# Exploring the Mutations on Improving Oxidative Stability of Tobacco Etch Virus Protease for Tag-removal in Refolding of Two Disulfide-rich Proteins

**Jun Fan** ( fanjun@ahau.edu.cn )
Anhui Agricultural University
**Enkhtuya Bayar**
Anhui Agricultural University
**Yuanyuan Ren**
Anhui Agricultural University
**Yafang Hu**
Anhui Agricultural University
**Yinghua Chen**
Anhui Agricultural University
**Xuelian Yu**
Anhui Agricultural University

**Research Article**

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

**Posted Date:** May 11th, 2021

**DOI:** https://doi.org/10.21203/rs.3.rs-473834/v1

# Abstract

Tobacco etch virus protease (TEVp) is a useful tool for removing fusion tag, but wild type TEVp shows less oxidative stability, which limits its application under the oxidized redox state to facilitate disulfide bonds formation for refolding disulfide-bonded proteins. Previously, we combined six mutations into the TEVp to generate the TEVp$^{5M}$ for obviously increasing the protein solubility and decreasing the auto-cleavage. In this work, we introduced and combined C19S, C110S and C130S mutations into the TEVp$^{5M}$ to generate seven variants, analyzed protein solubility and the cleavage activity of the constructs in each of three *E. coli* strains including BL21(DE3), BL21(DE3)pLys, and Rossetta(DE3), and those of the optimized soluble variants in the oxidative cytoplasm of Origami(DE3) under the same induction conditions. The results suggested that desirable protein solubility, cleavage activity and oxidative stability are not combined. Unlike that of the C19S, introduction of the C110S and/or C130S less affected protein solubility but increased tolerance to the oxidative redox state. Use of the TEVp$^{5M}$C110S/C130S variant, the refolded disulfide-rich bovine enteropeptidase or maize peroxidase was released via cleaving the sequence between the target protein and the cellulose-binding module bound to regenerated amorphous cellulose.


# Introduction

Recombinant protein is often produced in *E. coli* as insoluble aggregates, also called as inclusion bodies (IBs). Fusion tags as solubility enhancers are used to improve protein solubility, and other attached carriers are applied as affinity tags for rapid purification [1]. Inactive enzymes in IBs are usually refolded in vitro to form active ones in soluble form [2]. Several fusion passengers are developed to improve refolding efficiency [3]. On-column refolding dependent on the fused affinity tag increases yields of the refolded proteins [4]. The cellulose-binding module (CBM) at the C-terminal appendix linked to the single-chain antibody is effective for renaturation of the construct bound to the cellulose [5]. For refolding disulfide-rich proteins, the buffer is prepared with the specified ratio of the reduced and oxidized glutathione (GSH and GSSG), or that of the cysteine and cystine to yield the oxidative redox state for assisting the disulfide bonds formation [6].

Ideally, fusion protein tags are required to be removed, due to their affecting enzyme activity, protein crystallization, and producing unnecessary immunological reaction. This process is performed by the specific protease for proteolysis of the designed sequence incorporated in the fusion proteins. On the other hand, removal of the fusion tag with the specific proteases releases highly active target enzymes, for instance, diaminopropionate ammonia-lyase from *E. coli* (eDAL) and *Salmonella typhimurium* (sDAL), and *E. coli* phospholipase A, facilitating the designed fusion protein switch to detect the protease activity by the coupled assay [7,8]. Among the commonly-used sequence-specific protease tools, thrombin, enterokinase (EK, also called enteropeptidase), and Factor Xa containing the disulfide bonds are not functionally produced in *E. coli* cytoplasm. In contrast, the TEVp, as a thiol protease dependent with the catalytic triad including Cys151, displays several advantages, like high sequence specificity for cleaving the sequence ENLYFQG (tevS) between Q and G to retain only extra glycine residue in target protein,

functioned in *E. coli* cytoplasm, and in broad range of pH or temperature. In addition, it retains the catalytic activity in various buffers and in the presence of diverse additives, and supply of reductants such as dithiothreitol or 2-mercaptoethanol suppresses the TEVp activity loss owing to prohibition of cysteine oxidation [10]. For overcoming TEVp disadvantages, substitution of S219 with N219, V219 or P219 protects the auto-cleavage [11], and mutations of L56V/S135G and T17S/N68D/I77V increases the solubility [12,13]. The TEVp is inactivated in the endoplasmic reticulum of HEK-239T cells, but introduction of N23Q/C130S/T173G activates the cleavage efficiency. Additionally, the activity is decreased by introducing the C110S in the triple mutant, and blocked by the C19S [14]. Mutation of certain amino acid residues improves the TEVp characters, but sometimes impairs protein folding concurrently. For example, directed evolution changes protease specificity or improves the catalytic efficiency, but production of the soluble variants requires induction at low temperature or fusion of the powerful solubility enhancer, *E. coli* maltose binding protein [15,16]. So far, the highly active variant for site specific tag-removal under the oxidizing redox state to benefit disulfide bond formation during protein refolding is still lacking.

Previously, we mutated the S219V for inhibiting protease auto-cleavage and five mutated amino acid residues for enhancing protein solubility to generate the TEVp variant, TEVp$^{5M}$ [17]. Use of two *E. coli* expression strains BL21(DE3) and Rossetta(DE3) for expressing the four TEVp missense mutants under two induction conditions, we identified that desirable soluble production and stability were not shared in the TEVp variants [18]. In the effort to create the TEVp variant with the improved oxidative stability, and in order to better understanding the mutations on protein solubility and conformational quality, we mutated each of three cysteine residues to serine ones, and combined the mutations in the TEVp$^{5M}$ to construct seven variants and used three *E. coli* host expression strains for testing protein solubility and the cleavage activity of TEVp variants under the same induction conditions. The Origami (DE3) strain with disruption of the genes encoding thioredoxin reductase and glutathione reductase was chosen for investigating the solubility and activity of the optimized soluble TEVp variants in the oxidative cytoplasm. Based on the current data, we confirmed that mutations of C110S and/or C130S improved the oxidative stability of the TEVp$^{5M}$, and protein solubility was little impacted, different that of the C19S. The TEVp$^{5M}$C110S/C130S variant was efficient for separating the bovine EK (bEK) catalytic domain and maize peroxidase (mPex), through cleaving the fusion protein containing the CBM tag bound to regenerated amorphous cellulose (RAC) for matrix-assisted refolding of two disulfide-rich proteins.

