biophysica Acta, 1770, 1360–1368.

31. Zhang, L., Patel, H. N., Lappe, J. W., & Wachter, R. M. (2006). Reaction progress of chromophore biogenesis in green fluorescent protein. Journal of the American Chemical Society, 128, 4766–4772.
32. Lee, S. K., & Keasling, J. D. (2006). Propionate-regulated high-yield protein production in. *Escherichia coli*. *Biotechnology and Bioengineering*, 93, 912–918.

33. Phan, J., Zdanov, A., Evdokimov, A. G., Tropea, J. E., Peters, H. K. 3rd, Kapust, R. B., Li, M., Wlodawer, A., & Waugh, D. S. (2002). Structural basis for the substrate specificity of tobacco etch virus protease. *Journal of Biological Chemistry*, 277, 50564–50572.

34. Nunn, C. M., Jeeves, M., Cliff, M. J., Urquhart, G. T., George, R. R., Chao, L. H., & Tscuchia, Djordjevic, S. (2005). Crystal structure of tobacco etch virus protease shows the protein C terminus bound within the active site. *Journal of Molecular Biology*, 350, 145–155.

35. Kosuri, P., Alegre-Cebollada, J., Feng, J., Kaplan, A., Inglés-Prieto, A., Badilla, C. L., Stockwell, B. R., Sanchez-Ruiz, J. M., Holmgren, A., & Fernández, J. M. (2012). Protein folding drives disulfide formation. *Cell*, 151, 794–806.

36. Qin, M., Wang, W., & Thirumalai, D. (2015). Protein folding guides disulfide bond formation. *Proceedings of the National Academy of Sciences of the United States of America*, 112, 11241–11246.

37. Aronson, D. E., Costantini, L. M., & Snapp, E. L. (2011). Superfolder GFP is fluorescent in oxidizing environments when targeted via the Sec translocon. *Traffic (Copenhagen, Denmark)*, 12, 543–548.

38. Zhang, W., Zheng, W., Mao, M., & Yang, Y. (2014). Highly efficient folding of multi-disulfide proteins in superoxidizing *Escherichia coli* cytoplasm. *Biotechnology and Bioengineering*, 111, 2520–2527.

39. Hoseki, J., Oishi, A., Fujimura, T., & Sakai, Y. (2016). Development of a stable ERroGFP variant suitable for monitoring redox dynamics in the ER. *Bioscience Reports.*, 36, e00316.

40. Schuh, A. K., Rahbari, M., Heimsch, K. C., Mohring, F., Gabryszewski, S. J., Weder, S., Buchholz, K., Rahlfs, S., Fidock, D. A., & Becker, K. (2018). Stable integration and comparison of hGrx1-roGFP2 and sfroGFP2 redox probes in the malaria parasite *Plasmodium falciparum*. *ACS Infectious Diseases*, 4, 1601–1612.

41. Xu, Y., Zhang, L., Ma, B., Hu, L., Lu, H., Dou, T., Chen, J., & Zhu, J. (2018). Intermolecular disulfide bonds between unpaired cysteines retard the C-terminal trans-cleavage of Npu DnaE. *Enzyme and Microbial Technology*, 118, 6–12.

42. Gill, R. T., Cha, H. J., Jain, A., Rao, G., & Bentley, W. E. (1998). Generating controlled reducing environments in aerobic recombinant *Escherichia coli* fermentations: effects on cell growth, oxygen uptake, heat shock protein expression, and in vivo CAT activity. *Biotechnology and Bioengineering*, 59, 248–259.

43. Siurkus, J., & Neubauer, P. (2011). Reducing conditions are the key for efficient production of active ribonuclease inhibitor in *Escherichia coli*. *Microbial Cell Factories*, 10, 31.

