Overexpression in *Escherichia coli*, Purification and Characterization of *Thermoanaerobacter tengcongensis* Elongation Factor G with Multiple Rare Codons

Liqiang Zhang, Peng Xue and Hongjie Zhang*

Center of System Biology, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Beijing 100101, China

Abstract: The *fus*A gene encoding a thermophilic protein EF-G with multiple rare condons was cloned from *Thermoanaerobacter tengcongensis* (*Tte*EF-G) and overexpressed in *Escherichia coli* by cotransfering a RIG plasmid to overcome the potential codon-bias problem originated from Arg, Ile and Gly. The recombinant protein was identified by MALDI-TOF-MS for molecular mass with approximation of 76 kDa and by trypsin digestion coupled LC-MS/MS for peptide sequence coverage of 61.3%. The *in vivo* complementary assay indicates that *Tte*EF-G could significantly rescue the *E. coli* LJ14 (*frr*^ts) at the non-permission temperature of 42°C in the bi-transformant of *Tte*RRF and *Tte*EF-G. This study indicated that coexpression of rare codons' cognate tRNA is a useful method for protein overexpression in *E. coli*.

Keywords: Elongation factor G, RIG plasmid, gene expression, purification and characterization, *Thermoanaerobacter tengcongensis*.

INTRODUCTION

Elongation factor G (EF-G) is an essential factor for protein synthesis in bacteria and in eukaryotic organelles, promoting the GTP-dependent translocation of the nascent protein chain from the A-site to the P-site of the ribosome [1, 2]. EF-G from bacteria has been a drug target for treating bacterial infection for several decades [3, 4]. Beyond that it has been identified as an essential protein in concerted interaction with ribosome recycling factor (RRF) during disassembly of the posttermination complex, freeing the ribosome for the next round of translation [5-7]. The interaction between EF-G and RRF has showed strong species specificity. To date, all reported EF-Gs are highly homologous in amino acid sequence, and only *Thermus thermophilus* EF-G crystal structure has been resolved in different forms [8-10]. There are four domains in EF-G, each well separated, and domains III and IV have been identified to participate in the specific interaction between EF-G and RRF domain II [6].

Previous work from our and foreign laboratories has showed that thermophilic proteins in relation with protein biosynthesis are good candidates for elucidating the mechanism of peptide biosynthesis. One advantage is that they can tolerate extensive mutation or truncation to a certain degree while keeping the protein folded in the assay temperature range in a temperature sensitive strain *in vivo*. This property has endowed it a powerful tool for providing useful structural and functional insight on protein biosynthesis. For example the expression and function study of *Tte*RRF and two chimeras with *Eco*RRF in *E. coli* have been served for the successful identification of each domain's function in RRF [7]. Another example is that C-terminal deletion showed

enhanced complementary activity of *Tth*RRF *in vivo* [11, 12]. In order to elucidate the post-termination complex disassembly mechanism, both RRF and EF-G structural information is essential. As a first step in understanding the structural bases of the species-specific interaction between *Tte*EF-G and *Tte*RRF, it is necessary to obtain sufficient quantities of purified *Tte*EF-G.

There are two *fus*A genes encoding EF-G like proteins in *T. tengcongensis* genome [13]. Gene *fusA* (NCBI locus: NP_623834) encodes a 690 amino acid peptide chain, and Gene *fusA*2 (NCBI locus: NP_623870) encodes a 700 amino acid peptide chain. Homologous analysis indicates that *fusA* encoded protein has share 83.9% similarity with *Tth*EF-G, significantly higher than that of *fusA2* encoded protein, which shares 66% similarity with *Tth*EF-G. A significant characteristic is that the EF-G is encoded extensively by rare used codons in protein expression system of *Escherichia coli*. In this paper *Tte*EF-G was highly expressed in *E. coli*, and its biological function has been identified. The results indicate that *Tte*EF-G encoded by *fusA* gene has elongation factor G's property and could rescue the temperature sensitive *E. coli* strain at the non-permission temperature when *Tte*EF-G and *Tte*RRF were co-transformed into that cells.

MATERIALS AND METHODS

Bacterial Strains, Plasmids and Chemicals

Thermoanaerobacter tengcongensis MB4 is a rod-shaped, gram-negative bacterium that propagates optimally at 75 °C. It was isolated recently from a hot spring in *Tengcong*, China [13]. It was used to prepare the genomic DNA for amplifying the *T. tengcongensis* EF-G gene in this study. *E. coli* LJ14 is a MC1061 strain (*frr*^ts), of which the chromosomal wild-type *frr* allele is replaced with a mutant allele (V117D) [14], allowing *E. coli* LJ14 to grow at 30°C but not at 42°C. The strain was used for the *in vivo* activity assays of the concerted interaction of *Tte*EF-G with *Tte*RRF by com-

*Address correspondence to this author at the Center of System Biology, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Beijing 100101, China; Tel: +86-10-64837257;
E-mail: hjzhang@sun5.ibp.ac.cn

plementation analysis. pET-28a is a T7 promoter initiated high expression vector for cloned gene. It was used for *Tte*EF-G expression. pQE-60 is a vector with a strong T_5 promoter purchased from QIAGEN. pSTV-28 was purchased from Takara, Japan, which encodes Cm^r, and was used to cotransform *Tte*EF-G gene into *E. coli* LJ14 strain for *in vivo* complementary assay. RIG plamid was a kind gift from W.G.J. Hol [15]. Restriction endonucleases and DNA ligation kit were from NEB and Takara, respectively. All other reagents were of analytic grade.

