Enhanced solubility of heterologous proteins by fusion expression using stress-induced Escherichia coli protein, Tsf

Kyung-Yeon Han, Jong-Am Song, Keum-Young Ahn, Jin-Seung Park, Hyuk-Seong Seo & Jeewon Lee

Department of Chemical and Biological Engineering, Korea University, Seoul, South Korea

Correspondence: Jeewon Lee, Department of Chemical and Biological Engineering, Korea University, Anam-Dong 5-1, Sungbuk-Ku, Seoul 136-713, South Korea. Tel.: +82 2 3290 3304; fax: +82 2 926 6102; e-mail: leejw@korea.ac.kr

Received 22 March 2007; revised 11 May 2007; accepted 24 May 2007.
First published online 3 July 2007.
DOI:10.1111/j.1574-6968.2007.00824.x

Abstract

Through two-dimensional electrophoresis, Escherichia coli proteome response to a protein denaturant, guanidine hydrochloride, was analyzed and elongation factor Ts (Tsf) detected as a stress-induced protein. Many host proteins aggregated, or their synthesis levels decreased significantly under conditions of protein denaturation as 34 out of 699 soluble proteins knocked out and 63 proteins decreased by over 2.5-fold. Interestingly, the expression level of Tsf increased 1.61-fold compared with a nonstress condition. Contrary to direct expression, various heterologous proteins were solubly expressed in E. coli when subjected to N-terminus fusions of Tsf. Owing most likely to an intrinsic high folding efficiency, Tsf seemed to play critical roles in sequestering interactive surfaces of heterologous proteins from nonspecific protein–protein interactions leading to formation of inclusion bodies. It has been also demonstrated that Tsf is effective in aiding the production of a biologically active bacterial cutinase, which could be of interest to biotechnology and commercial applications.

Keywords
Escherichia coli BL21 (DE3); proteome; stress response; Tsf; solubility enhancer.

Introduction

The major advantages of the Escherichia coli system for recombinant protein expression are established genetic manipulation techniques, rapid large-scale production capability for industrial and medical proteins, and a relatively low cost of maintenance (Davis et al., 1999; Baneyx & Mujacic, 2004). In spite of many successful applications as a host bacterial strain for heterologous therapeutic proteins in the biotechnology industry (Davis et al., 1999), the formation of undesirable insoluble protein aggregates, known as inclusion bodies, is a critical bottleneck in bioactive and functional recombinant protein production (De Bernardez Clark, 1998; Lillie et al., 1998; Christendat et al., 2000; Nallamsetty & Waugh, 2006). Precise timing during protein synthesis and the co- or posttranslational folding is required to reach a native three-dimensional structure with timely interactions between nascent polypeptides and folding chaperones to prevent irreversible aggregation and misfolding (Baneyx & Mujacic, 2004; Maier et al., 2005; Kaiser et al., 2006). Several elongation factors, such as elongation factor Tu (EF-Tu), have also been reported to play a role in DNA repair, replication, transcription, and RNA processing (Malki et al., 2002). Moreover, EF-Tu is involved in the refolding of unfolded proteins similar to molecular chaperones (Caldas et al., 1998) and displays protein disulfide isomerase activity in vitro (Richarme, 1998). In addition, Caldas et al. (2000) reported that elongation factor G (EF-G) and initiation factor 2 (IF2) also possess chaperone-like properties that facilitate protein folding.

In the present study, elongation factor Ts (Tsf) was identified as an aggregation-resistant protein through an E. coli proteome-wide analysis in response to a protein denaturant. The synthesis level of soluble Tsf was significantly increased even in the presence of a strong protein-denaturing environment induced by guanidine hydrochloride (GdnHCl). Tsf has been used as an N-terminus fusion partner protein in the synthesis of aggregation-prone proteins (see Table 1.) in E. coli cytoplasm. Tsf dramatically increased the cytoplasmic solubility of the heterologous proteins, most likely because of the activity of a cis-acting folding enhancer. Moreover, the efficacy of Tsf has been demonstrated in the production of a biologically active bacterial enzyme (cutinase) that may yield numerous biotechnology and commercial applications.
Table 1. Target heterologous proteins and Tsf fusion mutants used in this study

