ORIGINAL ARTICLE

High yields and soluble expression of superoxide dismutases in *Escherichia coli* due to the HIV-1 Tat peptide via increases in mRNA transcription

Yangdong Sun[1,4], Qiao Ye[2,4], Min Wu[3], Yonghong Wu[2], Chenggang Zhang[2] and Weiqun Yan[1]

This study aimed to validate the high yield and soluble expression of proteins carrying the transactivator of transcription (Tat) peptide tag, and further explored the potential mechanism by which the Tat tag increases expression. *Escherichia coli* superoxide dismutase (SOD) proteins, including SodA, SodB and SodC, were selected for analysis. As expected, the yields and the solubility of Tat-tagged proteins were higher than those of Tat-free proteins, and similar results were observed for the total SOD enzyme activity. Bacterial cells that overexpressed Tat-tagged proteins exhibited increased anti-paraquat activity compared with those expressing Tat-free proteins that manifested as SodA > SodC > SodB. When compared with an MG1655 wild-type strain, the growth of a ΔSodA mutant strain was found to be inhibited after paraquat treatment; the growth of ΔSodB and ΔSodC mutant strains was also slightly inhibited. The mRNA transcript level of genes encoding Tat-tagged proteins was higher than that of genes encoding Tat-free proteins. Furthermore, the α-helix and turn of Tat-tagged proteins were higher than those of Tat-free proteins, but the β-sheet and random coil content was lower. These results indicated that the incorporation of the Tat core peptide as a significant basic membrane transduction peptide in fusion proteins could increase mRNA transcripts and promote the high yield and soluble expression of heterologous proteins in *E. coli*.

INTRODUCTION

With the rapid development of biotechnology and bioengineering, *Escherichia coli* has become widely used as a conventional host to express heterologous proteins[1–3] because of its characteristic of including a clear genetic background, easy and inexpensive culturing, fast growth and production of adequate yields of protein.[4,5] Nevertheless, its application is still limited because of the degradation of heterologous proteins by cellular proteases[6] and/or the formation of inclusion bodies consisting of aggregates of misfolded protein.[7,8] To address this, several approaches, such as optimization of the genetic code,[9,10] increasing the transcription of the appropriate mRNA,[11] optimization of Shine–Dalgarno sequences[12,13] and alteration of the bacterial growth state,[14] have been analyzed. However, additional efficient strategies for obtaining high yields and soluble expression of heterologous proteins are urgently needed. For instance, Wu *et al.*[15] documented that HIV-1 virus-encoded transactivator of transcription (Tat) core peptide could promote high yields and soluble expression of

heterologous protein in *E. coli*, and showed an important application of this tag in the development of protein-associated drugs in the field of biology. Therefore, the development of a fusion tag for high heterologous protein yields and soluble expression may be a novel strategy.

The Tat protein is a member of the protein transduction domain superfamily.[16,17] Depending on the virus strain, the total length of Tat is 86–101 amino acids, comprising five domains, including the N-terminal activation domain, cysteine-rich domain, central domain, basic amino-rich domain and glutamine-rich domain.[18] As a cell-penetrating peptide, the Tat core peptide (YGRKKRRQRRR, pI = 12.8) is encoded in the genome of the HIV-1 virus and is enriched with basic amino acids. The basic domain of Tat is one of the shortest known cell-penetrating peptides leading to major accumulation in the nucleus that could deliver several heterologous macromolecules across biomembranes without the loss of their bioactivity, for example, enhanced green fluorescent protein,[19,20] p53,[21,22] cyclin-dependent kinase inhibitor p27,[23]

[1]Department of Biological Engineering, College of Pharmacy, Jilin University, Changchun, China; [2]Beijing Institute of Radiation Medicine, State Key Laboratory of Proteomics, Cognitive and Mental Health Research Center, Beijing, China and [3]Institute of Protein Research, Tongji University, Shanghai, China
[4]Co-first authors.
Correspondence: Dr W Yan, Department of Biological Engineering, College of Pharmacy, Jilin University, Changchun 130021, China.
E-mail: weiqunyan@jlu.edu.cn

and Cu-Zn superoxide dismutase (SOD).[24] Tat also exhibits a novel ability in providing high yield and soluble expression of heterologous protein in *E. coli*. To elucidate the mechanism of this process, the expression of *E. coli* Sod superfamily genes was examined in this study, including SodA, SodB and SodC. Consistent with our hypothesis, we found that the Tat core peptide, as a significant basic peptide, could also promote high yields and soluble expression of the heterologous proteins SodA, SodB and SodC in *E. coli*. As expected, the yields and solubility of Tat-tagged proteins were higher than that of Tat-free proteins, and the results of total SOD activity showed a similar pattern. The bacteria in which Tat-tagged proteins were overexpressed had a significant increase in anti-paraquat activity compared with those in which Tat-free proteins were expressed. In addition, Tat core peptides could increase mRNA transcriptional levels to promote high yields and soluble expression of heterologous proteins in *E. coli*. Furthermore, the amounts of α-helix and turn of Tat-tagged proteins were higher than those of Tat-free proteins, and the β-angle and random coil content was lower. This study provides a reference for the promotion of high yields and soluble expression of heterologous proteins for biotechnology and bioengineering.

MATERIALS AND METHODS
Construction of prokaryotic expression vectors
The prokaryotic expression vectors pET28b and pET28b-TAT were developed in our laboratory. The *E. coli* superoxide dismutase superfamily genes, including *sodA*, *sodB* and *sodC*, were chosen, and their complimentary DNA (cDNA) fragments were amplified from the *E. coli* BL21(DE3) strain by PCR using the primers shown in Table 1. PCR products were gel-extracted using a DNA extraction kit (TOYOBO, Osaka, Japan) and were then digested by the restriction enzymes *Bam*HI and *Xho*I (TaKaRa, Kusatsu, Japan), and subsequently subcloned into pET28b-Tat and pET28b to perform direct sequencing shown in Table 2.

Expression of the Tat-SodA/SodA, Tat-SodB/SodB and Tat-SodC/SodC peptides
The prokaryotic expression vectors pET28b-Tat-SodA/Tat-SodB/Tat-SodC and pET28b-SodA/SodB/SodC were transformed into the *E. coli* BL21(DE3) cells and incubated in a 37 °C incubator for 10–12 h until positive clones were visible. One clone from each group was picked and used to inoculate 5 ml of LB medium containing kanamycin (100 µg ml^{-1}), and allowed to grow at 37 °C (with shaking at 220 r.p.m.) to a logarithmic growth phase. The cells were then diluted to $OD_{600 \, nm} = 0.8$ and 5 ml of the cells ($\sim 3.2 \times 10^8$ cells) were inoculated in 150 ml of LB medium containing kanamycin (100 µg ml^{-1}) and incubated at 37 °C (with shaking at 220 r.p.m.) for 4–5 h until an $OD_{600 \, nm}$ of 0.6–1.0 was achieved. Expression of heterologous protein was induced by application of isopropyl β-D-1-thiogalactopyranoside (IPTG; 1 mM) at 30 °C for 6 h. The cells were harvested at 0, 2, 4 and 6 h after induction of protein expression and subjected to sonication in ice-cold phosphate-buffered saline, and then centrifuged at 12 000 r.p.m. for 10 min and filtered by passage through a 0.45 µm filter. Equal volumes of samples were then prepared and fractionated by SDS–polyacrylamide gel electrophoresis (SDS-PAGE) and western blotting. All experiments were repeated at least three times.