Cloning of the EF-G Gene from *T. tengcongensis*

Based on the putative EF-G gene sequence of *T. tengcongensis* [13] identified by bioinformatics, a forward, 5´-AGACAGC**CATATG**CCAAGGGATTTCAGCTTAGA TA AAGTTAGG-3´ and a reverse, 5´-GCAA**GGATCC**TTATT TTTTTGCTGATAAT ATTTGCTCAGC-3´ primer containing *Nde*I and *Bam*HI sites (bold) respectively, were designed to amplify the *T. tengcongensis* EF-G gene by PCR using *T. tengcongensis* genomic DNA as a template. The PCR product was inserted into the *Nde*I and *Bam*HI sites of pET-28a after digestion with the same enzymes, resulting in pET28-*Tte*EF-G for expression of *Tte*EF-G. The EF-G gene was also cloned into the pSTV-28 vector between the *Kpn*I and *Bam*HI sites, resulting in pSTV-*Tte*EF-G for the complementation assay.

Expression and Purification of *Tte*EF-G

pET-28-*Tte*EF-G plasmid and RIG plamid were cotransformed into *E. coli* BL21(DE3)pLysS cells. Transformants were grown in LB medium supplemented with 50 μg/ml kanamycin and 25 μg/ml chloramphenicol overnight at 37°C for small scale culture. The overnight culture (20 ml) was then diluted 50-fold with the same medium, and incubated at 37°C until a turbidity of 0.8 at 600 nm was obtained prior to induction by 0.5 mM IPTG for 4 h. The cell pellet was collected by centrifugation at $6,000 \times g$ for 20 min at 4°C, resuspended in 45 ml of chelating buffer (25 mM Tris-HCl, pH 8.0, 500 mM NaCl, and 10 mM imidazole), and frozen at −20°C until later use. Because twenty amino acid residues including a His_6-tag and a thrombin cleavage site were fused in frame to the N-terminus when the *Tte*EF-G gene was cloned into pET28a vector, the *Tte*EF-G protein was first expressed as a His_6-tagged protein. To purify the His_6-tagged *Tte*RRF protein, the above bacterial cells were thawed at 37°C, ultrasonicated on ice for 1 min × 5 to break the cells, and then the supernatant (40 ml) was recovered after the lysed bacteria were centrifuged at $37,000 \times g$ at 4°C for 30 min. The supernatant was applied directly onto a 5×1 cm² of chelating Sepharose Fast Flow column at a flow rate of 1.0 ml/min, which was charged with Ni^{2+} prepared according to the producer's protocol (Amershan Pharmacia Biotech) and equilibrated with the chelating buffer. The column was washed first with 50 ml of chelating buffer followed by 50 ml of washing buffer (25 mM Tris-HCl, pH 8.0, 500 mM NaCl, 80 mM imidazole). Finally, the column was eluted with elution buffer (25 mM Tris-HCl, pH 8.0, 500 mM NaCl, 200 mM imidazole), and monitored by absorbance at 280 nm. The major peak of protein was collected and then dialyzed against distilled water at 4°C. The purified fusion protein was then incubated with thrombin (Sigma T4648) in

the ratio of 10 units of thrombin mg^{-1} protein in cleavage buffer (20 mM Tris-HCl, pH 8.0, 2.5 mM $CaCl_2$, 10 mM $MgCl_2$) at 4°C for 8 h. After digestion, the cleavage mixture was applied onto the second metal chelating chromatography column to remove the fusion partner. Further purification was performed using Sephacry S100 column after the N-terminal peptide was removed by thrombin cleavage. Protein purity was checked by SDS-Page.

Mass Spectrographic Analysis

For MALDI-TOF-MS analysis, a protein sample (10 pmol) was dissolved in 40 μl of a solution containing 50% MeCN and 0.1% trifluoroacetic acid for MALDI-TOF-MS analysis. A freshly prepared solution of α-cyano-4-hydroxycinnamic acid (10 mg/ml) in 70% MeCN and 0.1% trifluoroacetic acid was mixed with the same volume of the above sample solution. A total of 1.0 μl of sample mixture was applied to the MALDI target and allowed to dry. A MALDI-TOF mass spectrum was acquired on an AXIMA-CFP plus mass spectrometer equipped with a 337.1-nm nitrogen laser. Data from 100 laser shots were averaged for each spectrum.

