Manipulation of the *tuf* gene provides clues to the localization of sequence element(s) involved in the thermal stability of *Thermotoga maritima* elongation factor Tu

Anna M. Sanangelantoni,¹ Piero Cammarano² and Orsola Tiboni†

Author for correspondence: Anna M. Sanangelantoni. Tel. +39 382 505578. Fax: +39 382 528496. e-mail: Sanan@ipvgen.unipv.it

Truncated versions of the *tuf* gene for elongation factor Tu (EF-Tu; 400 aa) from the hyperthermophilic bacterium *Thermotoga maritima* have been produced by progressive 3' → 5' trimming. The truncated genes have been expressed in *Escherichia coli* and the thermal stability of the gene products has been assayed by monitoring their GDP-binding capacity after preheating the cell-free extracts at various temperatures (65–95°C). One of the truncated proteins, corresponding to the nucleotide-binding domain (G domain; aa 1–200) appears to be only slightly less stable than the full-length EF-Tu. Replacement of the first 90 N-terminal residues of both the full-length *Thermotoga* EF-Tu and the isolated G domain with the corresponding sequence of the mesophilic bacterium *E. coli*, drastically destabilizes both the complete and the truncated protein, indicating that sequence element(s) that are crucial for the attainment of a thermally stable conformation of the *Thermotoga* EF-Tu lie well within the initial portion of the G domain between residues 1 and 90. The relevant residues defy identification, however, as no amino acid preferences, or exclusive sequence element(s), appear to distinguish the N-terminal region of the thermophilic proteins from those of mesophilic counterparts. It is suggested that the thermal stability of *Thermotoga* EF-Tu is critically dependent upon unique tertiary structural interactions involving certain N-terminal residues of the molecule.

**Keywords:** elongation factor Tu, extreme thermophiles, thermostability, protein engineering, genetic manipulation

INTRODUCTION

Proteins from hyperthermophilic micro-organisms maintain their heat resistance upon expression in mesophilic hosts (Love & Stciff, 1987; Tiboni *et al.*, 1989; Sanangelantoni *et al.*, 1992), indicating that heat stability is solely due to intrinsic features of the molecule rather than to extrinsic factors (Fabry *et al.*, 1989; Schultes *et al.*, 1990). Nonetheless, no unique amino acid substitution scheme, or exclusive sequence motif, distinguishes the thermophile proteins from their mesophile counterparts (Adams, 1993). This implies that thermal resistance may be brought about by subtle (and elusive) primary structural changes resulting in increased hydrophobicity, chain rigidity and overall tightness of intramolecular packing (Nosoh & Sekiguchi, 1990). However, evidence for this is scanty.

The elongation factors ubiquitously involved in aminoacyl-tRNA binding to ribosomes (EF-Tu for Bacteria; EF-1a for Archaea and Eucarya) offer a good case in protein stability studies. First, the available sequences (about 400 relatively conserved residues) span a host of evolutionarily diverse organisms including the hyperthermophilic bacterium *Thermotoga maritima* (Bachleitner *et al.*, 1989) and an as yet unpublished sequence of *Aquifex pyrophilus* (W. Ludwig, Technische Universität München, Germany, personal communication). Secondly, the stability difference between the *T. maritima* factor, which withstands temperatures up to 90–95°C, and that of the mesophile *E. coli*, which is inactivated at temperatures slightly above 50°C, is far greater than that seen for most other proteins. Thirdly, the *T. maritima* gene for EF-Tu...
(tuf) can be overexpressed in *E. coli* and the heterologous gene product can be easily purified via the selective denaturation of the (heat-labile) host-cell proteins (Tiboni et al., 1989). Lastly, the EF-Tu N-terminal domain involved in GTP binding and hydrolysis (about 200 residues) retains its GTP-binding capacity following removal of the middle and C-terminal domains involved in the binding of aminoacyl-tRNA, elongation factor Ts and kifamycin-like antibiotics (Weiiland et al., 1992). In principle this enables one to selectively modify tuf regions corresponding to individual EF-Tu domains and to assess the stability of the encoded protein by means of simple functional assays. This approach may thus allow the localization of sequence element(s) that are responsible for the increased overall stability of the thermophile protein.