Western blot assay
The concentration of total proteins was measured with a BCA protein assay (Thermo Fisher Scientific Inc., Waltham, MA, USA) and equal amounts of samples (100 mg protein) were separated by electrophoresis using 15% polyacrylamide gels and transferred to a polyvinylidene difluoride membrane (GE Healthcare, Pittsburgh, PA, USA) following the manufacturer's instructions. The membranes were incubated with mouse-derived anti-GAPDH antibodies (1:500 in Tris-Buffered Saline with Tween-20 (TBST), Beyotime Institute of Biotechnology, Shanghai, China) and mouse-derived anti-His antibodies (1:3000 in TBST, Sigma, Santa Clara, CA, USA) for 1.5 h at room temperature, followed by incubation with horseradish peroxidase-conjugated goat anti-mouse secondary antibodies (1:5000 in TBST, Beyotime Institute of Biotechnology) at room temperature for 1 h. Reaction with chemiluminescence substrate luminal reagent (GE Healthcare) and exposure to X-ray film were used to examine the immunolabeled bands. The optical density of the bands was scanned and quantified with ImageJ software version 1.40g (http://rsb.info.nih.gov/ij/, NIH), and

Table 1 Primers used in this study

Primer name	Primer sequence
Escherichia coli	
pU: *Bam*HI-SodA	5'-CGC*GGATCC*GATGAGCTATACCCTG-3'
pD: *Xho*I-SodA	5'-CCG*CTCGAG*TTTTTTCGCCGCAA-3'
pU: *Bam*HI-SodB	5'-CGC*GGATCC*GATGTCATTCGAATT-3'
pD: *Xho*I-SodB	5'-CCG*CTCGAG*TGCAGCGAGATTTT-3'
pU: *Bam*HI-SodC	5'-CGC*GGATCC*GATGAAACGTTTTAG-3'
pD: *Xho*I-SodC	5'-CCG*CTCGAG*CTTAATTACACCAC-3'

Table 2 The characteristics of prokaryotic expression vectors used in this study

Vectors	Feature			
	cDNA fragment of the heterologous proteins	Size of fusion proteins coding region (bp)	Fusion proteins to be expressed	Size of fusion proteins to be expressed (kD)
pET28b-Tat-SodA	*Bam*HI-Tat-SodA-*Xho*I	743	Tat-SodA-6×His	27.23
pET28b-SodA	*Bam*HI-SodA-*Xho*I	756	6×His-SodA-6×His	27.69
pET28b-Tat-SodB	*Bam*HI-Tat-SodB-*Xho*I	693	Tat-SodB-6×His	25.42
pET28b-SodB	*Bam*HI-SodB-*Xho*I	705	6×His-SodB-6×His	25.87
pET28b-Tat-SodC	*Bam*HI-Tat-SodC-*Xho*I	597	Tat-SodC-6×His	21.83
pET28b-SodC	*Bam*HI-SodC-*Xho*I	609	6×His-SodC-6×His	22.29

histogram analysis using the Origin 9.5 software (http://www.originlab.com/).

Total superoxide dismutase activity assay

The above induced bacterial cells were harvested at 0, 2, 4 and 6 h, lysed with a cell lysis solution (50 mM Tris-HCl (pH 6.8), 15 mM NaCl, 5 mM EDTA, 0.5% Nonidet P-40 and 1 mM phenylmethanesulfonyl fluoride) on ice for 1 h, and centrifuged at 12 000 r.p.m. for 10 min at 4 °C. The concentration of total proteins was measured with a BCA protein assay (Thermo Fisher Scientific Inc.) and equal amounts of samples (100 mg protein) were used to detect the total superoxide dismutase activity using a Total Superoxide Dismutase Assay Kit (S0102, Beyotime Institute of Biotechnology) following the manufacturer's instructions.

Bacterial proliferation activity assay

The prokaryotic expression plasmids pET28b-Tat-Sod-A/Tat-SodB/Tat-SodC and pET28b-Sod-A/SodB/SodC were transformed into *E. coli* BL21(DE3) cells, and positive colonies were used to inoculate 5 ml LB medium containing kanamycin ($100\,\mu\text{g}\,\text{ml}^{-1}$), and grown to the logarithmic growth phase at 37 °C with shaking at 220 r.p.m. The cells were diluted to $OD_{600\,\text{nm}} = 0.4$, and the inducer IPTG and paraquat were added to final concentrations of 1 and 4 mM, respectively. Cells were then inoculated onto bacterial growth plates (350 μl per well) and incubated at 30 °C with shaking for 24 h; bacterial growth curves were measured with Bioscreen C (Oy Growth Curves Ab Ltd, Raisio, Finland). The data were analyzed using Microsoft Excel and Origin 9.5 software (http://www.originlab.com/). The growth curves of wild-type (MG1655), ΔSod-A, ΔSod-B and ΔSod-C strains were also measured and analyzed using the same method.

mRNA transcript level assay

The prokaryotic expression plasmids pET28b-Tat-SodA/Tat-SodB/Tat-SodC and pET28b-SodA/SodB/SodC were transformed into *E. coli* BL21(DE3) cells, and positive colonies used to inoculate 5 ml LB medium containing kanamycin ($100\,\mu\text{g}\,\text{ml}^{-1}$), and grown to logarithmic growth phase at 37 °C with shaking at 220 r.p.m. The cells were diluted to $OD_{600\,\text{nm}} = 0.8$, and 5 ml of cells ($\sim 3.2 \times 10^8$ cells) were inoculated into 150 ml LB medium containing kanamycin ($100\,\mu\text{g}\,\text{ml}^{-1}$) and incubated at 37 °C (with shaking at 220 r.p.m. for 4–5 h until $OD_{600\,\text{nm}} = 0.6$–1.0. The incubation was followed by IPTG (1 mM) induction at 30 °C for 6 h. Cells were harvested at 0, 3 and 6 h after induction of protein expression, and total RNA was extracted using the Total RNA extraction kit (TOYOBO) following the manufacturer's instructions.

Total RNA was used as a template in a reverse transcription reaction using a ReverTra Ace qPCR RT kit (TOYOBO), according to the manufacturer's instructions. The reaction mixtures, including 10 μl 2× loading buffer, 1.2 μl oligo(dT), 2 μl RNA, 0.2 μl MMLV and 6.6 μl DEPC-treated ddH$_2$O, were prepared and incubated at 65 °C for 30 min, followed by 42 °C for 30 min, and then 85 °C for 10 min. Subsequently, a total of 100 ng cDNA was used as the template in a quantitative real-time polymerase chain reaction (RT-qPCR) reaction using the SYBR Premix Ex Taq kit (TaKaRa, Kusatsu, Japan) according to the manufacturers' instructions. The reaction mixtures, including 10 μl 2× Master Mix, 0.08 μl forward primer, 0.08 μl reverse primer, 2 μl cDNA, 0.4 μl Taq DNA polymerase and 7.44 μl ddH$_2$O, were prepared and RT-qPCR performed according to the following program, using the primers shown in Table 1: one cycle of

95 °C for 3 min and then 40 cycles of 95 °C for 12 s, 62 °C for 30 s and 72 °C for 30 s. The results were analyzed using SDS 1.4 software (Applied Biosystems, Waltham, MA, USA) based on the $2^{-\Delta\Delta Ct}$ method, and histogram analysis was performed using Origin 9.5 software (http://www.originlab.com/).

Purification of proteins and secondary structure determination by circular dichroism (CD)

Cells correctly expressing Tat-tagged and Tat-free proteins were centrifuged at 12 000 r.p.m. for 10 min at 4 °C, and the pellets washed three times in phosphate-buffered saline. The suspensions were sonicated by ultrasonic disruptor and centrifuged at 4 °C, 12 000 r.p.m. for 10 min. The supernatants were added with adequate loading buffer (40 mM Na$_3$PO$_4$+0.5 M NaCl+10 mM imidazole, pH 7.4) to balance the pH and ionic strength. The protein samples were balanced with 4–5 column volumes of the Ni-NTA column; sample loading buffer was added until the loading baseline and then eluted from the column by imidazole-containing buffer with gradient concentration (40 mM Na$_3$PO$_4$+0.5 M NaCl+gradient concentration of imidazole, pH 7.4). Finally, the protein samples were collected according to the elution peak and identified by SDS-PAGE.