LC-MS/MS analysis sample was first digested in SDS gel using reported conditions. In brief, protein band was soaked in wash condition (50% MeCN, 2.5 mM NH_4HCO_3) for 24 h at 4°C and then washed and dried. Freshly prepared trypsin solution (10 μl; 0.02 μg/μl trypsin in 50 mM NH_4HCO_3) was added to plug for in-gel protein digestion at 37°C for 16 h. Peptides were extracted by sonication in 10 μl of 50% MeCN, 0.1% Formaric acid for 20 min and concentrated to dryness before use [16].

10 μl of the trypsin digested sample was loaded to a C_{18}-coated capillary column with a free tip after dissolving in 0.1% formaric acid buffer. Then a linear gradient elution process was used up to 60% MeCN. Mass spectra were captured using a LTQ mass spectrometry produced by Thermo Group Company. The MS/MS data were analyzed using software Bioworks 3.2. The criteria of $X_{corr}>1.8$ for single charged, $X_{corr}>2.5$ for two charged and $X_{corr}>3.75$ for three charged peptides with DeltaCN≥0.1000, RSP≤5 and SP≥500 was applied. NCBI protein database was used.

In Vivo Complementary Assay of the in Concerted Activity of *Tte*EF-G

The *in vivo* assay was performed by monitoring complementation of the *E. coli* LJ14 (frr^{ts}) with *Tte*RRF and *Tte*EF-G or with only *Tte*RRF. Plasmid pQE-*Tte*RRF and pQE-*Tte*EF-G, which expresses *Tte*RRF and *Tte*EF-G, respectively, were cotransformed into *E. coli* LJ14 (frr^{ts}), or as a contrast only pQE-*Tte*RRF was transformed. Cultures were started with 0.06% (v/v) inoculums from freshly growth (at 30°C) overnight cultures, and growth at both permissive (30°C) and non-permissive (42°C) temperatures was monitored by recording culture turbidities at 600 nm at regular intervals [17].

RESULTS AND DISCUSSION

Construction of *Tte*EF-G Expression Plasmid

As described in *Materials and methods*, the *Tte*EF-G coding sequence was synthesized by PCR and cloned into

vector pET28, resulting in pET28-*Tte*EF-G for the expression of the *Tte*EF-G gene. The constructed protein contains a His$_6$-tag and thrombin cleavage site coding sequences in the expression vector, which simplify the purification of the *Tte*EF-G protein. On this case three amino acid residues (GSH) was extended at EF-G's N terminus.

High Soluble Expression and Purification of the His$_6$-tagged *Tte*EF-G Protein

Fig. **1** lane A and lane B showed the expression of cloned *Tte*EF-G gene in *E. coli* without or with IPTG induction, respectively. In both cases, there were no significant expressions of the cloned *Tte*EF-G gene in *E. coli*. Checking the nucleotide sequence of *Tte*EF-G, it could be found that certain amino acid are extensively encoded by rarely used codons by *E. coli* in highly expressed genes [15]. Fig. **2** and Table **1** indicate the rare codons in *fusA* from *T. tengcongensis*. Arg has been extensively encode by rare codon AGA/AGG with the appearance of 91.4%; Ile by rare codon AUA (ATA in DNA code) with the appearance of 52.6%, and Gly by GGA with the appearance of 42.4%. Moreover these rarely used codons by *E. coli* are appeared highly at the translation starting terminus and also appeared in doublets, and even tetralets (Table **1**). Rare codons have been shown to greatly diminish expression of recombinant protein in *E. coli* because of translational stalling [18, 19]. One method of overcoming this codon bias is to re-engineer the *Tte*EF-G gene to be overexpressed so that it uses the preferred codons of *E. coli*. Obviously this method is both costly and time consuming [20, 21]. Recently Baca and Hol tried to cotransform *E. coli* with a plasmid (RIG) carrying the rare codon's cognate tRNA and could successfully express several parasite genes in *E. coli* highly [15]. In order to know whether this strategy could solve our problem, the *Tte*EF-G gene was cotransformed with RIG-plasmid in *E. coli* BL21(DE). Expression of the His$_6$-tagged *Tte*EF-G gene in *E. coli* was analyzed by SDS-PAGE. Fig. **1** lane C and lane D show the 10% SDS-PAGE pattern of the cell lysate from *E. coli* harboring pET28-*Tte*EF-G before and after induction with IPTG for 4 h. A very intense staining protein band can be observed in the induced total cell lysate (Fig. **1**, lane D), but not in the un-induced cell lysate (Fig. **1**, lane C). Further examination indicates that the expressed *Tte*EF-G is mainly in soluble form (data not shown). His$_6$-tagged *Tte*EF-G was first purified to apparent homogeneity in one step by Ni^{2+} chelating affinity chromatography (Fig. **1**, lane E) except several weak bands with lower molecular weight. This His$_6$-tag was further cleaved by thrombin and then re-eluted by Ni^{2+} chelating affinity column, and the unbound fraction was collected for further assay use.