In this report, truncated versions of the *T. maritima* EF-Tu, and chimeric (mesophile/thermophile) versions of both the full-length protein and the G domain, have been generated by manipulation of the tuf gene; the modified gene sequences have been expressed in *E. coli* and the thermal stability of the gene products has been monitored by means of GDP-binding assays. The results suggest that sequence element(s) located in the N-terminal portion of the G domain play a crucial role in maintaining the thermally stable conformation of the *T. maritima* protein, although no distinguishing primary structural features can be identified through comparison of thermophile and mesophile EF-Tu.

**METHODS**

**Bacterial strains and plasmids.** Plasmids pTM6 (Tiboni et al., 1989) and AB80 (Bernardi & Bernardi, 1979), containing the *tuf* genes of *T. maritima* and *E. coli*, respectively, were used to prepare the truncated and chimeric genes. Plasmids pDR720 (Russel & Benett, 1982), pBR322 (Bolivar et al., 1977) and pUC19 (Vieira & Messing, 1982) were used as the vectors and *E. coli* HB101 as the host. Plasmid-containing strains were grown either in LB medium or in M9 medium supplemented with 10 mM magnesium sulphate, 0.2% glucose and 0.3% Casamino acids; 3β-indole-3-acrylic acid was then added to a concentration of 50 μg mL⁻¹ and the incubation was continued for a further 2.5 h. Cells were collected by centrifugation at 5000 g for 10 min and washed with buffer A (50 mM Tris/Cl, pH 8.0, 10 mM MgCl₂, 50 mM KCl, 70 mM 2-mercaptoethanol) and stored at −20°C until use.

**Preparation of cell-free extracts and GDP-binding assays.** Cells were broken by grinding with twice their own mass of alumina powder. The paste was extracted with buffer A (30 mM per g wet cells). After centrifugation at 30000 g for 30 min to remove alumina and cell debris, the clear extracts were centrifuged at 100000 g for 2 h and 100 μl aliquots of the resultant ribosome-free supernatant (S100, containing 8–12 mg protein mL⁻¹, depending on the preparation), were incubated for 10 min at various temperatures between 65°C and 95°C. As noted elsewhere (Tiboni et al., 1989), a preliminary 10 min incubation of the S100 at 65°C is essential for the activation of the *T. maritima* EF-Tu expressed by transformed *E. coli* cells and for denaturation of most *E. coli* host-cell proteins.

After removal of precipitated proteins (10 min centrifugation at 100000 g) the GDP-binding activity was assayed by measuring the amount of [³H]GDP (Amersham; sp. act. 459 GBq mmol⁻¹) bound by 10 μL of the heat-treated extracts, during a 10 min incubation at 30 or 65°C (Tiboni et al., 1978).

**Separation of chimeric G domain.** All operations were carried out at 4°C. Cell-free extracts from 15 g of cells were prepared in buffer B (50 mM Tris/Cl, pH 7.4, 10 mM MgCl₂, 80 mM KCl, 70 mM 2-mercaptoethanol, 15%, v/v, glycerol) as described previously. The protein solution (35 mL) was applied to a 2.5×35 cm DEAE-cellulose DE52 column equilibrated with buffer B. After adsorption of the sample, the column was developed with 11 80–320 mM KCl gradient in buffer B at a constant flow rate of 50 mL h⁻¹ while collecting fractions of 12 mL. Fractions were assayed for GDP-binding activity at 30°C and the protein profile was determined by SDS-PAGE. Positive fractions were pooled, dialysed against buffer B without glycerol, and solid ammonium sulphate was added to 75% saturation. After centrifugation at 20000 g for 15 min the protein precipitate was dissolved in buffer B containing 10 mM GDP.