# Materials And Methods

### Bacterial Strains, Plasmids and Reagents

*E. coli* strains DH5α, BL21(DE3), BL21(DE3)pLys, Rossetta(DE3) and Origami(DE3), the plasmids pET-22b and pET-28b are products of Novagen (Madison, WI). The plasmid pET32a-pEK was kindly provided by Professor Zhao Zhongbao [19]. The plasmids for expressing the double His6-tagged TEVp$^{5M}$, the emerald green fluorescent protein (EmGFP) as the C-terminal solubility reporter, the GST-tevS-eDAL, the CBM tag

fused to the linker GGTGGS around the tevS, and the GST-tevS-sDAL were constructed in our laboratory [8,17,20]. MutanBEST Kit for site-directed mutation and reagents for plasmid construction and protein overexpression were supplied by 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). Mouse anti-His6 monoclonal antibody and horseradish peroxidase (HRP) conjugated anti-mouse IgG antibody were obtained from GenScript, China. The compound 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 (DL-DAP), *o*-phenylenediamine (OPA), 2,4-dinitrophneylhydrazine (2,4-DNP) were bought from Sigma (USA).

## Plasmids Construction

Each and combined mutations C19S, C110S and C130S were introduced into the TEVp$^{5M}$ using the primer pairs C19S1 and C19S2, C110S1 and C110S2, C130S1 and C130S2 (Tab. S1), and pET28-TEVp$^{5M}$ as the template. After amplification, the PCR products were phosphorylated, ligated, and sequenced. The sequence encoding the TEVp variant was excised with *Nco* I/*Xho* I and subcloned into *Nco* I/*Sal* I sites of pET28-GFP for expressing the TEVp variant fused to the EmGFP. Since the Origami (DE3) strain bears kanamycin resistance, the tagged TEVp coding sequence with *Xba* I and *Xho* I excision was inserted into the *Xba* I/*Xho* I sites of the pET-22b vector.

According to comparison of the mature HRP amino acid sequence (Fig. S1), the fragment encoding the mature mPex Q45-S350 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 inserted into the *Bam*H I-*Xho* I site of the plasmid encoding the CBM tag. The plasmid for expressing the CBM tagged bEK was constructed as the same procedure.

## Production of the TEVp variants in different *E. coli* strains

Except where noted, induction and extraction of recombinant proteins in this study were conducted as follows. The plasmids were transformed into the *E. coli* strain. The recombinant cells were cultured overnight at 37 °C in 5 ml of lysogeny broth (LB), diluted to 50-fold and grown at 37°C. When $OD_{600}$ value was reached about 0.5, as measured on a U-2900 spectrometer (Hitachi, Japan), the target protein was induced by use of 0.5 mM isopropylthio-β-D-galactoside (IPTG). After cultured at 28 °C for 12 h in 10 ml liquid culture of a 50-ml shake flask at 220 rpm, cells were collected by centrifugation (4000 *g*, 10 min, 25 °C), washed with buffer A (20 mM Tris-HCl, pH 8.0, 100 mM NaCl), sonicated with 3 s for 198 times, and 10 s interval at 4 °C. Followed by centrifugation (12000 *g*, 15 min, 4 °C), the soluble and insoluble fractions were collected, and the pellet was washed twice with buffer A, and dissolved in 8 M urea, and centrifuged (12000 *g*, 15 min, 25 °C) to remove the precipitant. Protein amounts were determined by Bradford method, using bovine serum albumin as the reference. Protein samples were separated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). After electrophoresis, proteins on the gel were transferred to polyvinylidene fluoride membrane, immunoblotted with anti-His6 monoclonal

antibodies, and treated with affinity purified HRP-conjugated goat anti-mouse IgG. The band representing the target protein was visualized by addition of 4-chloro-1-naphthol solution dissolved in 20% methanol and 0.08% $H_2O_2$.

### Solubility Analysis of the TEVp Variants

The C-terminally fused EmGFP reporter was 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 [17].

### Coupled Assay of the Cleavage Activity

In vitro cleavage activity was assayed using the GST-tevS-eDAL purified by Ni-NTA [18]. 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. The cleavage was reacted at 30 °C for 1 h, and the activity was determined by coupled assay. This enzyme depends on the PLP cofactor and catalyzes DL-DAP to pyruvate and ammonia. The reaction mixture for testing eDAL activity contained 50 μM PLP and 10 mM DL-DAP, and the prepared eDAL 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 (12000 *g*, 10 min, 25 °C), absorbance at 520 nm representing pyruvate amounts was measured.

### Purification of the Soluble TEVp Variants

The expression plasmids were transformed into the Rossetta (DE3) cells. After induction at 28 °C for 12 h, cells in 500 ml LB culture 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 with 3 s for 198 times, and 10 s interval at 4 °C. The supernatant was loaded on a column containing 4 ml Ni-NTA with pre-equilibration with 40 ml buffer B, washed twice with 40 ml in buffer B (pH 8.0) containing 40 mM imidazole, and eluted with 40 ml buffer B (pH 8.0) containing 250 mM imidazole. Purified protein was concentrated by use of the Ultra-15 centrifugal filter tube equipped with the Ultracel-10 membrane, and exchanged with buffer A. Purity of the TEVp variants were characterized by SDS-PAGE.

### Modification of Free Cysteine Residues in Purified TEVp proteins

The free cysteine residues of the TEVp constructs were labeled with AMS as the described method [21]. Purified TEVp variants was incubated with either 150 μM $CuCl_2$ for oxidizing the free cysteine residues on the TEVp surface or 1 mM dithiothreitol (DTT) for reducing the oxidized cysteine ones at 25 °C for 1 h, and precipitated by trichloroacetic acid. After it was washed twice, centrifuged (12000 *g*, 1 min, 25 °C) and re-suspended with buffer A, the precipitated protein was labeled with AMS, a maleimidyl reagent specifically alkylating free thiol group of cysteine to increase molecular weight up to 0.5 kDa [22]. When the reaction was lasted for 20 min at room temperature, the mixture was centrifuged (12000 *g*, 10 min, 4

°C) and washed with buffer A. The labeled TEVp variants were incubated with the SDS-PAGE loading buffer in absence of DTT at 100 °C for 10 min, and separated by 12% SDS-PAGE under non-reducing condition.

### Refolding of the Fusion Proteins and Release of the Target Enzymes via Tag Removal

The IBs from the BL21(DE3) cells carrying the plasmids encoding the CBM tagged bEK or mPex were collected, washed and re-suspended with buffer C [30 mM Tris/HCl, 150 mM NaCl, 10% (v/v) glycerol, 0.5% (v/v) Triton X-100, pH 7.5]. Then, IBs were re-suspended with buffer C in the absence of Triton X-100. The tagged bEK in the prepared IBs was solubilized with buffer D (30 mM Tris-HCl, 200 mM NaCl, 8 M urea, 5 mM DTT, and 10 mM EDTA-$Na_2$). To ensure sufficient amounts of IBs to be solubilized, and the mismatched disulfide bonds in the proteins were reduced, the mixture was incubated at room temperature for 2 h and centrifuged (18000 *g*, 30 min, 25 °C) to remove the pellet [23]. The mPex in the IBs was solubilized with buffer E (40 mM Tris-HCl, pH 9.0, 4.5 M urea, 5 mM DTT) to a protein concentration of 0.3 mg/mL, according to the published report with slight modification [24]. The suspension was centrifuged (18000 *g*, 30 min, 25 °C) and the solubilized protein was collected.