Subsequently, the concentrations of purified proteins, including Tat-SodA/SodA, Tat-SodB/SodB and Tat-SodC/SodC, were detected by a BCA protein assay (Thermo Fisher Scientific Inc.). The proteins were diluted to 5 mM in dilution buffer (10 mM Tris-HCl, pH 7.4), and secondary structure determination was performed using CD, followed by histogram analysis using Excel software.

Statistical analysis

The data were statistically analyzed using SPSS software (version 21.0, SPSS, Chicago, IL, USA, http://spss.en.softonic.com/); groups were compared using Student's *t*-test with significant differences defined as $P < 0.05$, whereas $P < 0.01$ represented a highly significant difference.

RESULTS

Preparation of the prokaryotic expression vectors, pET28b-Tat-Sod-A/Tat-SodB/Tat-SodC and pET28b-Sod-A/SodB/SodC

To elucidate the effect of Tat protein tags on heterologous protein expression in *E. coli*, the *E. coli* Sod superfamily genes, including *sodA*, *sodB* and *sodC*, were selected and cloned into pET28b-Tat and pET28b to produce the prokaryotic expression vectors pET28b-Tat-Sod-A/Tat-SodB/Tat-SodC and pET28b-Sod-A/SodB/SodC. The oligonucleotide sequences used to amplify these genes are shown in Table 1, and the schematic diagram of prokaryotic expression vector construction is shown in Figure 1. The correct construction of the prokaryotic expression vectors was confirmed by direct sequencing.

Tat tags promote high yield and soluble expression of SodA, SodB and SodC in *E. coli*