**PAGE.** Proteins were analysed by SDS-PAGE (Laemmli, 1970). The following markers were used: lysozyme (144 kDa), soybean trypsin inhibitor (215 kDa), carbonic anhydrase (31 kDa), ovalbumin (45 kDa), BSA (66.2 kDa) and phosphorylase b (92.5 kDa). Proteins were stained with Coomasie brilliant blue.

**Sequence analysis.** Analysis of the DNA sequence was performed using the DNA STRIDER 1.0 program (Marck, 1988).
RESULTS AND DISCUSSION

The *T. maritima* EF-Tu sequence was aligned with all of the bacterial and archaeal sequences presently available in the databases, and with an as yet unpublished EF-Tu sequence from the hyperthermophilic bacterium *Aquificps pyrophilus* (courtesy of W. Ludwig) probably representing the deepest known branching in the bacterial tree (Burggraf et al., 1992; Bocchetta et al., 1995). For concision of presentation only two mesophilic and two thermophilic sequences are shown in Fig. 1 (a global alignment of the *T. maritima* and *Thermus* EF-Tu with all mesophilic counterparts is available from the authors upon request). All of the bacterial sequences appear to be highly conserved (65-82% overall sequence identity), especially within the N-terminal region harbouring the three consensus elements (GXGGXGK, DXG and PTUQ) in all of the available archael sequences (from both mesophilic and thermophilic sources). This prevented site-directed mutagenesis experiments aimed at identifying single residues that could be responsible for the high stability of the thermophile protein. Consequently, the sequence element(s) involved in the thermal stability of *T. maritima* EF-Tu were investigated by constructing progressively deleted, or chimeric, tuf genes, and testing the heat resistance of the resultant gene products.

Inspection of the restriction map of the plasmid pTu1, containing the *Thermotoga* tuf gene, revealed a unique BglII site situated 263 bp upstream from the termination codon of *tuf*, which created ends compatible with those produced by the BamHI site of the pUC19 polylinker. Accordingly, a first deletion clone (termed pTu2) was produced by digestion of pTu1 with restriction enzymes BglII and BamHI followed by ligation (Fig. 1). Additional progressive deletions were produced by the 'Erase-a-Base' system starting from a BglII/KpnI-digested pTu1 plasmid. The resultant recombinant plasmids were analysed for insert length and only inserts having a suitable size were sequenced at the 3' end. Clones having a termination codon immediately downstream of the *T. maritima* tuf sequence, namely pTu3, pTu4, pTu5 (whose encoded polypeptide is indicated by an arrow on the last amino acid in the alignment shown in Fig. 1) and pTu6 (in which the entire tuf gene had been deleted), were selected for further analysis.

Expression of the truncated genes and thermal stability of the gene products

As Fig. 1 shows, all of the polypeptides encoded by the truncated tuf genes contain the three consensus elements required to form the nucleotide-binding pocket. Therefore, the presence of the expressed proteins could be monitored by measuring the GDP-binding capacity of the cell-free extracts at 65 °C, a temperature treatment that allows offsetting of host-cell EF-Tu activity and is also required to activate the entire (overexpressed) *T. maritima* EF-Tu (Tiboni et al., 1989). In essence, cell-free extracts were prepared from *E. coli* HB101 cells transformed with plasmids harbouring the entire *T. maritima* tuf gene (plasmid pTu1), or the progressively deleted versions thereof (plasmids pTu2 to pTu5), and comparable extracts
Expression in *E. coli* of progressively deleted *Thermotoga tuf* genes. (a) SDS-PAGE analysis on 10 μl aliquots of the extracts of *E. coli* transformed with the following plasmids: pT6 (control), pT1 (harbouring the gene encoding the *T. maritima* EF-Tu), pT2, pT3 and pT4 (encoding the isolated G domain), heated at the indicated temperatures (°C). NT, Not treated. The gel electrophoretic patterns are representative of all 5100 preparations obtained. Identical gel patterns and similar proportions of heterologous proteins were observed in different preparations irrespective of differences in the GDP-binding capacity (see b). Lane M contains molecular mass markers. GDP-binding activity of extracts treated and assayed as reported in Methods (7500 c.p.m. correspond to 1 pmol of bound GDP). Note that in the pT1 and pT4 assays the maximum preincubation temperature was 95 °C instead of 90 °C. The results of four to six separate assays are shown.