For increasing more amounts of the CBM tag than the cellulose, RAC was prepared, as previous described procedure [20]. For the bEK refolding, the protein was diluted with buffer F (100 mM Tris–HCl, 6 M urea, 10 mM cystine, pH 8.0), and RAC was added. The mixture was diluted slowly with buffer G [80 mM Tris–HCl, 0.7 M urea, 15% (v/v) glycerol, 0.5 mM cysteine, 5 mM cysteine, 2 mM $CaCl_2$]. For the mPex refolding, the denatured proteins were diluted with buffer H [40 mM Tris-HCl, pH 8.5, 0.5 M urea, 5% glycerol (V/V), 2 μM hemin, 2 mM $CaCl_2$, 0.5 mM GSH and 5 mM GSSG]. After the refolding process was finished, the mixture was centrifuged (3000 *g*, 10 min, 25 °C), the resin was washed three times with buffer G for the bound bEK, or four times with buffer H in absence of hemin for the bound mPex. Purified TEVp C110S/C130S variant was incubated with the refolded protein bound to RAC with mass ratio of 1: 10 at 10 °C for 24 h. Then, Ni-NTA resin was added and incubated for 2 h at room temperature. Followed by centrifugation (3000 *g*, 10 min, 25 °C), the supernatant was collected.

### Activity Assay of the Refolded Enzymes

The prepared protein samples were subjected to SDS-PAGE analysis. The bEK cleaving the GST tagged sDAL with incorporation of the D4K as the bEK recognition sequence was analyzed, based on the coupled assay of the sDAL activity. The bEK can cleave the eDAL at the undesired site to inactivate the eDAL [8]. The mPex catalyzes the degradation of $H_2O_2$ using OPA as a hydrogen donor, which turns yellow upon oxidation [25]. The freshly prepared mPex was incubated in the buffer I (20 mM Tris-HCl, pH 7.5, 50 μg/mL OPA, 10 mM or 30 mM $H_2O_2$) at 37 °C for 30 min, and absorption at 411 nm was measured.

### Statistical Analysis

For solubility analysis, five samples from different colonies randomly selected for culturing were conducted and the data from the closest value of three biological replicates were calculated. For activity

assay, the data were from three technique replicates. Data were indicated as means ± standard deviations (SD). Data were evaluated using a one-tailed t-test. The data were analyzed using SPSS ver. 22 (SPSS Inc., USA).

# Results

### Production of the TEVp Variants in Different *E. coli* Strains

To avoid obvious activity loss under the oxidized redox state, we introduced each or combined C19S, C110S and C130S into the TEVp$^{5M}$. Three constructs include M1-M3 with one cysteine residue replaced with serine one. The M4–M6 constructs contained the combined two mutations, and M7 contain three mutations (Tab. 1).

For exploring the introduced mutations on the protein solubility, activity, and the enhanced oxidative stability, we selected four *E. coli* expression strains for production of the variants at the same induction conditions. As detected by SDS-PAGE (Fig. 1a), four variants including the M1, M4, M5 and M7 containing the C19S mutation are less soluble than the TEVp$^{5M}$ in *E. coli* BL21 (DE3) cells, and more amounts of variant proteins were aggregated than those of the TEVp$^{5M}$. Western blot analysis showed that insoluble and insoluble TEVp were produced. Protein folding in the M1 was impaired by the C19 mutation, and not recovered with further induction of the C110S and/or C130S in the M4, M5 and M7 variant. In the BL21(DE3) pLysS strain for inhibiting the background expression, four variants containing the C19S also showed the less solubility (Fig. 1b), whereas the insoluble expression levels were decreased as well, in contrast to the correspondent production in the BL21(DE3). The observation suggested that decreased protein production was not efficient for improving the variant solubility. Different from the BL21(DE3) pLysS, augment of rare tRNA levels in the Rossetta (DE3) usually enhances protein production, but in this work, it did not augment soluble yields of the constructs (Fig. 1c), although several rare arginine codons are distributed in the TEVp coding sequence [26]. Considering the C19S mutation decreased protein solubility, we expressed other three TEVp variants in the Origami (DE3) and testified the enhanced soluble production than the TEVp$^{5M}$ under the oxidative environment (Fig. 1d), suggesting that the C110S and/or C130S mutations confered the TEVp$^{5M}$ with the improved oxidative stability.

### Quantitative Analysis of Protein Solubility

Using the C-terminal EmGFP reporter, the constructed TEVp solubility was quantitatively analyzed. The TEVp$^{5M}$ containing the C110S and/or C130S mutations showed higher solubility than the TEVp$^{5M}$ in the BL21(DE3) stain (Fig. 2a), but soluble productivity was comparable to the TEVp$^{5M}$ in the BL21(DE3)pLys strain (Fig. 2b). The other constructs in the former host strain were less soluble than the latter one, suggesting that decreasing protein production level assisted protein folding, similar to the lowering induction temperature for enhancing soluble production of the TEVp variant with the impaired protein folding [15,18]. Supply of rare tRNAs increased soluble yields of the TEVp$^{5M}$ variant, but decreased those of all constructs (Fig. 2c). Soluble production of the variants containing the C110S and/or C130S

mutations was higher than the TEVp$^{5M}$ in the Origami (DE3), probably due to the improved oxidative stability (Fig. 2d). SDS-PAGE analysis also displayed that soluble production of the GFP tagged TEVp$^{5M}$ and the constructs containing substitution of the C110S and/or C130S (Figs. S2, a-d). It is noted that one molecule of fluorophore biosynthesis in the GFP is accompanied by the generation of one molecule of $H_2O_2$ [27]. So, the GFP is not an ideal reporter for in vivo detecting the solubility of the target protein at high expression levels.

## Cleavage Activity Analysis of the TEVp variants in Clear Lysates

Because the fused GFP reporter possibly affects the oxidative stability of the TEVp variants, we selected the His6-tagged TEVp variants for assay of the cleavage activity. Each of the TEVp$^{5M}$, M2, M3 and M6 variants in supernatants cleaved purified GST-tevS-eDAL partially, more efficiently than the other ones (Figs. S3, a-d). By comparison of the mutants in the selected *E. coli* strain, the M3 variant in the BL21(DE3), TEVp$^{5M}$ in BL21(DE3)pLysS and Rossetta (DE3), and the M6 variant in Origami (DE3) showed the highest activity (Figs. 3a-3d). On the other hand, the M6 was most soluble in the BL21(DE3), BL21(DE3)pLysS and Origami (DE3) strains, and the TEVp$^{5M}$ and the M3 showed highest solubility in the Rossetta (DE3) strain (Figs. 2a-2d). The results indicated that, protein solubility, conformational quality and oxidative stability are divergent features, and all parameters cannot be simultaneously enhanced in the recombinant TEVp. The correlation among the characters analyzed in the study provides a consistent view of mutational effects on the TEVp constructs.