To investigate the effect of Tat tags on heterologous protein expression in *E. coli*, bacterial cells were harvested at different time points including 0, 2, 4 and 6 h post induction followed by preparation of total protein, including supernatant and pelleted proteins. As demonstrated by SDS-PAGE (Figure 2a) and quantification of SDS-PAGE analysis (Figure 2b), the expression level of Tat-tagged proteins (including Tat-SodA, Tat-SodB and

4
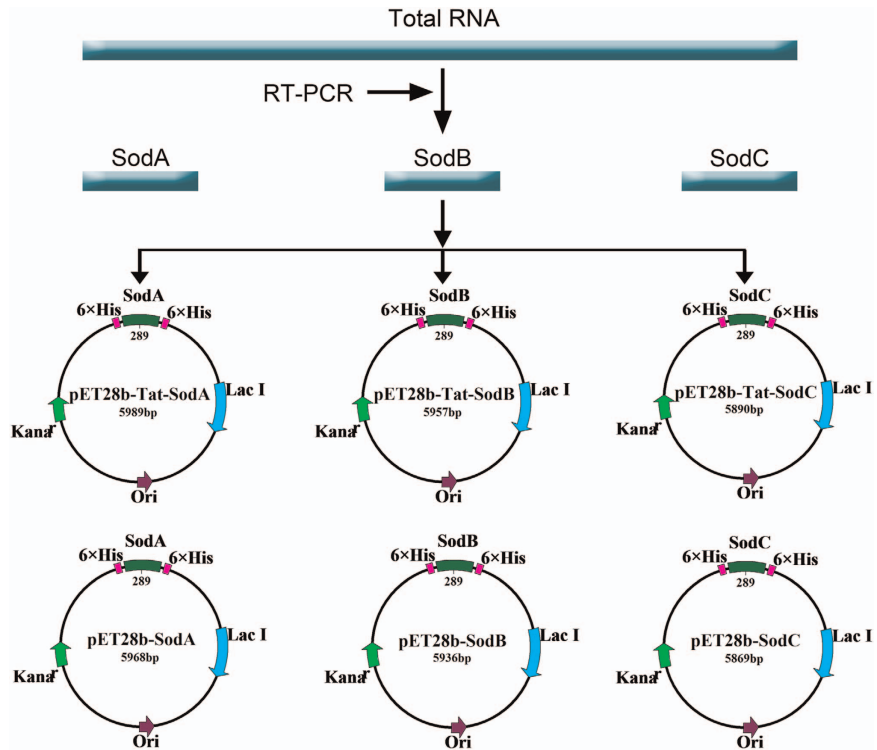
Figure 1 Schematic diagram of prokaryotic expression vector construction.
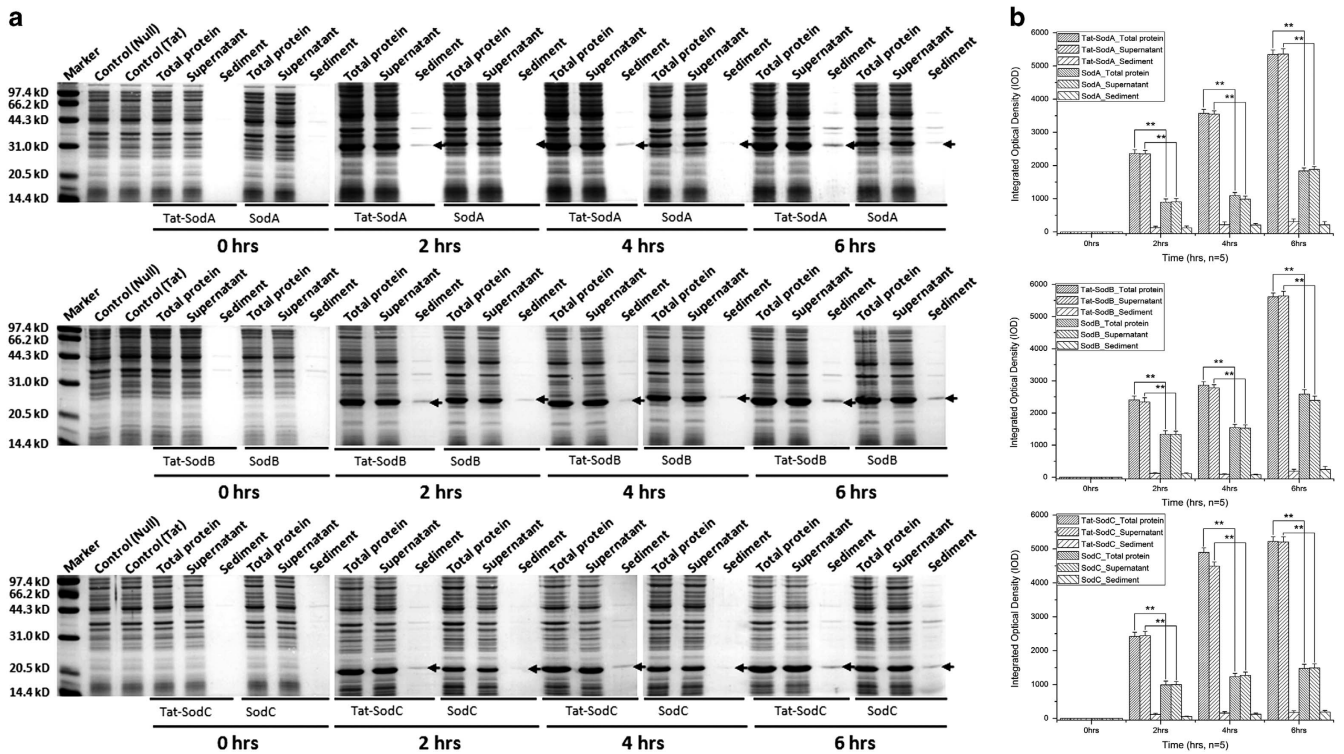
Figure 2 Protein expression level assay of transactivator of transcription (Tat)-tagged and Tat-free proteins by SDS–polyacrylamide gel electrophoresis (SDS-PAGE). (**a**) The protein expression level assay of Tat-tag proteins and Tat-free proteins by SDS-PAGE. (**b**) Quantitative analysis of the protein expression level of Tat-tagged and Tat-free proteins. Asterisks indicate statistically significant differences between the strains (**P<0.01).
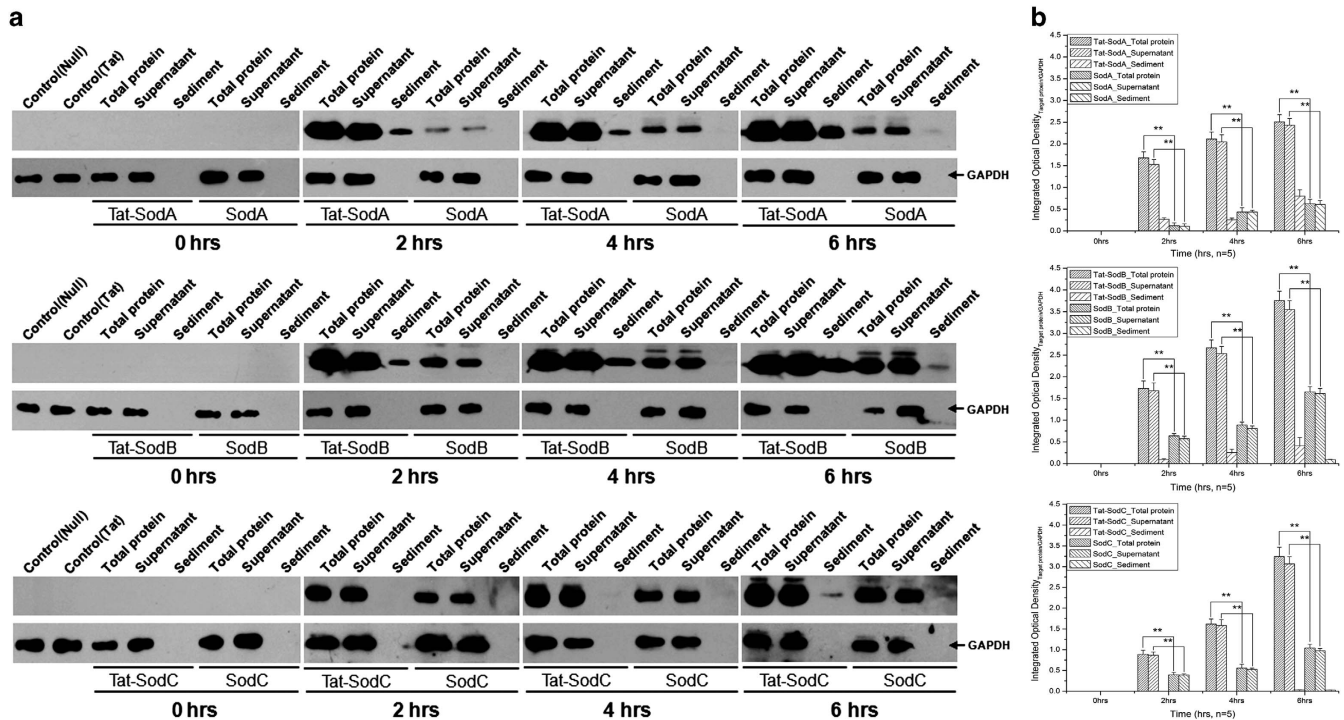
Figure 3 (a) Protein expression level assay of transactivator of transcription (Tat)-tagged and Tat-free proteins by western blotting. (b) Quantification of the protein expression level of Tat-tagged and Tat-free proteins. Asterisks indicate statistically significant differences between strains (**$P < 0.01$).
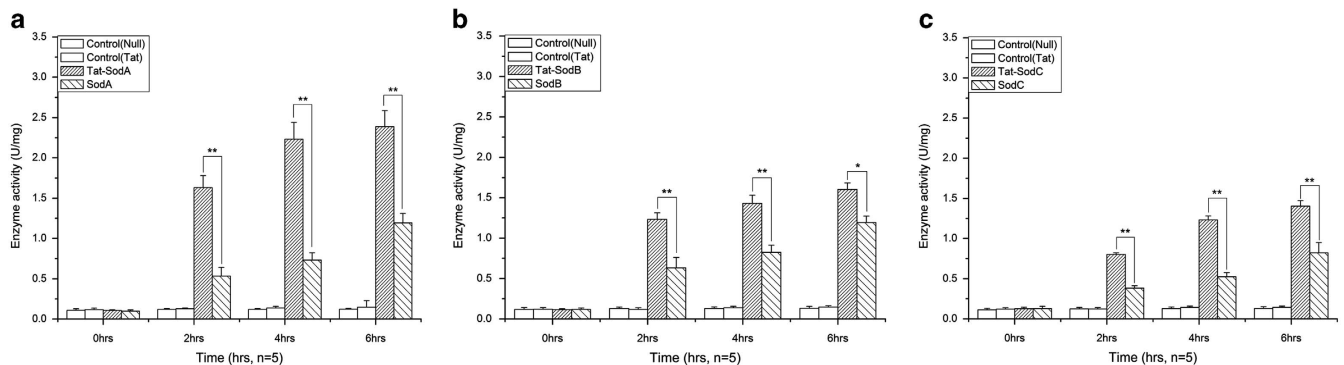