Identification of the *Tte*EF-G Protein

MALDI-TOF-MS analysis shows two major peaks at *m/z* 76, 376.2 and 38, 205.4, corresponding to single and double protonated *Tte*EF-G, respectively (Fig. **3**). The molecular mass of *Tte*EF-G obtained by mass spectrometry is consistent with the theoretical value calculated from amino acid composition of the protein (76, 955.95) using the software of Paws. The relative deviation of the measured value versus its theoretical value is less than 0.8%. This relative large deviation may partly be due to the measurement having surpassed

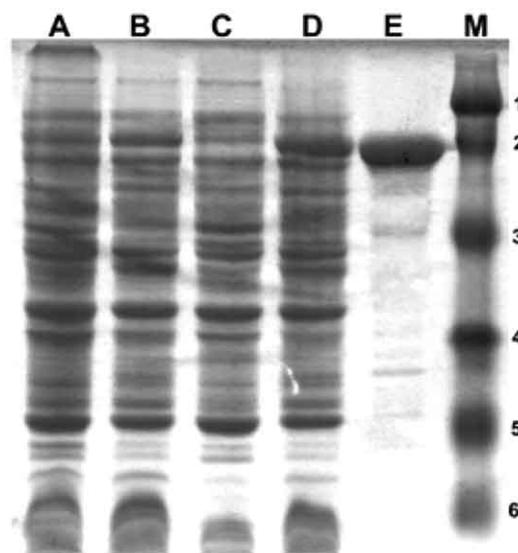the standard marker range of BSA, of which the average molecular mass is 66, 430.09 Da.

Figure 1. 10% SDS-Page analysis of *Tte*EF-G at different condition. Lane A. total protein extract from BL21 (DE3) harboring pET28-*Tte*EF-G alone before IPTG induction. Lane B. total protein extract from BL21 (DE3) harboring pET28-*Tte*EF-G alone after IPTG induction for 4.0 h. Lane C. total protein extract from BL21 (DE3) harboring pET28-*Tte*EF-G and RIG plasmid before IPTG induction. Lane D. total protein extract from BL21 (DE3) harboring pET28-*Tte*EF-G and RIG plasmid after IPTG induction for 4.0 h. Lane E. purified *Tte*EF-G protein after chelating affinity chromatography. Lane M. protein molecular weight marker (Fermentas) with different molecular weight, band 1, 118 kDa; band 2, 86 kDa; band 3, 47 kDa; band 4, 34 kDa; band 5, 26 kDa and band 6, 19 kDa, respectively.