## Free Cysteine Residues Labelling in Purified TEVp Constructs

The free cysteine residues in the purified target proteins produced in *E. coli* form the disulfide bonds via $CuCl_2$ oxidation [21], or sulphenic acids through $H_2O_2$ oxidation [22]. The oxidizing process in vitro is detected by using AMS attributable to failure of its association with the oxidized cysteine [21,22]. To label the TEVp, we detected the purified TEVp purity by SDS-PAGE analysis under reducing condition (Fig. 4a). The AMS-labelled TEVp$^{5M}$ exhibited several bands in non-reducing SDS-PAGE (Fig. 4b), 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$^{5M}$, thanks 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. 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 the Oxidized Redox State on the Activity of Purified TEVp constructs

The purified TEVp constructs cleaved the partial fusion protein (Fig. 5a). The specific activity of the TEVp constructs was comparable to the TEVp$^{5M}$, except for the M3 showing the slightly decreased activity (Fig. 5b). With addition of 2 mM DTT, the three constructs showed the cleavage activity equivalent with the TEVp$^{5M}$ (Fig. 5c). On the contrary, addition of compounds for forming the oxidative state including 5 mM

cystine and 0.5 mM cysteine, or 5 mM GSSG and 0.5 mM GSH inhibited the TEVp$^{5M}$ activity, but the inhibitory effect was less for the TEVp$^{5M}$C110S/C130S (Fig. 5c). The concentrations of the compounds such as cysteine and GSSG containing the disulfide bonds are higher than those of cysteine and GSH containing the thiol group, leading to the oxidative redox state in the buffer. The data suggested that the engineered TEVp variant was more insensitive to the oxidized redox medium.

### Removal of fusion tags in two refolded proteins containing multiple disulfide bonds

For the constructed TEVp application, we used two proteins as the fusion partner for refolding. The tagged or untagged bEK and HRP are produced as the IBs[19, 23,24], same as insoluble aggregation of the CBM attached bEK or Mper (Fig. 6a). The fusion protein was refolded on RAC via the CBM tag assaciation, and the target protein was detached from the resin by using the purified M6 variant (Fig. 6b). The refolded bEK cleaved the GST tagged sDAL into two parts, and cleavage activity was assayed based on the released sDAL catalysis (Fig. 6c). With addition the refolded mPex, 10 and 30 mM $H_2O_2$ were partially transformed into yellow compounds in the presence of OPA, in contrast to the heat-inactivated enzyme (Fig. 6d). Based on the color change in the mixture containing 30 mM $H_2O_2$ as the control, auto-oxidation was most likely reacted spontaneously. Because of the CBM tag associated with RAC, the buffer is exchanged by centrifugations for allowing the other TEVp construct cleavage. Use of the other fusion tag for refolding the target protein, for example, the small molecular chaperone [28], the M6 as the TEVp$^{5M}$C100S/C130S is an ideal tool, because buffer exchange requires long time.

## Discussion

In this study, we engineered one to three cysteine residues in the TEVp$^{5M}$ to test the mutations effects in different *E. coli* strains, and identified that desirable protein productivity, solubility, cleavage activity representing the conformational quality, and the oxidative stability were not simultaneously favored in the recombinant *E. coli* cells. As the folding and solubility reporter in *E. coli*, different GFP variants have been used in vivo and in vitro. A linear correlation between cell fluorescence and soluble amounts of the engineered protein variants is built by use of the fused enhanced GFP (eGFP) variant [29]. However, the correlation changes significantly for various GFP variants [30]. The insolubly aggregated GFP partner still emits the fluorescence [31]. Use of clear lysates containing the GFP reporter for determining the protein solubility is more accurate [32], but in this study, the fluorescence intensities of the crude extracts were changed irregularly for the different TEVp constructs fused to the EmGFP among the selected *E. coli* strains. This change is probably resulted from is the GFP toxicity, since high level production of the TEVp$^{5M}$ fused to the EmGFP at 37°C induction for over 4 h inhibits cell growth of *E. coli* Rosetta (DE3) [17].

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. In this work, we used the

TEVp$^{5M}$ with the 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 [37]. The eGFP forms non-fluorescent oligomer via the mismatached disulfides, leading to loss of fluorescence in *E. coli* periplasm by the SecYEG transportation pathway. In contrast, the superfolder GFP variant, or the red fluorescent protein in absence of cysteine residue is brightly fluorescent in the *E. coli* periplasm. Although we created the TEVp$^{5M}$C110S/C130S with the increased oxidative stability, the catalytic C151 in the TEVp construct is also prone to be oxidized. Introduction several mutations containing the C151S results in the TEVp transformed from the thiol protease to the serine one [38]. The C110S/C130S mutations are potentially used for further enhancing the stability of the TEVp variant as the serine protease under the oxidative environment. The C19 buried in the first strand in the TEVp is reluctant to form the uncorrected disulfide bridge with another TEVp molecule. In case of C19 oxidized to the sulphenic acid, protein folding will be probably impaired, due to the TEVp$^{5M}$C19S produced as inclusion bodies. The insoluble TEVp construct retains the cleavage activity [39]. Further study will focus on construction of the cysteine-free TEVp variant aggregate for tag-removal.

The CBM bound to RAC 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 [5]. Moreover, the refolded CBM tagged two enzymes on RAC were used as immobilizates for improving use efficiency and easy extraction from the reaction mixture. Once the TEVp variant constructed in the present work was incubated, the target protein is detached from the resin. Several TEVp variants as useful tools have been developed, including the significant inhibition of self-cleavage [11], the improved protein solubility [12,13], thermostability [40], and catalytic activity [16,41]. Our work expands the toolkit of TEVp variants for various applications and the developed variant will be used for removing the fusion tag for producing the tag-free proteins containing disulfide bonds as medicines.

In summary, the multiple mutations afforded the engineered TEVp$^{5M}$C110S/C130S variant with the enhanced solubility, and the obviously decreased self-inactivation, and the improved oxidative stability. Cleavage of the immobilized fusion proteins with the TEVp variant in the buffer containing the compounds for forming the oxidative redox state mediated release of the refolded disulfide-rich proteins from the prepared RAC resin.

# 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, Yinghua Chen, Jiong Hu, Xuelian Yu, Weiyu Wang, Jun Fan

### Corresponding author

Correspondence to Jun Fan.