<table>
<thead>
<tr>
<th>Aggregation-prone heterologous proteins</th>
<th>Fusion mutants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human minipro-insulin (mp-INS)</td>
<td>Tsf::mp-INS</td>
</tr>
<tr>
<td>Human epidermal growth factor (EGF)</td>
<td>Tsf::EGF</td>
</tr>
<tr>
<td>Human prepro-ghrelin (ppGRN)</td>
<td>Tsf::ppGRN</td>
</tr>
<tr>
<td>Human interleukin-2 (hIL-2)</td>
<td>Tsf::hIL-2</td>
</tr>
<tr>
<td>Human activation induced cytidine deaminase (AID)</td>
<td>Tsf::AID</td>
</tr>
<tr>
<td>Deletion mutant of human glutamate decarboxylase (GAD448–585)</td>
<td>Tsf::GAD448–585</td>
</tr>
<tr>
<td>Human ferritin light chain (hFTN-L)</td>
<td>Tsf::hFTN-L</td>
</tr>
<tr>
<td>Human granulocyte colony-stimulating factor (G-CSF)</td>
<td>Tsf::G-CSF</td>
</tr>
<tr>
<td>Human cold autoinflammatory syndrome 1 Protein (NALP3) NACHT</td>
<td>Tsf::NACHT</td>
</tr>
<tr>
<td>domain (NACHT)</td>
<td>Tsf::CUT</td>
</tr>
<tr>
<td>Human cutinase (CUT)</td>
<td>Tsf::CUT</td>
</tr>
</tbody>
</table>

Materials and methods

Bacterial strain and plasmids

Escherichia coli strain BL21 (DE3) [F- ompT hsdSB (rB mB')] was selected under both nonstress and GdnHCl-stress conditions for the two-dimensional electrophoresis (2-DE) analysis. To assess recombinant protein expression, PCR amplification using appropriate primers was performed and each recombinant gene and various fusion mutants (see Table 1.) were inserted into the Ndel–HindIII site of pT7-7 plasmids (Novagen) to construct an expression vector. After complete DNA sequencing of all gel-purified plasmid vectors, the E. coli strain BL21 (DE3) [F- ompT hsdSB (rB mB')] was transformed with the plasmid expression vectors, and ampicillin-resistant transformants were subsequently selected using Luria–Bertani (LB) agar plates supplemented with ampicillin (100 mg L⁻¹).

Recombinant E. coli culture and gene expression

For shake flask experiments, 250 mL Erlenmeyer flasks containing LB media at 100 mg ampicillin L⁻¹ of culture (37 °C) were used. When the culture turbidity (OD₆₀₀ nm) reached 0.5, gene expression was induced with the addition of isopropyl-β-D-thiogalactopyranoside (1 mM) and after a further 3–4 h of cultivation, the recombinant cells were harvested by centrifugation at 16 000 × g for 10 min. The isolated inclusion bodies, if any, were washed twice with 1% Triton X-100. Cell-free supernatants and the washed inclusion bodies were subjected to polyacrylamide gel electrophoresis (PAGE) analysis, using 14% Tris-glycine precast gels (Novex, San Diego, CA). Coomassie-stained protein bands were ultimately scanned and analyzed by a densitometer (Duoscan T1200, Bio-Rad, Hercules, CA).