Fig. 2. Expression in *E. coli* of progressively deleted *Thermotoga tuf* genes. (a) SDS-PAGE analysis on 10 μl aliquots of the extracts of *E. coli* transformed with the following plasmids: pT6 (control), pT1 (harbouring the gene encoding the *T. maritima* EF-Tu), pT2, pT3 and pT4 (encoding the isolated G domain), heated at the indicated temperatures (°C). NT, Not treated. The gel electrophoretic patterns are representative of all 5100 preparations obtained. Identical gel patterns and similar proportions of heterologous proteins were observed in different preparations irrespective of differences in the GDP-binding capacity (see b). Lane M contains molecular mass markers. GDP-binding activity of extracts treated and assayed as reported in Methods (7500 c.p.m. correspond to 1 pmol of bound GDP). Note that in the pT1 and pT4 assays the maximum preincubation temperature was 95 °C instead of 90 °C. The results of four to six separate assays are shown.

were also prepared from cells containing the control plasmid pT6. In preliminary assays, all extracts, except those derived from cells containing plasmids pT5 and pT6, were found to exhibit an appreciable GDP-binding capacity following incubation at 65 °C (data not shown). As the pT5-encoded protein (Fig. 1) still contains the three consensus elements that are required to form the nucleotide-binding pocket, but lacks GDP-binding activity, the minimum EF-Tu sequence required for GTP/GDP binding is borne by plasmid pT4. By comparison with the *E. coli* sequence (An & Friesen, 1980; Yokota et al., 1980), the polypeptide encoded by the latter plasmid was found to correspond to the EF-Tu G domain (residues 1-200) (Parmeggiani et al., 1987). Importantly, extracts from *E. coli* cells harboring the pT6 plasmid were completely devoid of GDP-binding activity at 65 °C, as expected if *E. coli* EF-Tu is completely denatured at this temperature.

In order to investigate the heat stability of the truncated EF-Tu proteins, aliquots of the S100 fraction (see Methods) were incubated at various temperatures between 65 °C (the temperature required for activation of overexpressed *T. maritima* EF-Tu) and 95 °C; the GDP-binding capacity of the preheated samples was then assayed at a constant temperature of 65 °C.

The gel electrophoretic patterns in Fig. 2(a) show that the products of the four recombinant plasmids are expressed to different degrees. Namely, the gel patterns of heat-treated S100 from *E. coli* cells harbouring either the pT1 or the pT4 plasmids display a prominent extra band that resists heating at 95 °C (pT1) and 85 °C (pT4). Although no extra bands are discernible in extracts prepared from cells harbouring either pT2 or pT3, the production of heterologous (*T. maritima*), thermally stable, protein by the two latter clones can still be inferred by the GDP-binding capacity of the extracts (Fig. 2b).

The results of GDP-binding assays carried out using several independent S100 preparations of each of the four recombinant clones are shown in Fig. 2(b). Variations in the GDP-binding capacity of the different preparations obtained from the same recombinant clone are probably accounted for by different activation levels attained by the
recombinant proteins during preheating and, to a minor extent, by differences in protein content of the parent cell-free extracts (8–12 µg protein µl⁻¹) (see Methods) or by different expression levels of the heterologous protein. Despite differences in the activity levels, however, independent preparations obtained from the same parent clone exhibit similar thermal-inactivation profiles. Indeed, the pTu1-encoded protein (full-length EF-Tu) is still apparent in the gel electrophoretic patterns of extracts heated up to 95 °C (Fig. 2a) and its GDP-binding activity is not abolished following heating (Fig. 2b).