### Author contributions

Designed the experiments and wrote the paper: Jun Fan. Performed the experiments: Enkhtuya Bayar, Yuanyuan Ren, Yafang Hu, Yinghua Chen, and Xuelian Yu. Analyzed the data: Yuanyuan Ren, Yafang Hu, Yinghua Chen, Xuelian Yu.

### Ethics declarations

### Ethical approval and consent to participate

All authors have read and agreed to the ethics for publishing the manuscript.

### Competing interests

The authors declare that they have no competing interests.

### Supplementary data

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

# References

1. Kosobokova, E.N., Skrypnik, K.A., Kosorukov, V.S. (2016). Overview of fusion tags for recombinant proteins. *Biochemistry (Mosc), 81,* 187-200.
2. 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.
3. Yamaguchi, H., Miyazaki, M. (2014). Refolding techniques for recovering biologically active recombinant proteins from inclusion bodies. *Biomolecules, 4,* 235–251.
4. Jungbauer, A., Kaar, W., Schlegl, R. (2004). Folding and refolding of proteins in chromatographic beds. *Current Opinion in Biotechnology, 15,* 487–494.
5. 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.

6. Wingfield, P.T., Palmer, I., Liang, S.M. (2014). Folding and purification of insoluble (inclusion body) proteins from *Escherichia coli. Current Protocols in Protein Science, 78,* 5.1–6.5.30.
7. Leach, C.A., Tian, X., Mattern, M.R., Nicholson, B. (2009). Detection and characterization of SUMO protease activity using a sensitive enzyme-based reporter assay. *Methods in Molecular Biology, 497,* 269-281.
8. 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,*
9. Waugh, D.S. (2011). An overview of enzymatic reagents for the removal of affinity tags. *Expression and Purification, 80,* 283–293.
10. Sun, C., Liang, J., Shi, R., Gao, X., Zhang, R., Hong, F., Yuan, Q., Wang, S. (2012). Tobacco etch virus protease retains its activity in various buffers and in the presence of diverse additives. *Protein Expression and Purification,* 82, 226–231.
11. Kapust, R.B., Tözsér, J., Fox, J.D., Anderson, D.E., Cherry, S., Copeland, T.D., Waugh, D.S. (2001). Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Engineering, 14,* 993–1000.
12. van den Berg, S., Löfdahl, P.A., Härd, T., Berglund, H. 2006. Improved solubility of TEV protease by directed evolution. . *Journal of Biotechnology,121,* 291–298.
13. Cabrita, L.D., Gilis, D., Robertson, A.L., Dehouck, Y., Rooman, M., Bottomley, S.P. (2007). Enhancing the stability and solubility of TEV protease using in silico design. *Protein Science, 16,* 2360–2367.
14. 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.
15. Verhoeven, K.D., Altstadt, O.C., Savinov, S.N. (2012). Intracellular detection and evolution of site-specific proteases using a genetic selection system. *Applied Biochemistry and Biotechnology, 166,* 1340–1354.
16. Sanchez MI, Ting AY. (2020). Directed evolution improves the catalytic efficiency of TEV protease. *Nature Methods, 17,* 167–174.
17. 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.
18. Fang, J., Chen, L., Cheng, B., Fan, J. (2013). Engineering soluble tobacco etch virus protease accompanies the loss of stability. *Protein Expression and Purification,92,* 29–35.
19. 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.
20. 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.

21. 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.
22. Iglesias-Baena, I., Barranco-Medina, S., Lázaro-Payo, A., López-Jaramillo, F.J., Sevilla, F., Lázaro, J.J. (2010). Characterization of plant sulfiredoxin and role of sulphinic form of 2-Cys peroxiredoxin. *Journal of Experimental Botany, 61,* 1509–1521.
23. 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.
24. 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.
25. 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.
26. Fang, J., Zou, L., Zhou, X., Cheng, B., Fan, J. (2014). Synonymous rare arginine codons and tRNA abundance affect protein production and quality of TEV protease variant. *PLoS One9*, e112254.
27. 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.
28. Yurkova, M.S., Sharapova, O.A., Zenin, V.A., Fedorov, A.N. (2019). Versatile format of minichaperone-based protein fusion system. *Scientific Reports. 9,*
29. 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.
30. Gupta, R. D. Tawfik, D. S. (2008). Directed enzyme evolution via small and effective neutral drift libraries. *Nature Methods, 5,* 939–942.
31. 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.
32. Hedhammar, M., Stenvall, M., Lönneborg, R., Nord, O., Sjölin, O., Brismar, H., Uhlén, M., Ottosson, J., Hober, 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.
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.