In order to identify the purified protein more confidently, the *Tte*EF-G sample in SDS page was digested with trypsin thoroughly. The peptide mixture was loaded on a C$_{18}$-coated silica column, and then eluted with a linear gradient solution with increasing concentration of MeCN. The eluted solution was electro-sprayed and then was transformed into an Ion-trap mass spectrometry. The ionized peptide was captured and its *m/z* value was measured. Then this captured peptide was fragmented again within ion-trap, and the produced fragments were measured again. This mode of mass spectrum is called as MS/MS or MS^2. Searching the collected data in NCBI protein sequence database by software of Bioworks, it could be found that there were 57 peptides could be matched by MS/MS spectra with the criteria set. Fig. **4** showed the MS/MS spectrum of the identified peptide with the parent ion *m/z*=1191.712 (sequence is ILDGAVAVFSAK), and the corresponding *y*-ions and *b*-ions were analyzed and labeled out (see Table **2** also). The other 56 peptides have been identified also using the same strategy. The total sequence coverage reaches to 61.3% (data not shown). On the contrary, none of the identified peptides occurred in *fus*A2 product (NP_623870). The high sequence coverage indicates that the expressed protein is highly possible the target *Tte*EF-G (*fus*A gene product, NP_623834) protein.

```
       M   P   R   D   F   S   L   D   K   V   R   N   I   G   I   M   A   H   I   D   A   G   K   T   T   T
  1  ATGCCAAGG GATTTCAGC TTAGATAAA GTTAGGAAC ATTGGTATA ATGGCGCAT ATAGATGCA GGAAAAACC ACTACTACC
       E   R   I   L   F   Y   T   G   K   T   H   K   L   G   E   V   H   E   G   T   A   T   M   D   W   M   V
 82  GAGAGAATC CTTTTTTAT ACTGGTAAA ACCCATAAG TTAGGTGAA GTCCACGAA GGAACTGCT ACGATGGAC TGGATGGTA
       Q   E   Q   E   R   G   I   T   I   T   S   A   A   T   T   C   Y   W   K   G   H   K   I   N   I   I   D
163  CAGGAGCAG GAAAGGGGT ATCACGATA ACCTCTGCG GCGACCACT TGTTACTGG AAAGGTCAT AAAATTAAC ATAATAGAT
       T   P   G   H   V   D   F   T   V   E   V   R   S   L   R   I   L   D   G   A   V   A   V   F   S   A
244  ACGCCGGGA CACGTGGAC TTTACTGTG GAAGTAGAG AGGTCCTTA AGGATATTA GATGGGCA GTTGCGGTT TTTTCTGCA
       K   E   G   V   E   P   Q   S   E   T   V   W   R   Q   A   D   K   Y   H   V   P   R   I   A   Y   V   N
325  AAAGAGGGA GTAGAGCCT CAGTCTGAA ACTGTGTGG AGACAGGCG GATAAATAC CATGTGCCT AGGATTGCC TATGTAAAC
       K   M   D   V   I   G   A   D   F   N   V   I   D   M   I   R   E   R   L   G   A   N   P   V   A   I
406  AAGATGGAT GTTATAGGT GCTGACTTT TTTAACGTA ATAGACATG ATAAGAGAA AGGCTTGGT GCAAATCCT GTTGCTATT
       Q   I   P   I   G   K   E   D   T   F   R   G   I   V   D   L   I   K   M   E   A   I   I   Y   K   D   D
487  CAAATTCCT ATAGGAAAA GAGGATACT TTCCGGGGA ATCGTAGAC CTCATAAAA ATGGAAGCA ATAATTTAT AAAGATGAC
       L   G   T   V   M   D   E   T   E   I   P   E   D   L   K   P   L   A   E   Y   R   E   K   L   L   E
568  TTGGGGACT GTTATGGAT GAGACAGAA ATACCAGAA GATTTAAAA CCTCTTGCA GAAGAGTAT AGAGAAAAA TTGCTGGAA
       A   I   A   E   V   D   E   T   I   M   E   K   Y   L   E   G   E   E   I   T   E   E   I   H   A   A
649  GCAATTGCT GAAGTAGAT GAGACTATC ATGGAAAAG TATTTGGAA GGGGAAGAG ATAACAGAA GAAGAAATT CACGCGGCT
       L   R   K   G   T   I   N   G   E   L   V   P   V   V   C   G   S   S   Y   K   N   K   G   V   Q   P   L
730  TTGAGGAAG GGGACAATA AATGGAGAA CTTGTACCT GTAGTTTGT GGGTCTTCC TACAAGAAC AAGGGAGTA CAGCCTCTT
       L   D   A   I   V   N   Y   L   P   S   P   L   D   L   P   P   V   K   G   M   S   I   D   G   E   E
811  TTGGATGCT ATTGTGAAT TATTTGCCG TCTCCTTTA GATTTGCCA CCTGTTAAA GGTATGTCT ATAGATGGT GGCGAGGAA
       L   E   R   K   P   D   N   E   P   F   S   A   L   F   K   I   M   A   D   P   Y   V   G   K   L
892  CTTGAGAGA AAGCCTGAT GACAATGAG CCTTTCTCA GCTCTTGCC TTCAAAATC ATGGCAGAC CCTTATGTA GGTAAACTG
       A   F   F   R   V   Y   S   G   T   L   K   A   G   S   Y   V   L   N   S   T   K   G   K   K   E   R   I
973  GCTTTTTTT AGAGTATAT TCAGGACA TTAAAGGCA GGCTCTTAT GTTTTAAAT TCTACAAAG GGAAAGAAA GAAAGAATA
       G   R   I   L   R   M   H   A   N   H   R   E   E   I   D   A   V   Y   T   G   D   I   A   A   A   V   G
1054 GGAAGAAATC TTGAGAATG CACGCAAAT CACAGAGAA GAGATTGAT GCCGTATAT ACTGGAGAC ATTGCTGCA GCAGTTGGA
       L   K   D   T   T   T   G   D   T   L   C   D   E   N   H   P   I   L   E   S   M   D   F   P   E   P
1135 CTTAAAGAT ACGACTACA GGAGACACT CTTTGCGAT GAAAATCAT CCAATACTA TTAGAGTCA ATGGACTTT CCTGAGCCG
       V   I   S   V   A   I   E   P   K   T   K   A   A   Q   E   K   M   S   I   A   L   S   K   L   A   E   E
1216 GTTATTTCT GTAGCAATT GAGCCAAAG ACAAAGGCA GCTCAGGAA AAGATGAGC ATAGCTTTG TCAAAACTT GCTGAAGAG
       D   P   T   F   K   T   Y   T   D   Q   E   T   G   Q   T   I   I   A   G   M   G   E   L   H   L   E   I
1297 GACCCAACT TTCAAGACG TACACTGAC CAGGAGACA GGGCAGACT ATAATAGCT GGAATGGGA GAATTGCAC CTTGAGATA
       I   V   D   R   L   R   R   E   F   N   V   E   C   N   V   G   K   P   Q   V   A   Y   K   E   T   I   T
1378 ATCGTGGAC AGGTTGAGA AGGGAATTT AACGTAGAG TGCAATGTG GGTAAGCCT CAGGTTGCT TACAAAGAA ACTATAACA
       K   P   V   R   V   E   G   K   F   I   R   Q   S   G   G   R   G   Q   Y   G   H   V   W   L   E   M   E
1459 AAACCTGTA AGAGTAGAA GGAAAGTTC ATAAGGCAG TCAGGCGGT CGCGGCCAG TATGGACAT GTATGGCTT GAGATGGAA
       P   A   P   R   G   E   G   Y   I   F   E   N   R   I   V   G   G   V   I   P   K   E   F   I   P   A   V
1540 CCTGCTCCA AGAGGAGAG GGCTATATA TTTGAAAAT AGAATTGTT GGAGGTGTC ATTCCAAAA GAATTTATT CCAGCTGTA
       D   A   G   I   Q   E   A   M   Q   N   G   V   L   G   G   Y   P   V   I   D   V   K   V   A   L   V   D
1621 GATGCAGGT ATACAGGAA GCTATGCAA AACGGCGTG CTAGGGGGC TACCCTGTG ATTGACGTT AAAGTAGCA TTAGTCGAC
       G   S   Y   H   E   V   D   S   S   D   M   A   F   K   I   A   G   S   I   A   F   R   E   G   M   K   K
1702 GGTTCTTAC CATGAGGTT GACTCTTCT GATATGGCT TTCAAAATT GCAGGTTCT ATAGCCTTT AGAGAAGGT ATGAAGAAG
       A   D   P   V   L   L   E   P   I   M   K   V   E   V   V   V   P   E   E   Y   M   G   D   V   I   G   D
1783 GCGGACCCT GTGCTCTTA GAGCCTATC ATGAAAGTA GAAGTAGTA GTTCCAGAA GAATACATG GGAGATGTC ATAGGAGAT
       L   N   A   R   R   G   K   V   E   G   M   E   T   R   S   G   A   R   V   I   R   A   F   V   P   L   A
1864 TTGAACGCA AGGCGTGGA AAAGTAGAA GGTATGGAG ACAAGGTCT GGAGCTAGG GTCATAAGA GCTTTTGTG CCTTTGGCT
       E   M   F   G   Y   A   T   D   L   R   S   K   T   Q   G   R   G   T   Y   T   M   Q   F   H   Y   E
1945 GAAATGTTT GGATACGCT ACTGATTTA AGGTCAAAG ACACAGGGT AGAGGAACT TACACGATG CAGTTCCAT CACTATGAA
       E   V   P   K   N   I   A   E   Q   I   L   S   A   K   K   *
2026 GAAGTGCCT AAAAATATT GCTGAGCAA ATATTATCA GCAAAAAAA TAA
```