Figure 4 Total superoxide dismutase (SOD) activity assay of transactivator of transcription (Tat)-tagged and Tat-free proteins. (a) Total SOD activity assay of Tat-SodA and SodA. (b) Total SOD activity assay of Tat-SodB and SodB. (c) Total SOD activity assay of Tat-SodC and SodC. Asterisks indicate statistically significant differences in activity between groups (*$P < 0.05$; **$P < 0.01$).

Tat-SodC) and Tat-free proteins (including SodA, SodB and SodC) increased with increases in induction times. The expression of Tat-tagged proteins was significantly (2–5-fold) higher than that of Tat-free proteins (**$P < 0.01$), and both Tat-tagged and Tat-free proteins were mainly expressed in the supernatant, with only a small amount detected in the pellet fraction. Similarly, as western blot (Figure 3a) and quantification of western blot results (Figure 3b) demonstrated, the expression levels of Tat-tagged proteins were significantly higher than those of Tat-free proteins, and both Tat-tagged and Tat-free proteins were mainly expressed in the supernatant and only a small amount was detected in the pellet.

The total superoxide dismutase activity of Tat-tagged proteins was significantly higher than that of Tat-free proteins

To elucidate the total superoxide dismutase activity assay of Tat-tagged proteins and Tat-free proteins, cells were harvested at different time points including 0, 2, 4 and 6 h post induction. With increasing induction times, the total superoxide dismutase activity of Tat-tagged proteins (including Tat-SodA, Tat-SodB and Tat-SodC) and Tat-free proteins (including SodA, SodB and SodC) were significantly increased, and the activity of Tat-tagged proteins was significantly higher than that of Tat-free
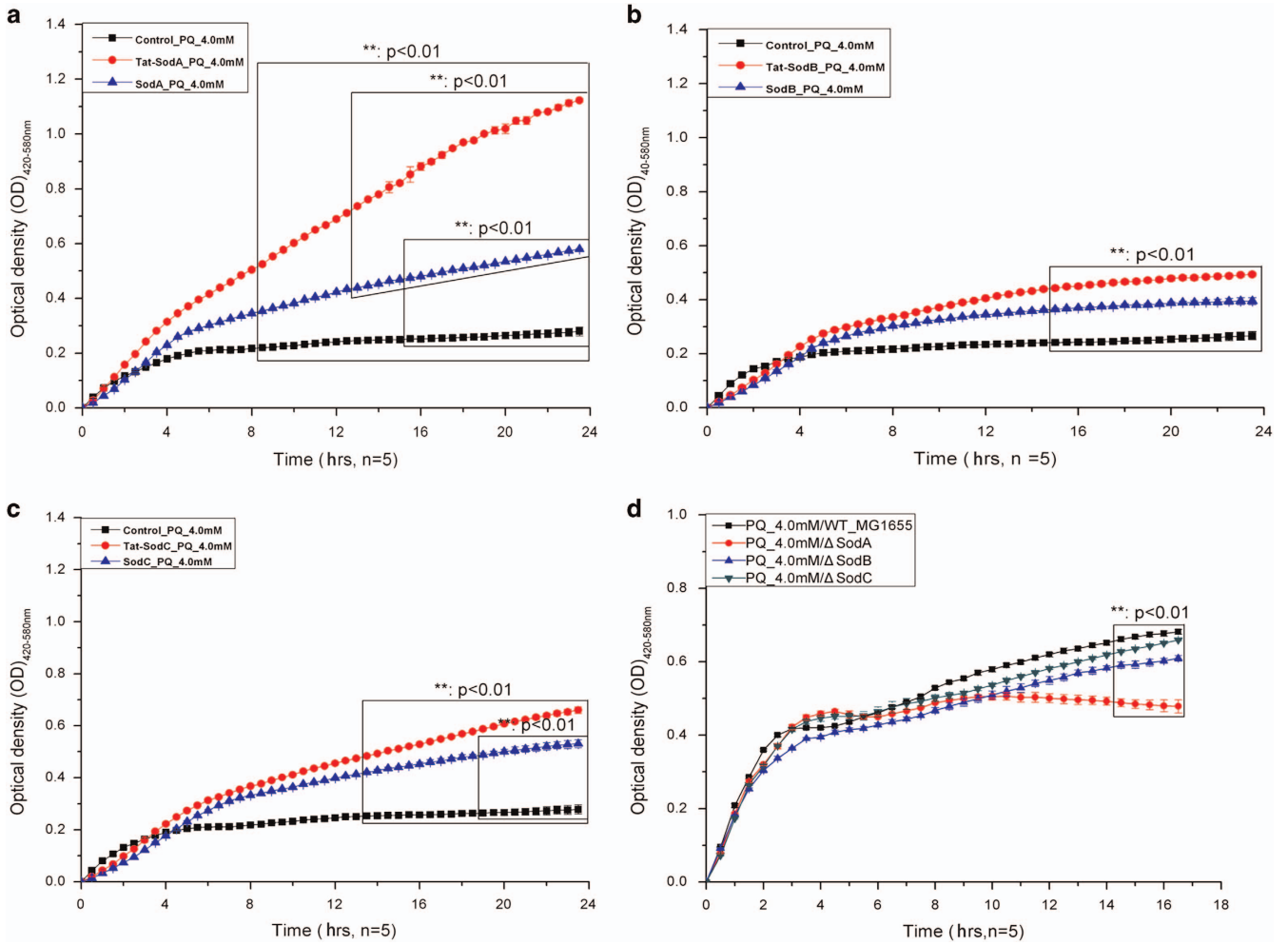
Figure 5 Bacterial anti-paraquat activity assay of *Escherichia coli* strains overexpressing SodA, SodB and SodC and of ΔSOD strain mutant strains. (**a**) The bacterial anti-paraquat activity assay of strains overexpressing Tat-SodA and SodA. (**b**) The bacterial anti-paraquat activity assay of strains overexpressing Tat-SodB and SodB. (**c**) The bacterial anti-paraquat activity assay of strains overexpressing Tat-SodC and SodC. (**d**) The bacterial anti-paraquat activity assay of ΔSodA, ΔSodB and ΔSodC mutant strains. Asterisks indicate statistically significant differences between strains (*$P<0.05$; **$P<0.01$). Tat, transactivator of transcription.
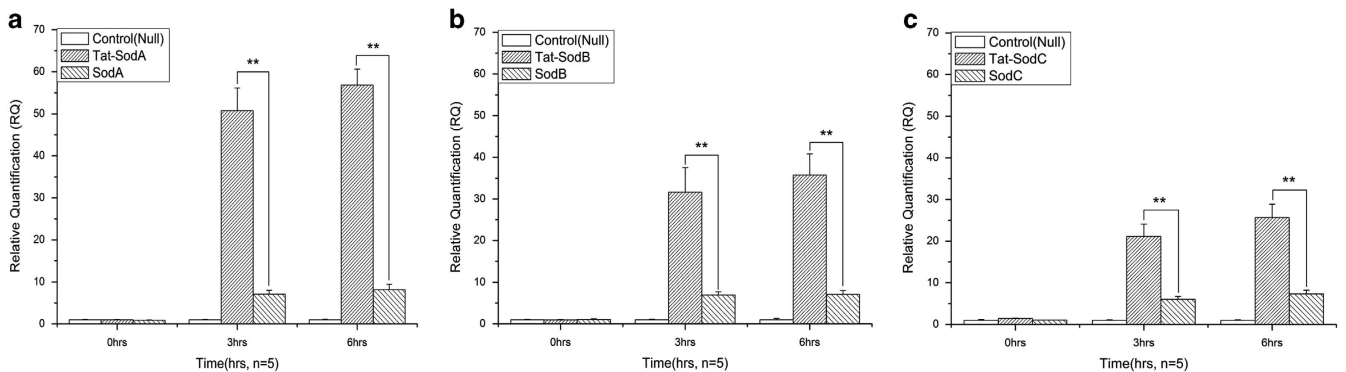
Figure 6 mRNA transcript level assay of transactivator of transcription (Tat)-tagged and Tat-free proteins by quantitative reverse transcription PCR (RT-qPCR). (**a**) The mRNA transcript level assay of Tat-SodA and SodA by RT-qPCR. (**b**) The mRNA transcript level assay of Tat-SodB and SodB by RT-qPCR. (**c**) The mRNA transcript level assay of Tat-SodC and SodC by RT-qPCR. Asterisks indicate statistically significant differences between transcript levels of proteins (*$P<0.05$; **$P<0.01$).

proteins (*$P < 0.05$; **$P < 0.01$, Figures 4a–c), and manifested as SodA > SodC > SodB.

Cells expressing Tat-tagged proteins had a significant increase in anti-paraquat activity compared with those expressing Tat-free proteins

To investigate the anti-paraquat activity of Tat- tagged and Tat-free proteins, the growth curves of *E. coli* cells overexpressing either Tat-tagged or Tat-free SOD proteins were measured over 24 h by Bioscreen C following paraquat treatment. After 4.0 mM paraquat treatment, the growth of cells expressing Tat-tagged proteins (including Tat-SodA, Tat-SodB and Tat-SodC) and Tat-free proteins (including SodA, SodB and SodC) significantly increased with increases in induction times in comparison with a control strain. In addition, the growth of cells expressing Tat-tagged proteins was higher than that of cells expressing Tat-free proteins (**$P < 0.01$, Figures 5a–c), and manifested as SodA > SodC > SodB. In addition, when compared with wild type (MG1655), the growth of ΔSodA was significantly inhibited after 6 h of culturing (**$P < 0.01$, Figure 5d), and that of ΔSodB and ΔSodC was slightly inhibited.

The presence of the Tat tag may increase the mRNA transcript level to promote high yield and soluble expression of heterologous proteins in *E. coli*

To reveal the mechanism by which the Tat-tag promotes heterologous protein expression in *E. coli*, the mRNA transcript levels of genes encoding Tat-tagged and Tat-free proteins were detected with a RT-qPCR assay. Cells were harvested at 0, 3 and 6 h after induction of protein expression and total RNA was extracted. As indicated by the RT-qPCR assay results, the mRNA transcript levels of *sod* genes encoding Tat-tagged (including Tat-SodA, Tat-SodB and Tat-SodC) and Tat-free proteins (including SodA, SodB and SodC) were significantly increased with increasing induction times, and mRNA transcript levels of genes encoding Tat-tagged proteins were 2–3-fold higher than those of genes encoding Tat-free proteins (**$P < 0.01$, Figures 6a–c).