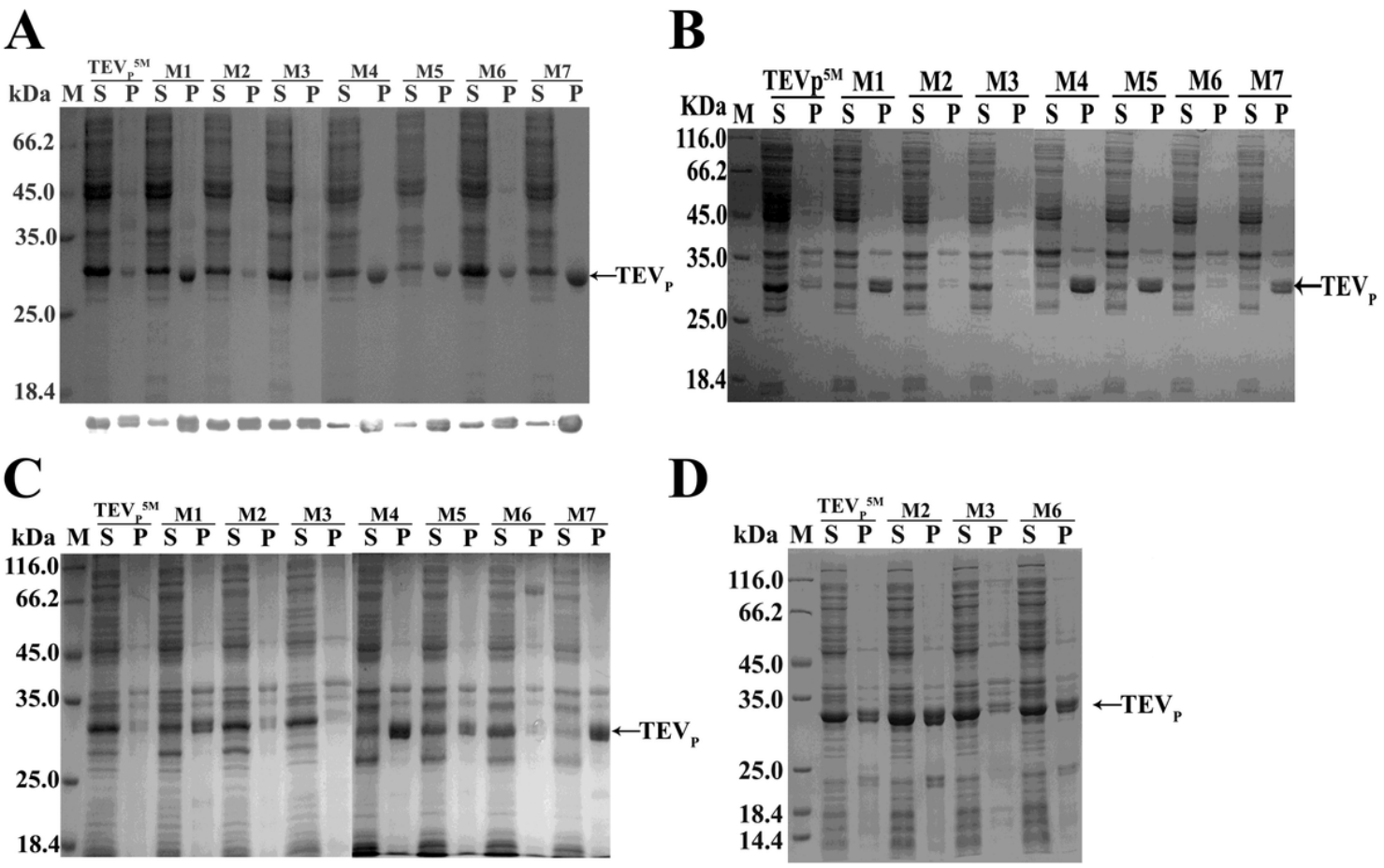
35. Kosuri, P., Alegre-Cebollada, J., Feng, J., Kaplan, A., Inglés-Prieto, A., Badilla, C.L., Stockwell, B.R., Sanchez-Ruiz, J.M., Holmgren, A., Fernández, J.M. (2012). Protein folding drives disulfide formation. *Cell, 151,* 794–806.
36. Qin, M., Wang, W., Thirumalai, D. (2015). Protein folding guides 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, 12,* 543–548.
38. Shafee, T., Gatti-Lafranconi, P., Minter, R., Hollfelder, F. (2015). Handicap-recover evolution leads to a chemically versatile, nucleophile-permissive protease. *Chembiochem, 16,* 1866–1869.
39. Li, G.Y., Xiao, Z.Z., Lu, H.P., Li, Y.Y., Zhou, X.H., Tan, X., Zhang, X.Y., Xia, X.L., Sun, H.C.(2016). A simple method for recombinant protein purification using self-assembling peptide-tagged tobacco etch virus protease. *Protein Expression and Purification,128,* 86–92.
40. 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,*
41. Denard, C.A., Paresi, C., Yaghi, R., McGinnis, N., Bennett, Z., Yi, L., Georgiou, G., Iverson, B.L. (2021). YESS 2.0, a tunable platform for enzyme evolution, yields highly active TEV protease variants. *ACS Synthetic Biology, 10,* 63–71.

# Tables

**Table 1. The variants constructed in this study**

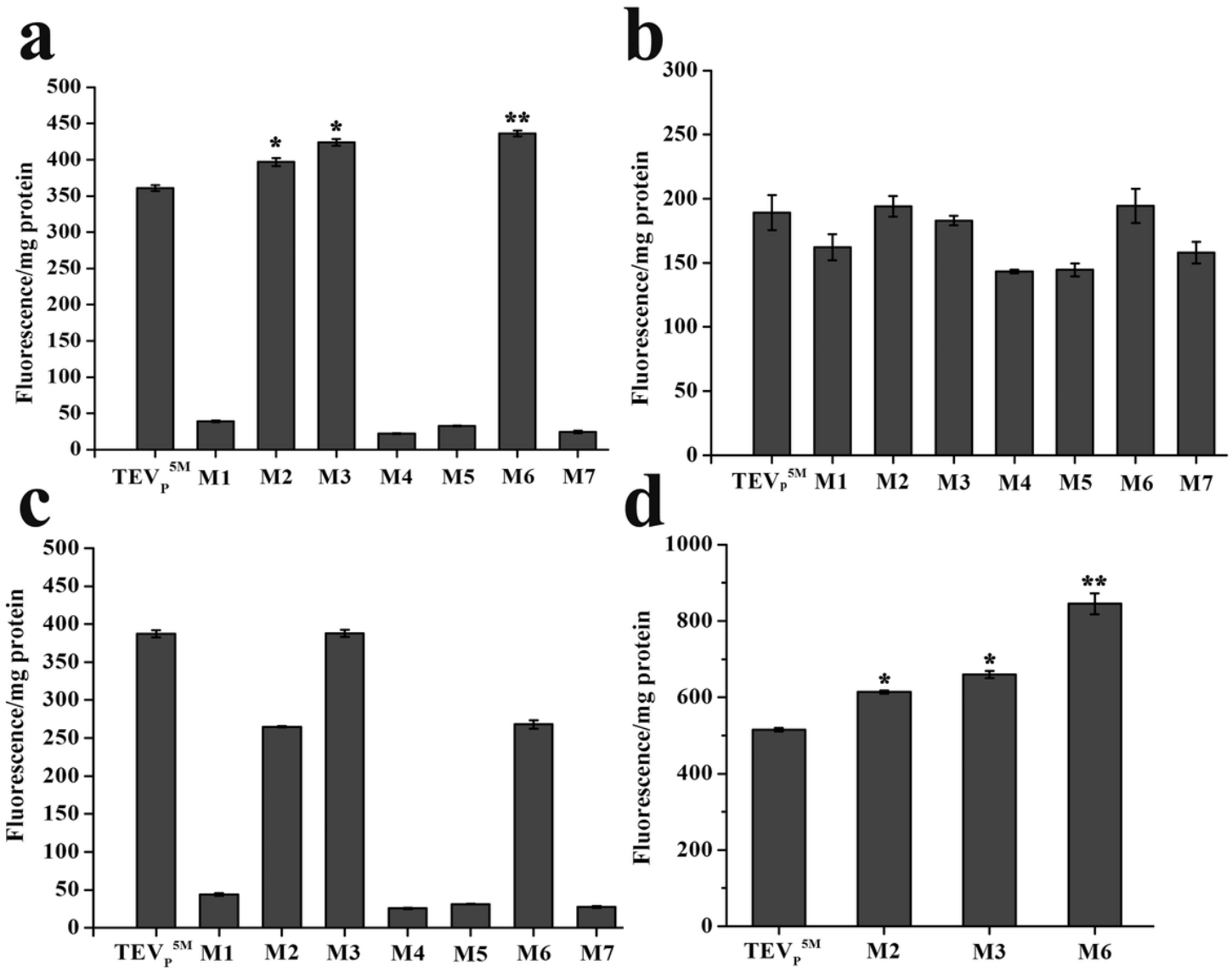
| M1 | M2 | M3 | M4 | M5 | M6 | M7 |
|---|---|---|---|---|---|---|
| C19S | C110S | C130S | C19S/C110S | C19S/C130S | C110S/C130S | C19S/C110S/C130S |