Figure 2. Sequence analysis of *fusA* gene from *T. tengcongensis*. The rare codons for *E. coli* (AGA/AGG for Arg, ATA for Ile and GGA for Gly) were impressed by boxes.

Table 1.   Usage of Rare Codon for Arg, Ile and Gly in *fus*A from *T. tengcongensis*

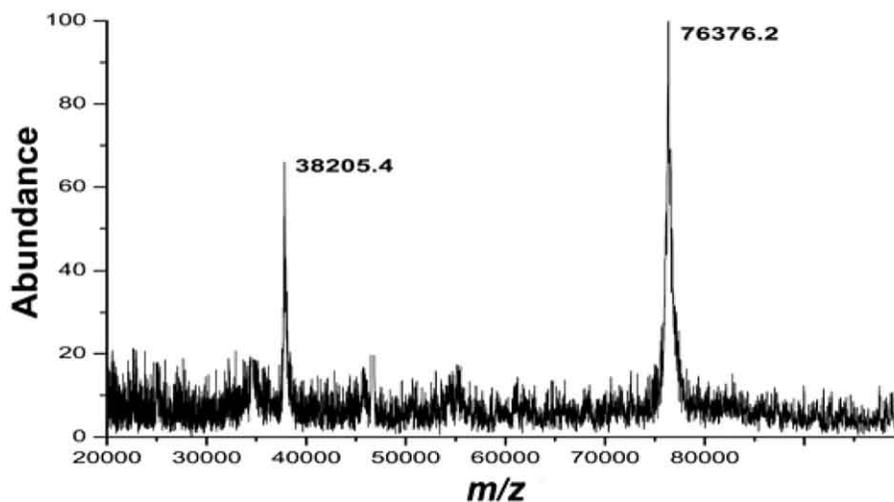|	Arg (AGA/AGG)	Ile (AUA)	Gly (GGA)	Number of Multiple rare codon		
				doublets	Triplets	Tetralets
No. of total codons	35	57	59			
No. of rare codons	32	30	25	11	0	1
Percentage of rare codons	91.4	52.6	42.4			