**Table**

| Table 1. The variants constructed in this study |
Figures

Figure 1

Analysis of overexpression of the double His6-tagged TEVp variants in the BL21(DE3) (a), BL21(DE3)pLysS (b), Rosetta (DE3) (c) and Origami (DE3) (d) strains. About 10 μg of soluble proteins, and 4 μg of proteins in pellets were separated by SDS-PAGE. In the figure 1a, the specific bands detected by Western blotting were displayed on bottom of the gel. M: protein molecular weight marker. S: soluble. P: pellet. Variant names are denoted above the gel. On each gel, the arrow indicated positions of TEVp5M and the different variants.
Figure 2

Soluble expression levels of the TEVp constructs in the BL21(DE3) (a), BL21(DE3)pLysS (b), Rosetta (DE3) (c) and Origami (DE3) (d) strains using the GFP reporter. Fluorescence from the cells harboring the empty plasmid was used as the control and subtracted. Data are means and standard deviation of three biological replicates. At least five samples were measured under the same induction conditions, and the closest values were counted.
Figure 3

The coupled assay of cleavage activity of the TEVp5M and its variants in four E. coli strains in the BL21(DE3) (a), BL21(DE3)pLysS (b), Rosetta (DE3) (c) and Origami (DE3) (d). Data are means and standard deviation of three technical replicates. The heat-inactivated TEVp construct was used as control, and the absorption was subtracted.
Figure 4

Analysis of the specific activity of the TEVp construct overexpressed in E. coli BL21(DE3) with adding DTT or H2O2 to the culture upon induction. (a) The TEVp5M or TEVp5MC110S or the TEVp5MC110S/C130S construct was extracted and cleavage activity was assayed with DTT addition to the culture at final concentrations of 2 mM or 10 mM. (b) Hydrogen peroxide addition to culture at 2 mM and 10 mM and the activity of the extracted TEVp was examined. (c) The relative ratio of roGFP fluorescence excited at 400 and 485 nm (F485/F400) from the BL21(DE3) cells treated with oxidant and reductant. The F485/F400 value was divided by the OD600 value. 1: control; 2: 2 mM H2O2; 3: 10 mM H2O2; 4: 2 mM DTT; 5: 10 mM DTT. Data are means and standard deviation of three technical replicates.

Figure 5

Labelling the TEVp constructs with AMS. (a) SDS-PAGE analysis of purified constructs under reducing condition. TEVp variant names were denoted above the gel. The arrow indicated positions of purified proteins. (b) The TEVp constructs treated with CuCl2 as the oxidative regent or DTT as the reducing agent, labeled with AMS, mixed with loading buffer in absence of the DTT, and subjected to SDS-PAGE analysis at the same amounts. All analyses were described in section of materials and methods.
**Figure 6**

In vitro analysis of the cleavage activity of the purified TEVp variants. (a) SDS-PAGE analysis of the fusion protein after TEVp cleavage. Variant names are denoted above the gel. The arrow indicated fusion protein and the cleaved products. (b) Specific activity of the purified TEVp constructs by the coupled assay. Data are means and standard deviation of three technical replicates. (c) The cleavage activity of purified TEVp constructs in the presence of 5 mM DTT, or 5 mM cystine plus 0.5 mM cysteine, or 5 mM GSSG plus 0.5 mM GSH.
Refolding the CBM tagged proteins and removal of the CBM tag. (a) SDS-PAGE analysis of the two CBM tagged proteins produced in E. coli BL21(DE3) strain. T: total protein. S: soluble fraction. P: Insoluble fraction. (b) SDS-PAGE analysis of the released target proteins after refolding from the RAC resins via CBM tag binding after incubation of the purified M6 variant. Lane 1: the released mPex. Lane 2: the released bEK. (c) Activity assay of the released bEK. Left: Cleavage of the GST-D4K-sDAL with the
refolded bEK. Lane C: the fusion protein incubated with the heat-inactive bEK. Lane 2: the fusion protein incubated with the refolded bEK. Middle: color display of the released sDAL for catalyzing the DL-DAP to generate pyruvate. The heat-inactive bEK was used as the control. The colored products represented the non-tagged sDAL catalysis. Right: The sDAL activity of purified GST-D4K-sDAL before and after cleavage reaction as indicated in the grey and black columns. (d) Activity of the released mPex. Left: The colored products representing the mPex catalysis. C: The heat-inactive mPerx (100 °C for 10 min) was added into the reaction mixture. 1: reaction mixture contained 10 mM H2O2. 2: reaction mixture contained 30 mM H2O2. Right: the measured activity of the mPex based on catalysis of 10 mM and 30 mM H2O2 as indicated in the grey and black columns. Data representing means and standard deviation were from three technical replicates.

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarymaterial.docx](#)