# Figures



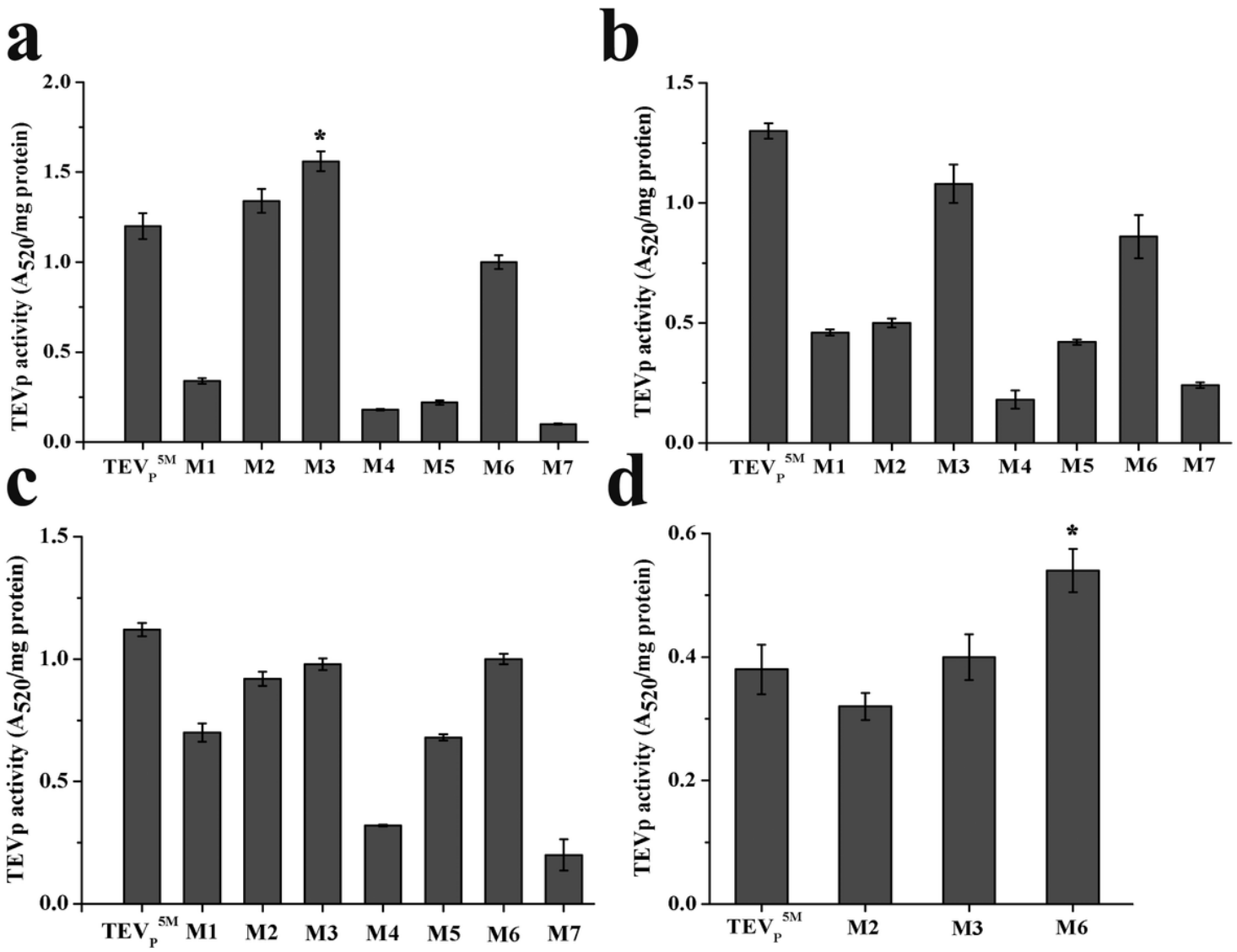

## Figure 1

Production of the constructed 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.



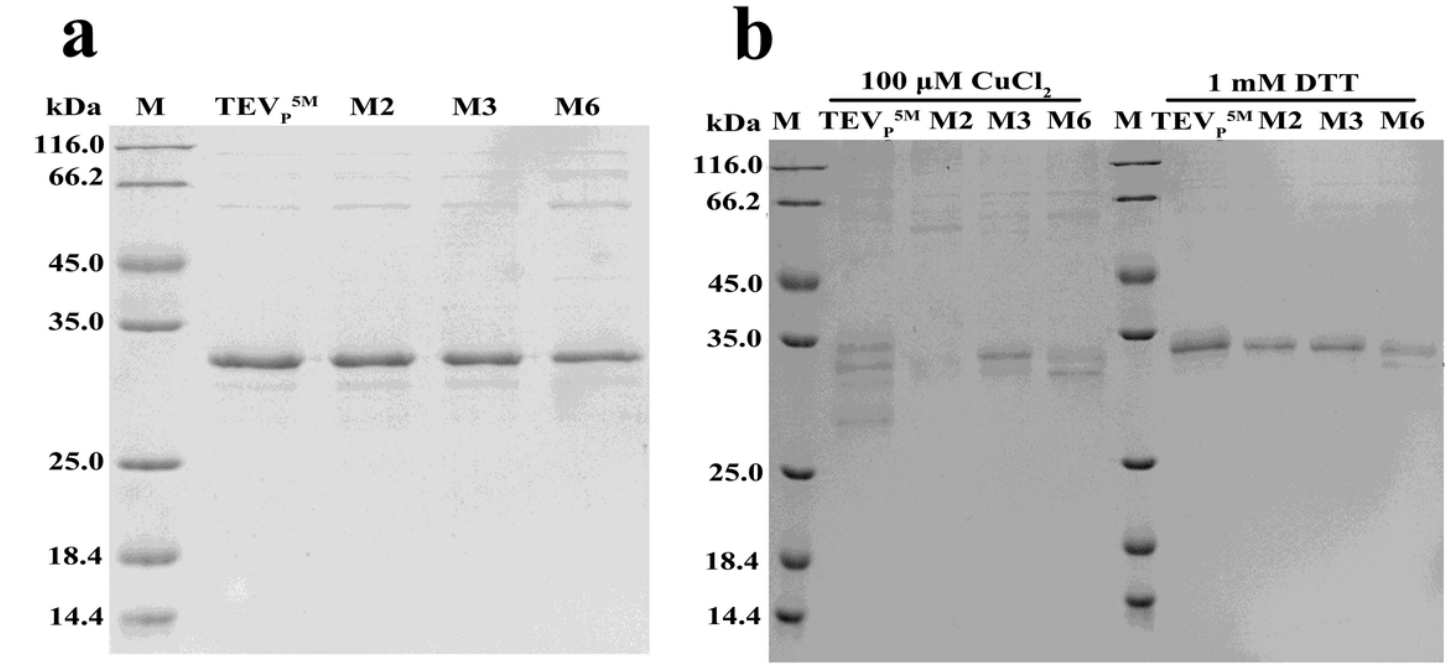

## Figure 2

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. At least five samples were measured under the same induction conditions, and the closest values were counted. The asterisk indicated significant differences; * $p < 0.01$.