Figure 3. MALDI-TOF-MS of *Tte*EF-G. The sample (10 pmol) was dissolved in 50% acetonitril and 0.1% (v/v) trifluoroacetic acid and applied onto the target by mixing it with α-cyano-4-hydroxycinnamic acid as the matrix. The single and double charged molecular peaks were labeled out.
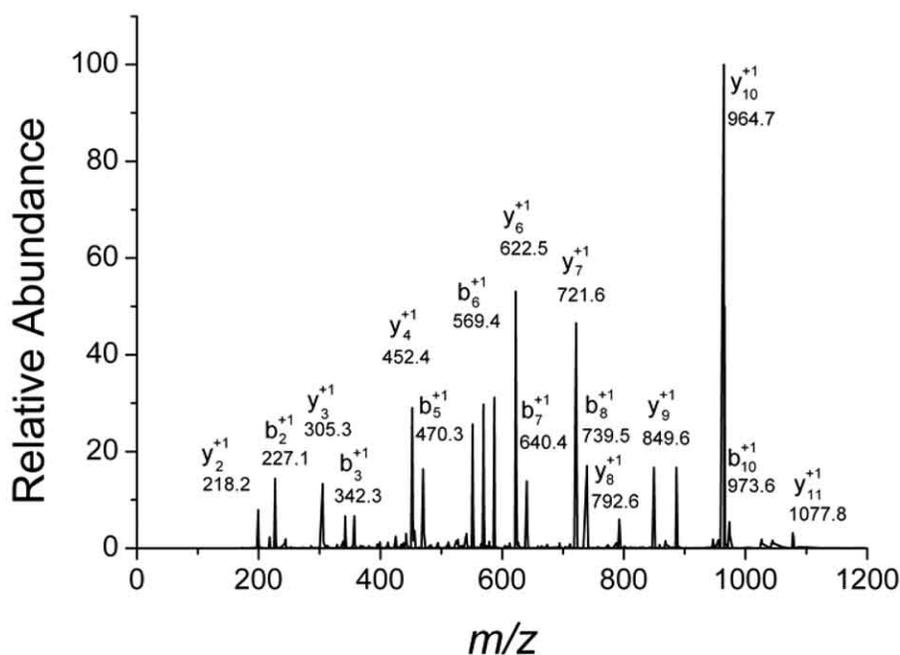

Figure 4. MS/MS spectrum acquired for the peptide. The parent ion selected is m/z=1191.712. The *b*-ions and *y*-ions were labeled out.

Table 2.   Calculated Masses for b and y Ions from MS/MS Measurements of the 1191.712 Da Tryptic Fragment

b-ion	MH^+ calc.	Residue	MH^+ calc.	y-ion
1	114.09	I	-	12
2	227.18	L	1077.59	11
3	342.10	D	964.51	10
4	399.22	G	849.48	9
5	470.26	A	792.46	8
6	569.33	V	721.42	7

(Table 2) contd….

b-ion	*MH*+ calc.	Residue	*MH*+ calc.	y-ion
7	640.37	A	622.36	6
8	739.43	V	551.32	5
9	886.50	F	452.25	4
10	973.54	S	305.18	3
11	1044.57	A	218.15	2
12	-	K	147.11	1

Functional Characterization of *Tte*EF-G

As reported recently, the concerted interaction between RRF and EF-G during disassembly of the translation post-termination complex showed strong species specificity. This provides a possible way to measure the biological activity of the expressed protein in semi-quantitative. In a previous publication from our group, we have reported that *Tte*RRF showed very low complementary activity in a temperature sensitive *E. coli* LJ14 (*frr*ts) transformant at 42°C. Only after prolonged incubation the *E. coli* strain growth could be observed. In contrast, the strain with *Eco*RRF transformant could proliferate very well at the same condition. The low *in vivo* activity of *Tte*RRF in *E. coli* is possibly due to the low efficient recognition between *Tte*RRF and *Eco*EF-G. To investigate whether *Tte*EF-G is able to complement *E. coli* LJ14 harboring *Tte*RRF gene, we monitored the growth of transforments harboring vector alone (pQE-60) or the re-combinants pQE-*Tte*EF-G and pQE-*Tte*RRF, which express *Tte*EF-G and *Tte*RRF, respectively, or the recombinant pQE-*Tte*EF-G and pQE-*Tte*RRF together. Although all the trans-forments grown in liquid cultures well at the permissive temperature (30°C) (data not shown), only the transformant harboring pQE-*Tte*EF-G plus pQE-*Tte*RRF could grow significantly after 12 h incubation at the non-permission tem-perature (42°C), while at that time the transformant with only *Eco*EF-G grows less, indicating that the expressed pro-tein *Tte*EF-G could recognition *Tte*RRF efficiently than *Eco*EF-G (Fig. **5**). The in concerted interaction of *Tte*EF-G and *Tte*RRF disassemble the translational posttermination complex, showing one of *Tte*EF-G's biological functions.

It should be pointed out that in the *in vivo* activity assay system, the *Tte*EF-G gene could also be expressed without the help of RIG-plasmid. Although the amount of *Tte*EF-G is very low, it is enough to rescue the *E. coli* LJ14 strain from heat shot at 42°C. This result indicates that *E. coli* itself has the ability of synthesis the rare codon's cognate tRNA, but its low amount limits the synthesis rate of the extrinsic gene and/or made it difficult to accumulate in cell due to degrada-tion. The existence of RIG-plasmid is mainly to increase the rate of peptide synthesis, not to alter the peptide synthesis mechanism. This study also indicated that coexpression of rare codons' cognate tRNA may be a widely applicable method for protein expression in *E. coli*.
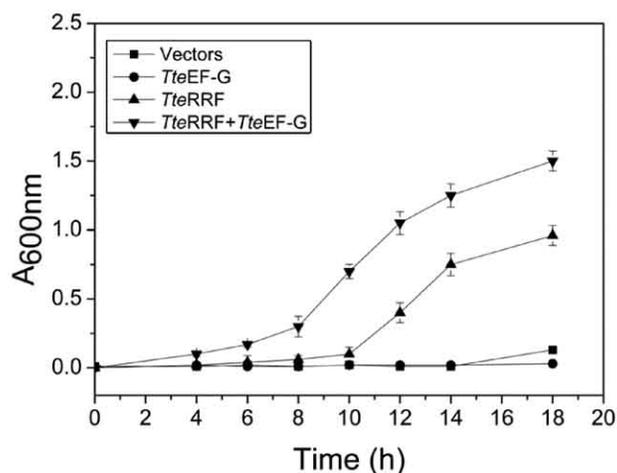