Sample preparation for proteome analysis and 2-DE

Flask culture conditions were the same as those for recombinant gene expressions. When the culture turbidity (OD₆₀₀ nm) reached 0.5, GdnHCl (100 mM) was added to the culture media to induce the stress condition. After a further 3-h cultivation, the cells were harvested by centrifugation at 3420 × g for 15 min (4 °C) and then washed twice with 40 mM Tris buffer (pH 8.0). Cell pellets were resuspended in 500 μL of lysis buffer (8 M Urea/4% (w/v) 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS)/40 mM Tris/Protease inhibitor cocktail; Roche Diagnostics GmbH, Mannheim, Germany) and disrupted by sonication. After sonication, the cell debris and the aggregated proteins were removed by centrifugation at 13 680 × g for 60 min (4 °C). The protein samples were resuspended in rehydration solution [8 M Urea/0.5% (v/v) Triton X-100/0.005% Orange G/1% (w/v) dithiothreitol/1% (v/v) carrier ampholyte, pH3-10; final volume, 320 μL]. Urea, CHAPS, Tris, dithiothreitol, orange G, Triton X-100, and sodium dodecyl sulfate (SDS) were purchased from Sigma (St Louis, MO). The first dimension of 2-DE was performed on a IPGphor Electrophoresis System (Amersham Bioscience, Uppsala, Sweden) at 20 °C. Linear immobilized pH gradient (IPG) gel strips, pH 4–7, were rehydrated for 12 h. Isoelectric focusing of rehydrated protein samples (45 μg) was performed at 500 V for 2 h, at 1000 V for 30 min, at 2000 V for 30 min, at 4000 V for 30 min, and finally maintained at 8000 V until 70 000 V h was achieved. For the second dimension, the IPG gel strips were equilibrated for 15 min in equilibration solution [50 mM Tris/HCl, pH 8.8/6 M Urea/30% (v/v) glycerol/2% (w/v) SDS/trace element bromophenol blue] in 1% dithiothreitol for 15 min, followed by 2.5% (w/v) iodoacetamide for 15 min. The second-dimensional separation was performed using a PROTEAN II Xi cell system (Bio-Rad) in a cold chamber at 4 °C on 12.5% polyacrylamide gels. SDS/PAGE was performed at 30 mA gel⁻¹ for 12 h. The silver-stained gels were scanned using a UMAX powerlook 1100 scanner.
MALDI–TOF-MS analysis and protein identification

Peptide mass fingerprints of protein spots were obtained using a MALDI–TOF-MS system (Voyager DE-STR, PE Biosystem, Framingham, MA, performed by the Korean Basic Science Institute, Seoul) and analyzed using MS-Fit (http://prospector.ucsf.edu/).

Bioactivity assay

The enzyme activity of the recombinant cutinase fusion mutant was assessed as described below. The hydrolysis reactions occurred in 96-well microplates at 37 °C for 15 min where each well contained 200 μL of enzyme/substrate solution (solution A) comprised of phosphate buffer, 106.7 μL (0.1 M, pH 8.0), Triton X-100 solution 13.3 μL (4 g L⁻¹), enzyme (cutinase) solution 13.3 μL, and substrate [p-nitrophenyl butyrate (PNB) or p-nitrophenyl palmitate (PNP)] reagent solution 66.7 μL. The reaction was initiated by adding 66.7 μL of substrate reagent solution to each well in the 96-well microplate. Absorbance changes (ΔOD₄₁₅ nm per minute) were measured using a Bio-Rad microplate reader (Tecan, Austria), and solution A above with no enzyme solution was used as blank. The eight wells in each of 12 columns of the 96-well microplate represented the same reaction condition and contained equal contents of enzyme and substrate. From the absorbance changes measured at each column, an average absorbance for a specific reaction condition could then be calculated.

Results and discussion

Proteome response upon GdnHCl-induced stress and identification of an aggregation-resistant protein

Solubility enhancement and avoidance of protein misfolding and aggregation during recombinant protein expression have been long-standing issues in the biotechnology industry that have yet to be satisfactorily settled. The main interest has been long-standing issues in the biotechnology industry. Solubility enhancement and avoidance of protein misfolding are key factors in ensuring the proper folding of newly synthesized polypeptides in the E. coli cytoplasm.