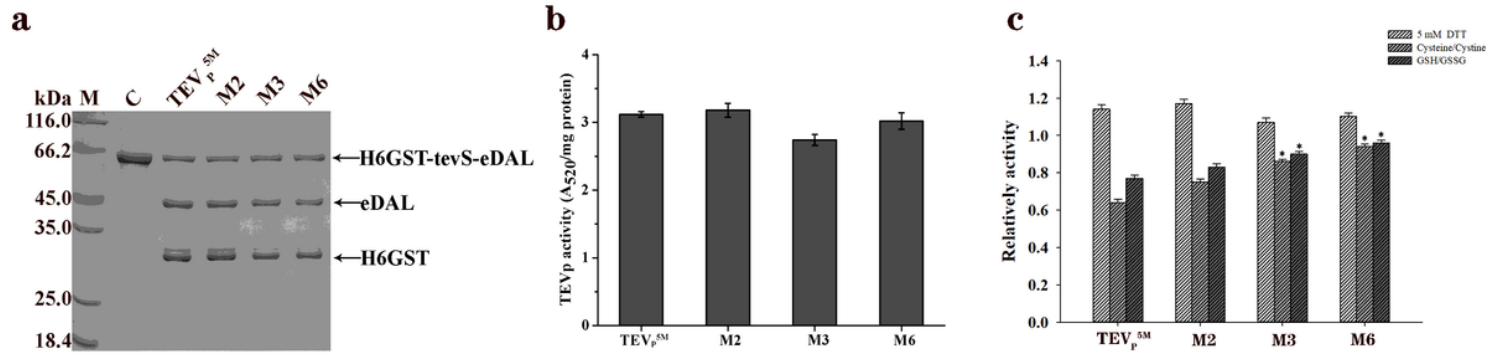

## Figure 3

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. The asterisk indicated significant differences; * $p < 0.01$.



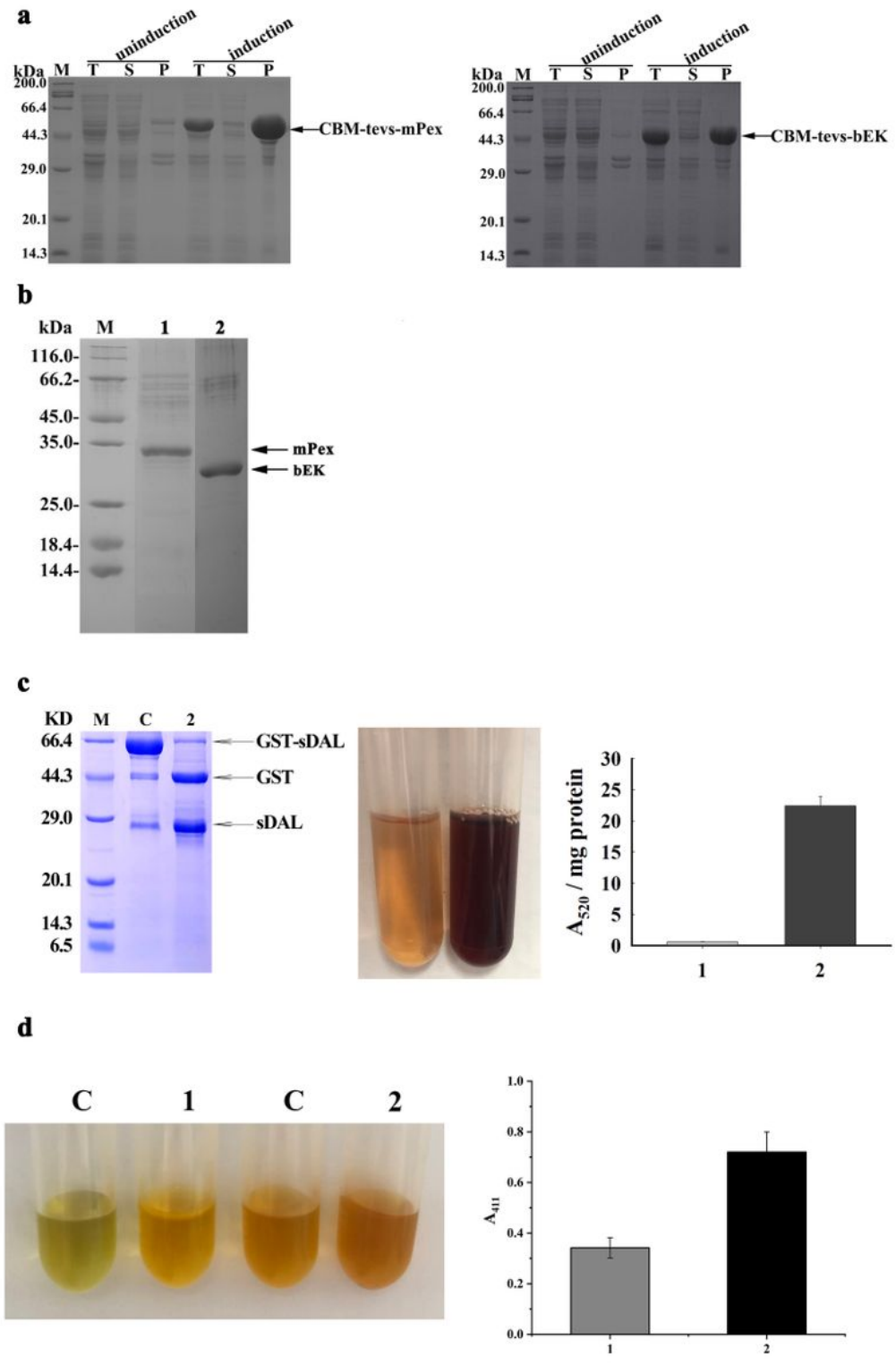

## Figure 4

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 5

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. The asterisk indicated significant differences; * p< 0.01.




## Figure 6

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 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.

## Supplementary Files

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

- Answertoreviwer.docx
- Supplementarymaterial.docx