Journal of Bioscience and Bioengineering
VOL. xx No. xx, 1–8, 2017

www.elsevier.com/locate/jbiosc

Evaluation of rice tetraticopeptide domain-containing thioredoxin as a novel solubility-enhancing fusion tag in *Escherichia coli*

Wenjun Xiao, Li Jiang, Weiyu Wang, Ruyue Wang, and Jun Fan*

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

Received 2 June 2017; accepted 24 August 2017
Available online xxx

Fusion of solubility-enhancing tag is frequently used for improving soluble production of target protein in *Escherichia coli*. The Arabidopsis tetraticopeptide domain-containing thioredoxin (TDX) has been documented to exhibit functions of disulfide reductase, foldase chaperone, and holdase chaperone. Here, we identified that fusion of rice TDX with the smaller size increased soluble expression levels of three fluorescent proteins with different fluorophores in the *E. coli* strain BL21(DE3) or the Rosetta (DE3) strain with coexpression of six rare tRNAs, but decreased conformational quality of certain fluorescent proteins, as comparison with the His6-tagged ones. Among five maize proteins, the rice TDX fusion carrier displayed higher solubility-enhancing activity than the yeast SUMO3 tag toward three proteins in both *E. coli* strains. Five fusion constructs were cleaved with the co-expressed TEV protease variant, but the released target proteins were partly insolubly aggregated in vivo. Attachment of the His6-tag to the TDX tagged proteins had little impact on protein solubility. After Ni-NTA purification, five His6-TDX tagged proteins displayed different apparent purities. Taken together, our work presents that rice TDX tag is a novel solubility enhancer.

[**Key words:** Rice tetraticopeptide domain-containing thioredoxin; Fusion tag; Soluble production; Fluorescent proteins; Maize proteins; *Escherichia coli*; In vivo cleavage; Purification]

As a host, *Escherichia coli* is widely applied for producing heterologous protein owing to its easy culture and genetic manipulation, fast growth and simple fermentation (1). However, protein of interest is often expressed mainly as insoluble aggregates, due to its poorly folded in *E. coli*. Fusion of solubility-enhancing tag is a simple and efficient approach to improve soluble production of target protein (2). The affinity tag such as His6-tag is another kind of fusion carrier for rapid purification. The tandem dual-tag containing the His6-tag and the solubility partner is constructed for increasing production and fast purification of the chimeric protein (3–6).

Except for the small ubiquitin related modifier 3 (SUMO3) from *Saccharomyces cerevisiae* (4), other commonly-used solubility-enhancing tags are from prokaryotes (2), such as *E. coli* maltose binding protein (MBP), thioredoxin (Trx), N-utilization substrate A (NusA) and *Schistosoma japonicum* glutathione *S*-transferase (GST). Several other fusion protein tags from prokaryotes are developed to rescue bacterially expressed aggregation-prone proteins (6–11). The *E. coli* molecular chaperones DnaK, GroEL, and trigger factor as the fusion partner improve soluble production of the difficult-to-express proteins (5,12). In addition to protein solubility enhancer, a few of peptides as N-terminal partners, also enhance soluble production of the selected proteins in *E. coli*, for example, the expressivity tag (EX-tag) containing N-terminal seven amino acid residues (MTDVTIK) of *E. coli* translation initiation factor II (13), and the designed SKIK peptide tag (14).

So far, sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis is frequently used to detect soluble expression level of target protein fused with solubility-enhancing tag (3–14). The C-terminal fusion peptide S-tag is applied as a tool to quantitatively measure soluble amounts of the chimeric proteins, so, the effects of different solubility enhancers on increasing certain target proteins are assessed (15). Fusion of the green fluorescent protein (GFP) variant at C-terminus is also applied to monitor fusion protein expression and solubility (16). Like GFP, uroporphyrinogen methyltransferase (UMT) as an orange-red fluorescent protein, and light-oxygen-voltage (LOV) domain as a flavin mononucleotide based fluorescent protein, are also used as fusion partner to detect soluble production of the recombinant proteins (17,18).

The tetraticopeptide domain-containing Trx (TDX) from *Arabidopsis thaliana* (AtTDX), a heat-stable and plant-specific Trx-like protein, exhibits the functions of disulfide reductase, foldase chaperone, and holdase chaperone (19). It contains three tetratricopeptide repeat domains and a Trx motif, and is located in the nucleus. The function of AtTDX is dependent with its oligomeric status, i.e., the protein in the low molecular weight form displays the predominately activities of disulfide reductase and foldase chaperone, while in the high molecular weight form it modulates holdase chaperone activity (19). Further study reveals that redox properties of the Trx motif in AtTDX is less thermostable than *E. coli* Trx (20).