Investigated after GdnHCl, a potent protein denaturant, was added to growing bacterial cultures when the OD₆₀₀ nm of the bacterial culture reached 0.5. As shown in Fig. 1, the cell growth stopped when the culture OD₆₀₀ nm reached 3.41 under the stress condition by GdnHCl, whereas the nonstressed cells continued to grow up to the culture OD₆₀₀ nm of 4.40. Thirty-four protein spots out of the 699 soluble proteins that were present in the nonstress proteome were totally knocked out and disappeared under the stress condition, while the synthesis level of 63 proteins out of the 699 soluble proteins decreased over 2.5-fold. Several molecular chaperones including DnaK, ClpB, and GroEL showed significant expression-level increases in response to protein denaturation and the aggregation induced by GdnHCl (data not shown). It is reasonable to presume that activation of molecular chaperone synthesis implies that under stress conditions, proper folding of newly synthesized polypeptides in the E. coli cytoplasm is severely hindered, and that proteins would be denatured without other folding enhancers. Nevertheless, the synthesis level of various functional proteins increased more than 1.5-fold compared with nonsstress conditions. It is interesting to note that about a 1.61-fold increase in the expression level (i.e. intensity of the protein spot estimated through the analysis of 2-DE gel image) of Tsf was observed even under the stress condition (Fig. 2, Table 2). In addition to the increased level of Tsf expression upon addition of GdnHCl, it was observed that thiol-specific oxidant 2-hydroxyethyl disulfide (10 mM) also led to a 2.1-fold increase in the expression of Tsf (data not shown). Tsf is a steric chaperone that can act as a folding template for the formation of correctly folded EF-Tu (Krab et al., 2001). EF-G and IF2 bind to the C-terminal domain of ribosomal protein [RimL (acyetylating enzyme for N-terminal ribosomal protein L7/L12) accession number P13857].
a protein that has structural homology with Tsf (Helgstrand et al., 2007). Such a similarity among proteins implies the possibility of either direct or indirect interactions of Tsf with EF-G and IF-2. Interestingly, EF-G and IF2 are also involved in interactions involving stress-induced unfolded proteins for protein renaturation and refolding similar to molecular chaperones (Caldas et al., 2000). As shown in Fig. 2, the synthesis level of Tsf, a member of elongation factors with chaperone-like functional characteristics, was significantly increased in response to a protein denaturant (GdnHCl), and therefore Tsf may be regarded as a stress-responsive and aggregation-resistant protein. Tsf might play an efficient role as a folding enhancer in the E. coli cytoplasm, probably because of its intrinsic high folding efficiency, when used as a fusion partner for the overexpression of heterologous proteins.

**Expression of aggregation-prone heterologous proteins with N-terminus fusion of Tsf as a cis-acting folding enhancer**

Tsf was used as an N-terminus fusion expression partner for the synthesis of several heterologous proteins [minipro-insulin (mp-INS), human epidermal growth factor (EGF), human prepro-ghrelin (ppGRN), human interleukin 2 (hIL-2), human activation-induced cytidine deaminase (AID), deletion mutant of human glutamate decarboxylase (GAD448–585), human ferritin light chain (hFTN-L), human granulocyte colony-stimulating factor (G-CSF), human cold autoinflammatory syndrome 1 protein (NALP3) NACHT domain (NACHT), and Pseudomonas putida cutinase (CUT)]. Synthesis of the hybrid protein NH2-[Tsf]::(heterologous protein)-COOH in E. coli cytoplasm was attempted. All outcomes of direct expression of the foreign proteins mentioned above without an N-terminus folding enhancer resulted in the formation of aggregated inclusion bodies and almost negligible solubility (Fig. 3). On the other hand, when expressed with the fusion of Tsf, the solubility of all the heterologous proteins dramatically improved (Fig. 3). These results indicated that E. coli Tsf as the fusion expression partner yielded a highly effective cis-acting folding enhancer in an E. coli system. As shown in Fig. 3, the N-terminus fusion of E. coli Tsf never decreased the synthesis level of the heterologous proteins at all and even led to a significant increase in the synthesis level in the case of fusion expression.

**Table 2. Result of identification of aggregation-resistant protein, Tsf**

<table>
<thead>
<tr>
<th>Gene name*</th>
<th>ExPASy accession no.*</th>
<th>Protein name*</th>
<th>pl/Mw (kDa)</th>
<th>Sequence coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tsf</td>
<td>P0A6P1</td>
<td>Elongation factor Ts</td>
<td>5.22/30.29</td>
<td>5.15/30.62</td>
</tr>
</tbody>
</table>

*Gene name, accession number, and protein name are based on the data from ExPASy Proteomics Server (http://www.expasy.org/).

1Theoretical values of pl and Mw were estimated using ‘Compute pl/Mw tool’ (http://www.expasy.org/tools/pi_tool.html).