The α-helix and turn of Tat-tagged proteins were significantly increased and the β-angle and random coil content was significantly decreased compared with that of Tat-free proteins

To elucidate the secondary structural changes of SodA, SodB and SodC when conjugated to the Tat tag, Tat-tagged proteins (including Tat-SodA, Tat-SodB and Tat-SodC) and Tat-free proteins (including SodA SodB, and SodC) were purified by Ni-NTA affinity chromatography. The chromatogram and the eluted proteins are shown in Figures 7a–f. Highly enriched, purified protein samples were prepared and the secondary structures of Tat-tagged and Tat-free proteins were determined by CD. As the histogram results showed, the secondary structures of the Tat-tagged proteins were different from those of the Tat-free proteins: the α-helix and turn of the Tat-tagged proteins were significantly increased compared with those of the Tat-free proteins (*$P < 0.05$; **$P < 0.01$, Figures 8a–c), whereas the β-angle and random coil content was significantly decreased (*$P < 0.05$; **$P < 0.01$, Figures 8a–c).

DISCUSSION

In our study, we demonstrated that the addition of a Tat tag promoted high yield and soluble expression of fusion proteins, such as SodA, SodB and SodC, in *E. coli*, and that the total SOD activity of Tat-tagged proteins was significantly higher than that of Tat-free proteins. In addition, the anti-paraquat activity of bacterial cells that overexpressed Tat-tagged SOD proteins was significantly higher than that of cells expressing Tat-free proteins, and the growth of a ΔSodA mutant strain was significantly inhibited by paraquat treatment. The mRNA transcript levels of genes encoding Tat-tagged proteins were significantly increased when compared with that of genes encoding Tat-free proteins. In addition, the α-helix and turn content of Tat-tagged proteins was significantly increased compared with that of Tat-free proteins, and the β-angle and random coil content was significantly decreased. These results indicated that the addition of a Tat tag could increase the mRNA transcript level and promote high yields and soluble expression of heterologous protein in *E. coli*, and could also change the secondary structure of its fusion protein.

The Tat protein encoded by HIV-1 is a member of the protein transduction domain family and is rich in basic amino acids. It is widely known that the full-length Tat peptide and the core domain (YGRKKRRQRRR) play important roles in the transduction of heterologous biological macromolecules such as proteins, peptides and nucleotides across all types of biomembranes *in vivo*.[25,26] The core region of Tat comprises a basic amino acid domain that contains six arginines and two lysines.[27–29] The current study showed that the transmembrane transduction efficiency of heterologous proteins was immediately affected by the number of amino acids. The prediction of protein secondary structure results showed that the α-helix structure could be formed in the Tat basic amino acid domain. CD analyses[30–32] and nuclear magnetic resonance imaging[33] have previously suggested that the basic amino acid domain of Tat protein transduction peptides could not form a functional structure. Secondary structures can be formally defined by the patterns of hydrogen bonds of the protein (such as α-helices, β-sheets, β-turns and random coils) that are observed in an atomic-resolution structure. The α-helices have particular significance in DNA-binding motifs, including helix-turn-helix motifs, leucine zipper motifs and zinc-finger motifs. The β-sheets are the most common protein structure element that is fundamental for correct folding. In our study, the CD results showed that the rates of α-helix and turn of Tat-tagged proteins were significantly increased compared with those for Tat-free proteins and that the β-angle and random coil content were significantly decreased, indicating that the addition of a Tat tag could change the secondary structure of its fusion protein. We hypothesized that these changes may be associated with high yields and soluble expression of these heterologous proteins, and our future work may focus on the effect of the
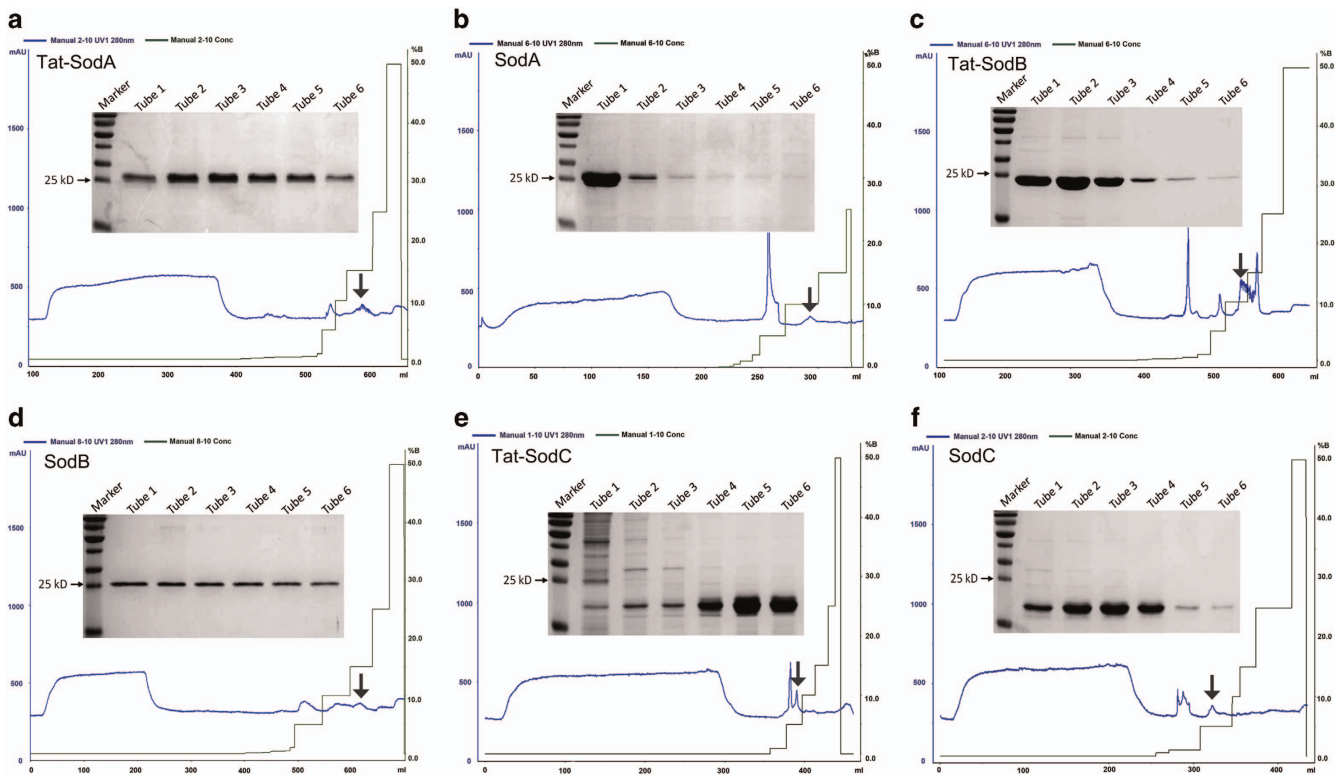
Figure 7 The chromatogram and eluted proteins identified by polyacrylamide gel electrophoresis (SDS-PAGE). (**a**) The chromatogram and the eluted proteins of Tat-SodA identified by SDS-PAGE. (**b**) The chromatogram and the eluted protein of SodA identified by SDS-PAGE. (**c**) The chromatogram and the eluted protein of Tat-SodB identified by SDS-PAGE. (**d**) The chromatogram and eluted protein of SodB identified by SDS-PAGE. (**e**) The chromatogram and eluted protein of Tat-SodC identified by SDS-PAGE. (**f**) The chromatogram and the eluted protein of SodC identified by SDS-PAGE. Tat, transactivator of transcription.
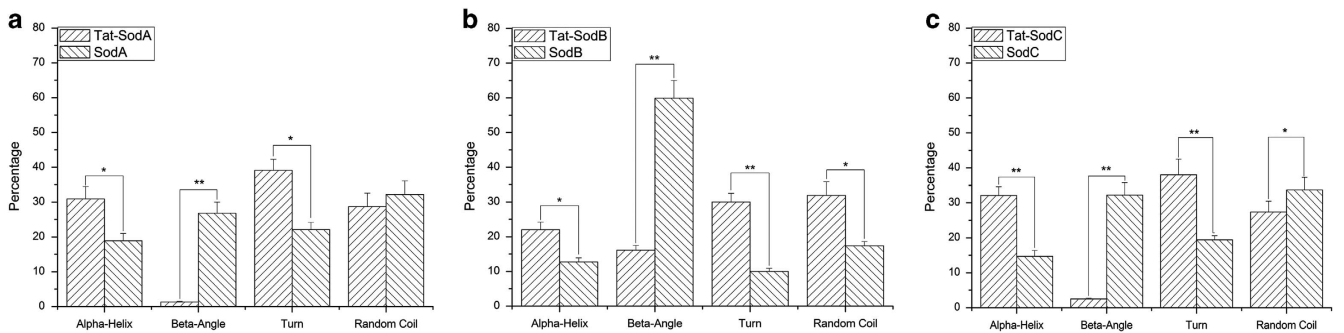
Figure 8 The secondary structure determination of transactivator of transcription (Tat)-tagged and Tat-free proteins by circular dichroism (CD). (**a**) The secondary structure determination of Tat-SodA and SodA by CD. (**b**) The secondary structure determination of Tat-SodB and SodB by CD. (**c**) The secondary structure determination of Tat-SodC and SodC by CD. Asterisks indicate statistically significant differences in the α-helix, turn, β-angle and random coil content between proteins (*$P<0.05$; **$P<0.01$).