Figure 5. Growth curves of various transformants of *E. coli* LJ14 (*frr*ts) at the non-permissive temperature (42 °C). Vector indicates the transformants with pQE-60 and pSTV-28; *Tte*EF-G, the trans-formants with pQE-60 and pSTV-*Tte*EF-G; *Tte*RRF, the transfor-mants with pQE-*Tte*RRF and pSTV-28; *Tte*RRF and *Tte*-EF-G, the transformants with pQE-*Tte*RRF and pSTV-*Tte*EF-G, respectively. All the transformants were cultured at 42 °C with the same start value of A$_{600nm}$ = 0.04.

REFERENCES

[1]　Wintermyer, W. and Rodnina, M.V. (**2000**) *Essays. Biochem., 35*, 117-129.
[2]　Valle, M., Zavialov, A., Sengupta, J., Rawat, U., Ehrenberg, M. and Frank, J. (**2003**) *Cell, 114*, 123-134.
[3]　Seo, H.S., Abedin, S., Kamp, D., Wilson, D.N., Nierhaus, K.H. and Cooperman, B.S. (**2006**) *Biochemistry 45*, 2504-14.
[4]　Besier, S., Ludwig, A., Brade, V. and Wichelhaus, T.A. (**2005**) *Antimicrob. Agents Chemother., 49*, 1426-31.
[5]　Rao, A.R. and Varshney, U. (**2001**) *EMBO J., 20*, 2977-2986.
[6]　Ito, K., Fujiwara, T., Toyoda, T. and Nakamura, Y. (**2002**) *Mol. Cell, 9*, 1263-1272.
[7]　Guo, P., Zhang, L.Q., Zhang, H.J., Feng, Y. and Jing, G. Z. (**2006**) *Biochem. J., 393*, 767-777.

[8] Czworkowski, J., Wang, J., Steitz, T. A. and Moore, P. B. (**1994**) *EMBO J., 13*, 3661-3668.

[9] Aevarsson, A., Brazhnikov, E., Garber, M., Zheltonosova, J., Chir-gadze, Y., al-Karadaghi, S., Svensson, L. A. and Liljas, A. (**1994**) *EMBO J., 13*, 3669-77.

[10] Hansson, S., Singh, R., Gudkov, A.T., Liljas, A. and Logan, D.T. (**2005**) *FEBS Lett., 579*, 4492-7.

[11] Fujiwara, T., Ito, K., Nakayashiki, T. and Nakamura, Y. (**1999**) *FEBS Lett., 447*, 297-302.

[12] Toyoda, T., Tin, O.F., Ito, K., Fujiwara, T., Kumasaka, T., Yama-moto, M., Garber, M. B. and Nakamura, Y. (**2000**) *RNA, 6***,** 1432-1444.

[13] Bao, Q., Tian, Y., Li, W., Xu, Z., Xuan, Z., Hu, S., Dong, W., Yang, J., Chen, Y., Xue ,Y., Xu, Y., Lai, X., Huang, L., Dong, X., Ma, Y., Ling, L., Tan, H., Chen, R., Wang, J., Yu, J. and Yang, H. (**2002**) *Genome Res., 12*, 689-700.

[14] Janosi, L., Mottagui-Tabar, S., Isaksson, L. A., Sekine, Y., Ohtsubo, E., Zhang, S., Goon, S., Nelken, S., Shuda, M. and Kaji, A. (**1998**) *EMBO J., 17*, 1141-1151.

[15] Baca, A.M. and Hol, W.G.J. (**2000**) *Int. J. Parasitol., 30*, 113-118.

[16] Tay, T.L., Lin, Q., Seow, T.K., Tan, K.H., Hew, C.L. and Gong, Z. (**2006**) *Proteomics, 6*, 3176-3188.

[17] Zhang, LQ., Guo, P., Zhang, H.J. and Jing, G.Z. (**2006**) *Arch. Bio-chem. Biophys., 450*, 191-202.

[18] Kane, J.F. (**1995**) *Curr. Opin. Biotechnol., 6*, 494-500.

[19] Kurland, C. and Gallant, J. (**1996**) *Curr. Opin. Biotechnol., 7*, 489-493.

[20] Prapunwattana, P., Sirawaraporn, W., Yuthavong, Y. and Santi, D.V. (**1996**) *Mol. Biochem. Parasitol., 83*, 93-106.

[21] Pan, W., Ravot, E., Tolle, R., Frank, R., Mosbach, R., Turbachova, I. and Bujard H. (**1999**) *Nucleic Acid Res., 27*, 1094-1103.