Amino acid sequence alignment of AtTDX with other three plant homologues reveals that rice TDX is smallest (19). Considering that high level production of the relatively small size of rice TDX

* Corresponding author. Tel.: +86 551 65786021; fax: +86 551 65786021.
E-mail address: fanjun@ahau.edu.cn (J. Fan).

1389-1723/$ – see front matter
http://dx.doi.org/10.1016/j.jbiosc.2017.08.016

Please cite this article in press as: Xiao, W., et al., Evaluation of rice tetraticopeptide domain-containing thioredoxin as a novel solubility-enhancing fusion tag in *Escherichia coli*, J. Biosci. Bioeng., (2017), http://dx.doi.org/10.1016/j.jbiosc.2017.08.016

probably alleviates the metabolic burden to the host strain, in this study, we evaluated the rice TDX tag as a novel solubility enhancer, as comparison with a few of well-known solubility-enhancing tags. The TDX tagged proteins were cleaved by the coexpressed TEV protease. Fusion of the N-terminal His6-tag facilitated the TDX tagged proteins to be rapidly purified.

MATERIALS AND METHODS

Plasmids construction Total RNAs were extracted from young leaves of the cultivated rice variety Zhong Hua 11 (*Oryza sativa* L. subsp. *japonica* Kato) and the first cDNA was synthesized using mouse reverse transcriptase (RNaseH-free) by addition of the oligo $(dT)_{18}$ primer (Takara, Dalian). Then, the cDNA encoding rice TDX (GenBank accession no. AK063980) was amplified by PCR. The purified PCR amplicon was inserted into the *Nde* I/*Bgl* II sites of the plasmid pET-29b (Novagen, USA), then, the sequence encoding the tobacco tech virus protease (TEVp) cut site (tevS) was inserted into *Bgl* II and *Bam*H I to yield the plasmid pTDX (Fig. S1A). The new *Bam*H I cut site was introduced into pET-28b plasmid by synonymous mutation and the original *Bam*H I cut site was eliminated to generate the plasmid pET-28b1 (Fig. S2). Each target gene was inserted into the plasmid pET-28b1 to express the N-terminally His6-tagged protein.

The proteins used for testing solubility-enhancing activity of the rice TDX tag and comparing the effect of the TDX tag with that of the other constructed fusion tags are listed in Table 1. The gene encoding EmGFP, mLOV, and mSR were cloned previously. Except for the Ex-tagged EmGFP, other two maize proteins were less expressed in *E. coli* BL21(DE3) than the His6-MBP tagged one (21). To distinguish the results in this study from the previous report, the EX-tag and the MBP tag in the former construction were substituted with the His6-tag and rice TDX tag respectively. The tevS in the MBP-mLOV was deleted. The sequence encoding H6D4 linker, tevS and mSR was constructed previously (22), excised with *Bgl* II and *Xho* I, and inserted into *Bgl* II/*Xho* I sites of the pTDX vector. The gene encoding mGFP5 was amplified using the vector pCAMBIA1302 as the template and inserted into *Bam*H I/*Xho* I sites of the plasmid pET-28b1 or the pTDX. The vector expressing the mUMT (L92-A369) was constructed previously (23). The gene was amplified and subcloned into the plasmid pET-28b1 or the pTDX.

The gene encoding the mature enzymes including mGLTR (I82-N531), mGSAT (E42-I474), mChlI (N98-S415) and mChlD (S50-S755) involved in chlorophyll biosynthesis were amplified using maize inbred B73 leaves cDNA library as template. The putative chloroplast transit peptide in each of four maize protein precursors was determined by comparison with the other corresponding plant homologues. The amplified products were digested with *Bam*H I and *Xho* I (*Not* I, for the mGLTR coding sequence) and inserted into the plasmid pET-28b1 or the pTDX with the same treatment. All these homologous proteins from other plant species are prone to be aggregated or poorly expressed in *E. coli*, which are improved by co-expression molecular chaperones or fusion of the solubility enhancer (24–27). The TEVp variant $TEVp^{5M}$ with enhanced solubility and activity, and its inactive variant $TEVp^{5M}C151A$ were constructed (21). The coding sequences were excised from the pET-28b derived plasmid, and inserted into *Xba* I and *Xho* I sites of the plasmid pET-22b, respectively. The wild type yeast SUMO3 (herein named as SUMO), as well as the GroEL tag was cloned and instead of the TDX tag. All constructed plasmids were sequenced, and primers used in this study are listed in Supplementary Table S1.