number and distribution of basic amino acids in the Tat tag on expression of heterologous proteins.

SOD is one of the main enzymes involved in the detoxification of reactive oxygen species in the cell, playing a key antioxidant role in metabolism. There are three major families of SOD, represented in *E. coli* by SodA, SodB and SodC. In this study, the total SOD activity assay showed that the total SOD activity of Tat-tagged proteins was significantly higher than that of Tat-free proteins. Interestingly, the expression level of SodA was higher than those of SodB and SodC. We hypothesized that

SodA may be a multifunctional protein, especially in its antioxidant role. Furthermore, we analyzed the antioxidant potential of Tat-tagged proteins and Tat-free proteins: the bacterial proliferation activity assay showed, with increases in induction times, that the anti-paraquat activity of Tat-SodA, Tat-SodB and Tat-SodC was increased in comparison with that of SodA, SodB and SodC, respectively, following paraquat treatment. In addition, the growth of a ΔSodA mutant strain was inhibited compared with the wild type following paraquat treatment, whereas the ΔSodB and ΔSodC mutants showed

normal growth. Therefore, SodA likely plays an important role in anti-paraquat activity.

The regulation of mRNA transcript levels requires effective transcription initiation, elongation and termination, for which transcription factors are major determining factors for the high-yield expression of exogenous proteins. Effective transcription initiation directly affects the mRNA transcript level.[34–36] Strong promoters and regulatory sequences must be selected to enable high-yield heterologous mRNA expression, and usually such constructs as the pET series of vectors containing the T7 promoter are utilized.[37,38] Effective transcript elongation and termination were also key factors for achieving high expression levels of exogenous protein, as they affect the integrity and specificity of mRNA transcription.[39,40] The deletion of negative-regulatory elements, such as attenuators, was usually performed to increase the integrity of the mRNA transcript, and strong transcriptional termination could prevent nonspecific mRNA transcription. The mRNA stability and its secondary structure were also directly related to the number of translation products. Therefore, we hypothesized that the mRNA transcript levels of Tat-tagged proteins may be significantly changed and that this change may directly affect the expression of heterologous proteins. As expected, after 3 h of culturing, the mRNA transcript level was significantly increased, and slightly increased after 6 h of culturing. In addition, expression of Tat-tagged proteins was higher than that of Tat-free proteins, indicating that the Tat tag could promote high yields and soluble expression of heterologous protein in *E. coli*.

The current literature states that the expression of membrane proteins and disulfide bond-rich proteins is extremely limited when using prokaryotic expression systems. The question of how to refold inclusion body proteins containing many disulfide bonds into their natural conformations with high efficiency remains a challenging problem at present and is becoming one of the bottlenecks in the bioengineering field. As the present study suggests that the addition of the Tat core peptide could promote high yields and soluble expression of heterologous proteins in *E. coli*, we also carried out the following experiment on the expression of the membrane protein LSECTin that contains two disulfide bonds. Preliminary experimental data showed that Tat-LSECTin was mainly expressed in the cell lysate supernatant, in the form of inclusion bodies, and that the Tat-LSECTin protein expression yield was significantly higher than that of LSECTin alone.[41] These results support the idea that Tat core peptides also promote high yield and soluble expression of membrane proteins containing disulfide bonds, and indicate that addition of a Tat tag could promote high yield and soluble expression of proteins in *E. coli* regardless of whether they were membrane proteins or disulfide bond-rich proteins, thereby providing a novel tool for heterologous protein expression in *E. coli*.

Our data show that HIV-1 virus-encoded Tat core peptide could promote high yields and soluble expression of heterologous proteins in *E. coli*. The expressed Tat-tagged protein was mainly localized in the supernatant and exhibited increased enzyme activity compared with untagged proteins; cells that showed overexpression of Tat-tagged proteins had significantly increased anti-paraquat activity compared with those that did not. The mRNA transcript level of Tat-tagged proteins was higher than that of Tat-free proteins; the α-helix and turn of Tat-tagged proteins was significantly increased and their β-angle and random coil content was significantly decreased. These results indicated that the addition of a Tat tag could promote high yields and soluble expression of heterologous protein in *E. coli* via increases in the mRNA transcript levels and alteration in the secondary structure of the resultant fusion proteins, and suggested significant applications in heterologous protein expression for biotechnology and bioengineering.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

1 Yin J, Bao L, Tian H, Gao X, Yao W. Quantitative relationship between the mRNA secondary structure of translational initiation region and the expression level of heterologous protein in Escherichia coli. *J Ind Microbiol Biotechnol* 2015; **13**: 1–6.

2 Wang CH, Zhao TX, Li M, Zhang C, Xing XH. Characterization of a novel Acinetobacter baumannii xanthine dehydrogenase expressed in Escherichia coli. *Biotechnol Lett* 2016; **38**: 337–344.

3 Vargas-Cortez T, Morones-Ramirez JR, Balderas-Renteria I, Zarate X. Expression and purification of recombinant proteins in *Escherichia coli* tagged with a small metal-binding protein from Nitrosomonas europaea. *Protein Expr Purif* 2015; **118**: 49–54.

4 Wu N, He L, Cui P, Wang W, Yuan Y, Liu S et al. Ranking of persister genes in the same *Escherichia coli* genetic background demonstrates varying importance of individual persister genes in tolerance to different antibiotics. *Front Microbiol* 2015; **6**: 1003.

5 Tormo A, Fenoll C. Influence of the genetic background on cell division and cell lysis: behaviour of different *Escherichia coli* strains carrying the ts-52 or the ftsA-3 mutation. *Microbios* 1985; **42**: 111–117.

6 Wang L, Zhou YJ, Ji D, Lin X, Liu Y, Zhang Y et al. Identification of UshA as a major enzyme for NAD degradation in Escherichia coli. *Enzyme Microb Technol* 2014; **58**: 75–79.

7 Betton JM, Sassoon N, Hofnung M, Laurent M. Degradation versus aggregation of misfolded maltose-binding protein in the periplasm of Escherichia coli. *J Biol Chem* 1998; **273**: 8897–8902.

8 Missiakas D, Schwager F, Betton JM, Georgopoulos C, Raina S. Identification and characterization of HsIV HsIU (ClpQ ClpY) proteins involved in overall proteolysis of misfolded proteins in Escherichia coli. *EMBO J* 1996; **15**: 6899–6909.

9 Hu S, Wang M, Cai G, He M. Genetic code-guided protein synthesis and folding in Escherichia coli. *J Biol Chem* 2013; **288**: 30855–30861.

10 Hoesl MG, Budisa N. Recent advances in genetic code engineering in Escherichia coli. *Curr Opin Biotechnol* 2012; **23**: 751–757.