The plots in Fig. 2(b) also show that the plasmid pTu2-encoded protein (aa 1–316), which lacks the last 86 C-terminal residues of the EF-Tu appears to be less heat stable than the full-length protein: its activity is completely abolished by heating at 90 °C and greatly reduced by heating at 85 °C (as deduced by GDP-binding assays). No further reduction of the protein’s stability occurs upon removal of 40 additional residues (plasmid pTu3, aa 1–276). Also, no further decrease in thermal stability is observed by extending deletions up to clone pTu4 which encodes the T. maritima G domain. In fact, the isolated G domain (aa 1–206) seems to be more heat stable than the pTu2- and pTu3-encoded proteins, although somewhat less stable than the full-length EF-Tu. This latter point is inferred by the fact that the protein is still abundant following heating of the extracts at 85 °C, but totally absent following heating at 95 °C.

The different stabilities of the three truncated proteins can be interpreted on the basis of the known three-dimensional structure of E. coli EF-Tu (Kjeldgaard & Nyborg, 1992). This appears to comprise three domains [aa 1–200 (G domain), aa 209–299 (domain I) and aa 300–393 (domain III)] whose interactions are important for the global function of the protein, and its overall structure is basically conserved in the thermophilic bacterium Thermus thermophilus (Berchtold et al., 1993). Although the native EF-Tu is required for tRNA binding to ribosomes, the three isolated domains do possess some of the EF-Tu activities. Namely, the E. coli G domain prepared using a genetically engineered tfAl (Parmeggiani et al., 1987) has been shown to retain certain properties of the intact molecule such as GDP and GTP-binding capacity and to display an intrinsic GTPase activity. Domain II/III (middle and C-terminal domains), isolated as a single polypeptide by tryptic digestion of the Thermus thermophilus EF-Tu, has been shown to maintain the ability to bind the elongation factor Ts and to be essentially refractory to proteolytic attack (Peter et al., 1990). As the isolated domains, although stable, are more prone to thermal denaturation than the intact protein, it seems that the ability of the thermophile EF-Tu to withstand temperatures up to 95 °C rests not only on specific amino acid preferences, but also on stabilizing co-operative interactions between the three domains. As the truncated Thermotoga maritima EF-Tu encoded by plasmids pTu2 and pTu3 lack either a segment of domain III (pTu2), or the entire domain III and part of domain II (pTu3), they are probably unable to attain the correct tertiary folding and, for this reason, they are more susceptible to thermal denaturation. In addition, as all the tfAl genes possess an identical upstream region comprising promoter and ribosome-binding site, the low expression levels of the recombinant protein in cells harbouring plasmids pTu2 and pTu3 can be accounted for by degradation of the gene products by proteases of the mesophile host. If this is the case, the relative abundance of T. maritima G domain in E. coli cells harbouring the pTu4 plasmid could reflect a more compact tertiary packing, rendering the protein less susceptible to proteolytic attack.

As expected, no GDP-binding activity, and no extra bands, are discernible in S100 extracts prepared from E. coli cells containing the control plasmid pTu6 (Fig. 2a, b).
Construction of chimeric \textit{tuf} genes and analysis of the encoded products

Since a unique \textit{BclI} site lies at an identical position (nucleotide 273) in both the \textit{T. maritima} (Buchleitner et al., 1989) and the \textit{E. coli} (An & Friesen, 1980; Yokota et al., 1980) \textit{tuf} gene sequences, hybrid genes were constructed to assess the contribution of the first 90 amino acids to the heat stability of both the \textit{T. maritima} \textit{EF-Tu} and the isolated \textit{G} domain. To this end, the recombinant plasmids \textit{pTu4} and \textit{pTuB} (harbouring the \textit{E. coli} \textit{tufB} gene) were digested with \textit{XhoI} and \textit{BclI} and the resultant fragments were ligated to generate the chimeras shown in Fig. 3.