Overexpression of the fusion proteins in the *E. coli* strains Except where noted, production and extraction of recombinant proteins were conducted as follows. The plasmids were transformed into *E. coli* BL21(DE3) or Rosetta (DE3) strain (Novagen, USA), and the recombinant cells were in a 1-ml lysogeny broth (LB) medium in a 15-ml shake tube with supply of kanamycin (50 μg/ml) for the BL21(DE3) cells, that of kanamycin (25 μg/ml) and chloramphenicol (25 μg/ml) for the Rosetta (DE3) cells. The transformants were diluted 200 fold, and cultured in a 5-ml LB medium in a 15-ml shake tube. When OD_{600} reached about 0.5, cells were induced by 0.5 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG), cultured at 28°C for 12 h, or at 37°C for 4 h, and collected by centrifugation. Cells were re-suspended with 0.6 ml buffer A (50 mM Tris/HCl, 100 mM NaCl, pH 8.0), sonicated at 4°C, and centrifuged. The pellets were solubilized with 0.2 ml of 8 M urea and centrifuged to remove cell debris.

Protein analyses Protein concentration was determined by Bradford method, using bovine serum albumin as standard. Protein samples in supernatants and pellets were separated by 13% SDS-PAGE. The gel was stained with Coomassie brilliant blue R250.

Cell fluorescence measurement Five recombinant colonies were randomly picked and cultured as mentioned above. For expressing the tagged mUMT, δ-aminolevulinic acid (ALA, Sigma, USA) as the pre-substrate of mUMT was dissolved in buffer A at pH 7.0 was added to the culture at the final concentration of 1 mM upon induction with IPTG at 28°C for 12 h, or at 37°C for 4 h. By catalysis of the *E. coli* genomic DNA encoding enzymes involved in heme biosyntesis, ALA is transformed into uroporphyrinogen III as UMT substrate (28). Fluorescent cells were photographed under irradiation of UV light at 350 nm. Cell fluorescence were measured using a F-4500 fluorescence spectrometer (Hitachi, Japan). The excitation and emission peaks are 473 nm and 509 nm for the mGFP5 (29), 488 nm and 515 nm for the EmGFP (30), 357 nm and 605 nm for the mUMT (27), 520 nm and 485 nm for the mLOV (21). Standard derivation was calculated from three biological replicates with similar resultant values. Cells were dried to constant weight through vacuum filtration at 60°C.

In vivo cleavage of the fusion proteins The BL21(DE3) cells coexpressing the TDX tagged maize protein with the active $TEVp^{5M}$ or inactive $TEVp^{5M}C151A$ were cultured in the presence of kanamycin (25 μg/ml) and ampicillin (25 μg/ml). After induction at 28°C for 12 h, protein samples in soluble and insoluble fractions were analyzed by SDS-PAGE.

Purification of the fusion proteins The His6-TDX tagged maize proteins produced in *E. coli* BL21(DE3) in 50 ml liquid culture of a 250-ml shake flask after induction at 28°C for 12 h were collected, re-suspended in buffer B (50 mM sodium phosphate, 300 mM NaCl, and 10 mM imidazole, pH 8.0), and sonicated. The supernatant was loaded on a nickel-nitrilotriacetic acid (Ni-NTA) agarose spin (Qiagen, Germany). Under centrifugal force at 380 ×*g*, the spin was washed three times with 0.6 ml buffer B, and 0.6 ml buffer B containing 40 mM imidazole (pH 8.0), respectively, and eluted with 0.3 ml buffer B containing 250 mM imidazole (pH 8.0).

RESULTS

Construction of fusion proteins for analyzing effects of rice TDX tag Two AtTDX proteins were searched in GenBank. AtTDX1, named in this study and as AtTDX previously (19), contains 380 amino acid residues (NM_112669.4). The AtTDX2 comprises of 373 amino acid residues (NM_001084706.1), which shares the same

TABLE 1. Constructed proteins in this work.

Protein name	Abbreviation	Gene ID	Fusion tag
Emerald GFP	EmGFP	KF293661.1	His6 TDX
Mutated GFP	mGFP5	KP893621.1	His6 TDX
Maize uroporphyrinogen III methyltransferase	mUMT	D83391	His6 TDX
Maize light-oxygen-voltage domain	mLOV	XR_564036.1	His6 MBP TDX
Maize serine racemase	mSR	JF502634.1	SUMO TDX
Maize glutamyl-tRNA reductase	mGLTR	103642104	SUMO TDX
Maize glutamate-1-semialdehyde 2,1-aminomutase	mGSAT	100193482	SUMO TDX
Maize magnesium chelatase subunits I	mChlI	DQ084025.1	SUMO TDX
Maize magnesium chelatase subunits D	mChlD	XM_008646181.2	GroEL TDX

Please cite this article in press as: Xiao, W., et al., Evaluation of rice tetraticopeptide domain-containing thioredoxin as a novel solubility-enhancing fusion tag in *Escherichia coli*, J. Biosci. Bioeng., (2017), http://dx.doi.org/10.1016/j.jbiosc.2017.08.016