2Experimental values of pl and Mw were estimated through 2-DE gel image analysis in this study.
of mp-INS and G-CSF. Therefore, it seems obvious that the enhanced solubility by Tsf fusion did not result from the reduced synthesis level of recombinant fusion proteins. It is surprising that the simple Tsf fusion avoided inclusion body formation by nonspecific interactions between partially folded intermediates of unfolded proteins when overexpressed in the highly congested environment of the bacterial cytoplasm. The effectiveness of Tsf in shielding the interactive surfaces of heterologous proteins associated with nonspecific protein–protein interactions likely prevented the formation of inclusion bodies and significantly increased the solubility of aggregation-prone target proteins. This folding-enhancer property of Tsf may be interpreted in three ways. Firstly, Tsf by itself may have an inherent cis-acting folding helper activity, (i.e. intermolecular chaperones) like the maltose-binding protein (MBP), thioredoxin, and glutathione S-transferase (GST), which successfully overexpress many heterologous proteins in *E. coli* (Smith and Johnson, 1998; LaVallie *et al.*, 1993; Davis *et al.*, 1999; Nallamsetty & Waugh, 2006). Secondly, Tsf may be a factor recruiting molecular chaperones or chaperone-like protein factors. Helgstrand *et al.* (2007) found that various translation factors, EF-Tu, EF-G, and IF2, can bind the C-terminus of RimL, which has a structural homology with Tsf. It is therefore presumed that EF-Tu, EF-G, or IF2, if they are recruited by Tsf, might circumvent the aggregation and misfolding of heterologous proteins. Thirdly, Tsf stabilizes EF-Tu, enabling the molecule to behave as a molecular chaperone. In a manner similar to DnaK, EF-Tu of *E. coli* could recognize hydrophobic patches of substrates and promote refolding of unfolded polypeptides (Malki *et al.*, 2002), and EF-Tu is poorly stabilized in the absence of Tsf (Krab *et al.*, 2001). Consequently, Tsf might stabilize EF-Tu, which in turn effectively behaves as a chaperone by interacting with aggregation-associated hydrophobic surfaces of heterologous polypeptides.

**Bioactivity of the recombinant fusion mutant of cutinase**

Cutinase, a hydrolytic enzyme that degrades cutin (polyester composed of hydroxyl and epoxy fatty acids), has largely been exploited for industrial applications (Kim *et al.*, 2007). Microbial cutinase has been used in oleochemical industries (Carvalho *et al.*, 1999) and as an ingredient for dishwashing and laundry detergents (Flipsen *et al.*, 1998). In addition, widespread studies on the discovery of biological tools for degrading environmental pollution caused by synthetic polymers have been carried out using cutinases (Shimao, 2001). This cutinase enzyme with potentially wide applications has an extremely low enzyme activity specifically on the *p*-nitrophenyl ester of long-chain fatty acids, such as...
PNP, while it has a high hydrolytic activity for a variety of esters, ranging from soluble p-nitrophenyl esters to insoluble long-chain triglycerides (Kim et al., 2003). Cloning was performed of the cutinase gene from P. putida, the gene in E. coli cytoplasm was expressed using Tsf as a fusion partner protein, and hydrolytic activities of the cutinase fusion mutant on PNB and PNP were assayed. (As shown in Fig. 3, the cytoplasmic solubility of nonTsf-fused cutinase was almost negligible, i.e. < 1%, and hence the analysis of enzyme activity of the non-Tsf-fused cutinase was not carried out in this study.) As shown in Fig. 4, the cutinase fusion mutant Tsf::CUT showed the same biological activity as native cutinase in that the fusion mutant only degraded PNB effectively, and not PNP. These results indicate that the recombinant cutinase with the fusion tag, E. coli Tsf, was not only correctly folded to a native structure but also retained biological activity. Therefore, the stress-responsive protein E. coli Tsf proved to be effective as a potent solubility enhancer for aggregation-prone cutinase. The use of E. coli Tsf as a potent cis-acting solubility enhancer in the industrial production of various aggregation-prone heterologous enzymes in their biologically active forms is a promising new development for recombinant protein expression.

Acknowledgements

The authors appreciate the support of Professor Hang Chul Shin at Soomnsil University, who kindly provided the gene clones of mp-INS and G-CSF. The authors also thank Professors Won Tae Lee and Hyun Soo Cho at Yonsei University for kindly donating the gene clones of ppGRN, AID, and NACHT, respectively. This study was supported by the Korea Health 21 R&D Project of the Ministry of Health & Welfare of the Republic of Korea (grant no. A050750). This work was also supported by the Second Brain Korea 21 Project. Further supports from the Korea Science and Engineering Foundation (grant no. R01-2005-000-10355-0), the Korea Research Foundation (grant no. KRF-2004-041-D00180), and the Microbial Genomics & Applications Center (Taejon, Republic of Korea) are also appreciated.

Authors contribution

K.-Y.H. and J.-A.S. contributed equally to this work.

References