11 Lin HH, Lin CH, Hwang SM, Tseng CP. High growth rate downregulates fumA mRNA transcription but is dramatically compensated by its mRNA stability in Escherichia coli. *Curr Microbiol* 2012; **64**: 412–417.

12 Jin W, Xu X, Jiang L, Zhang Z, Li S, Huang H. Putative carotenoid genes expressed under the regulation of Shine-Dalgarno regions in *Escherichia*

10

coli for efficient lycopene production. *Biotechnol Lett* 2015; **37**: 2303–2310.

13 Stefan A, Schwarz F, Bressanin D, Hochkoeppler A. Shine-Dalgarno sequence enhances the efficiency of lacZ repression by artificial anti-lac antisense RNAs in Escherichia coli. *J Biosci Bioeng* 2010; **110**: 523–528.

14 Raina S, Mabey L, Georgopoulos C. The *Escherichia coli* htrP gene product is essential for bacterial growth at high temperatures: mapping, cloning, sequencing, and transcriptional regulation of htrP. *J Bacteriol* 1991; **173**: 5999–6008.

15 Wu Y, Ren C, Gao Y, Hou B, Chen T, Zhang C. A novel method for promoting heterologous protein expression in *Escherichia coli* by fusion with the HIV-1 TAT core domain. *Amino Acids* 2010; **39**: 811–820.

16 Zhong J, Kang J, Wang X, Jiang W, Liao H, Yuan J. TAT-OSBP-1-MKK6(E), a novel TAT-fusion protein with high selectivity for human ovarian cancer, exhibits anti-tumor activity. *Med Oncol* 2015; **32**: 118.

17 Zhang X, Li Y, Cheng Y, Tan H, Li Z, Qu Y *et al.* Tat PTD-endostatin: a novel anti-angiogenesis protein with ocular barrier permeability via eye-drops. *Biochim Biophys Acta* 2015; **1850**: 1140–1149.

18 Frankel AD, Pabo CO. Cellular uptake of the tat protein from human immunodeficiency virus. *Cell* 1988; **55**: 1189–1193.

19 Cermenati G, Terracciano I, Castelli I, Giordana B, Rao R, Pennacchio F *et al.* The CPP Tat enhances eGFP cell internalization and transepithelial transport by the larval midgut of Bombyx mori (Lepidoptera, Bombycidae). *J Insect Physiol* 2011; **57**: 1689–1697.

20 Caron NJ, Torrente Y, Camirand G, Bujold M, Chapdelaine P, Leriche K *et al.* Intracellular delivery of a Tat-eGFP fusion protein into muscle cells. *Mol Ther* 2001; **3**: 310–318.

21 Zheng R, Yao Q, Xie G, Du S, Ren C, Wang Y *et al.* TAT-ODD-p53 enhances the radiosensitivity of hypoxic breast cancer cells by inhibiting Parkin-mediated mitophagy. *Oncotarget* 2015; **6**: 17417–17429.

22 Yoon CH, Kim SY, Byeon SE, Jeong Y, Lee J, Kim KP *et al.* p53-derived host restriction of HIV-1 replication by protein kinase R-mediated Tat phosphorylation and inactivation. *J Virol* 2015; **89**: 4262–4280.

23 Dai L, Liu Y, Liu J, Wen X, Xu Z, Wang Z *et al.* A novel cyclinE/cyclinA-CDK inhibitor targets p27(Kip1) degradation, cell cycle progression and cell survival: implications in cancer therapy. *Cancer Lett* 2013; **333**: 103–112.

24 Chou CM, Huang CJ, Shih CM, Chen YP, Liu TP, Chen CT. Identification of three mutations in the Cu,Zn-superoxide dismutase (Cu,Zn-SOD) gene with familial amyotrophic lateral sclerosis: transduction of human Cu,Zn-SOD into PC12 cells by HIV-1 TAT protein basic domain. *Ann NY Acad Sci* 2005; **1042**: 303–313.

25 Yan C, Gu J, Hou D, Jing H, Wang J, Guo Y *et al.* Improved tumor targetability of Tat-conjugated PAMAM dendrimers as a novel nanosized anti-tumor drug carrier. *Drug Dev Ind Pharm* 2015; **41**: 617–622.

26 Szyrwiel L, Shimura M, Shirataki J, Matsuyama S, Matsunaga A, Setner B *et al.* A novel branched TAT(47-57) peptide for selective Ni(2+) introduction into the human fibrosarcoma cell nucleus. *Metallomics* 2015; **7**: 1155–1162.

27 Zhao W, Liu K, Chen S, Zhu MM, Lv H, Hu J *et al.* Polyethylenimine derivate conjugated with RGD-TAT-NLS as a novel gene vector. *Biomed Mater Eng* 2014; **24**: 1933–1939.

28 Yu R, Yang Y, Cui Z, Zheng L, Zeng Z, Zhang H. Novel peptide VIP-TAT with higher affinity for PAC1 inhibited scopolamine induced amnesia. *Peptides* 2014; **60**: 41–50.

29 Xu R, Dong Y, Wang L, Tao X, Sun A, Wei D. TAT-RhoGDI2, a novel tumor metastasis suppressor fusion protein: expression, purification and functional evaluation. *Appl Microbiol Biotechnol* 2014; **98**: 9633–9641.

30 Nolandt OV, Walther TH, Roth S, Burck J, Ulrich AS. Structure analysis of the membrane protein TatC(d) from the Tat system of B. subtilis by circular dichroism. *Biochim Biophys Acta* 2009; **1788**: 2238–2244.

31 Loret EP, Georgel P, Johnson WC Jr, Ho PS. Circular dichroism and molecular modeling yield a structure for the complex of human immuno-deficiency virus type 1 trans-activation response RNA and the binding region of Tat, the trans-acting transcriptional activator. *Proc Natl Acad Sci USA* 1992; **89**: 9734–9738.

32 Loret EP, Vives E, Ho PS, Rochat H, Van Rietschoten J, Johnson WC Jr. Activating region of HIV-1 Tat protein: vacuum UV circular dichroism and energy minimization. *Biochemistry* 1991; **30**: 6013–6023.

33 Watkins JD, Campbell GR, Halimi H, Loret EP. Homonuclear 1H NMR and circular dichroism study of the HIV-1 Tat Eli variant. *Retrovirology* 2008; **5**: 83.

34 Zuo Y, Steitz TA. Crystal structures of the *E. coli* transcription initiation complexes with a complete bubble. *Mol Cell* 2015; **58**: 534–540.

35 Yang X, Chang HR, Yin YW. Yeast mitochondrial transcription factor Mtf1 determines the precision of promoter-directed initiation of RNA polymerase Rpo41. *PLoS ONE* 2015; **10**: e0136879.

36 Slobodin B, Agami R. Transcription initiation determines its end. *Mol Cell* 2015; **57**: 205–206.

37 Zhu D, Liu F, Xu H, Bai Y, Zhang X, Saris PE *et al.* Isolation of strong constitutive promoters from Lactococcus lactis subsp. lactis N8. *FEMS Microbiol Lett* 2015; **362**. pii: fnv107.

38 Schlabach MR, Hu JK, Li M, Elledge SJ. Synthetic design of strong promoters. *Proc Natl Acad Sci USA* 2010; **107**: 2538–2543.

39 Kriner MA, Groisman EA. The bacterial transcription termination factor Rho coordinates Mg homeostasis with translational signals. *J Mol Biol* 2015; **427**: 3834–3849.

40 Tate J, Gollnick P. The role of vaccinia termination factor and cis-acting elements in vaccinia virus early gene transcription termination. *Virology* 2015; **485**: 179–188.

41 Dong G, Wang C, Wu Y, Cong J, Cheng L, Wang M *et al.* Tat peptide-mediated soluble expression of the membrane protein LSECtin-CRD in Escherichia coli. *PLoS ONE* 2013; **8**: e83579.