SDS-PAGE analysis of crude extracts obtained from \textit{E. coli} cells transformed with recombinant plasmids encoding a chimeric \textit{EF-Tu}, or a chimeric \textit{G} domain, displayed an extra band of molecular mass about 48 or 30 kDa, respectively (Fig. 4b, d). Incubation of the extracts at increasing temperatures (up to 85–90°C) showed that replacement of the N-terminal portion (90 aa) of the thermophile proteins with the corresponding portion of their mesophile (\textit{E. coli}) counterpart does affect the protein’s thermal stability. As Fig. 4(a, b) shows, while the protein band (arrow in Fig. 4) corresponding to the \textit{T. maritima} \textit{EF-Tu} tolerates heating at 85°C, the chimeric protein begins to disappear following incubation of the extracts at 70°C. Therefore, this region of the \textit{T. maritima} \textit{EF-Tu} harbours sequence element(s) which are essential for the maintenance of its thermally stable conformation. When the same analysis was extended to the \textit{G} domain (Fig. 4c, d), which lacks the co-operative interactions with domains II and III, the effect of temperature on protein stability was more dramatic, the chimeric protein being as heat labile as the mesophile \textit{EF-Tu}.

The authenticity of the chimeric \textit{EF-Tu} was confirmed by testing the ability of cell-free extracts to bind GDP at 65°C (data not shown). The same experiment was not feasible in the case of the chimeric \textit{G} domain, as the protein was completely inactivated at 65°C and a GDP-binding assay done below this temperature did not allow discrimination of the chimeric \textit{G} domain and the host \textit{EF-Tu} activities. Accordingly, separation of this chimeric protein from the \textit{E. coli} \textit{EF-Tu} was undertaken.

Extracts prepared from \textit{E. coli} cells harbouring the recombinant plasmid encoding the chimeric \textit{G} domain were chromatographed on a DEAE-cellulose column and the fractions were tested for GDP-binding activity at 30°C. As Fig. 5 shows, two activity peaks were detected and identified as the \textit{E. coli} \textit{EF-Tu} and the chimeric \textit{G} domain, respectively, on the basis of their SDS-PAGE
On the whole, the results in Fig. 4 indicate that the thermal stability of both the thermophilic EF-Tu and the G domain are crucially dependent upon residues situated in the N-terminal region of the molecule, within the first half of the G domain. However, we have found that replacement of the first 90 residues of the mesophile (E. coli) EF-Tu with the corresponding region of the T. maritima protein does not result in increased heat stability (data not shown). We interpret this result as indicating that the remarkable heat resistance of the T. maritima EF-Tu is critically dependent upon unique tertiary structural interactions between certain N-terminal residues and distal residues that are absent in the E. coli moiety of the hybrid protein.

Fig. 5. DEAE-cellulose chromatography of extracts prepared from E. coli cells transformed with the plasmid bearing the chimeric G domain sequence. (a) GDP-binding activity of the eluted fractions assayed at 30 °C on 10 μl aliquots; (b) SDS-PAGE analysis of 10 μl aliquots of the same samples.

Fig. 6. Effect of preincubation at increasing temperatures on the GDP-binding capacity of the partially purified chimeric G domain. The samples were treated as described in Fig. 2 and the GDP-binding activity was assayed at 30 °C. NT, Not treated.

mobility. The partially purified chimeric G domain was then incubated at increasing temperatures (between 45 and 90 °C). As Fig. 6 shows, the GDP-binding activity of the chimeric protein was drastically reduced following incubation at 60 °C and disappeared altogether upon incubation at 65 °C in accordance with the gel electrophoretic patterns of crude extracts in Fig. 4(d).

Conclusions

Primary structural comparison and engineering of the tuf gene sequence of the thermophile T. maritima allow the following conclusions. (1) Interactions between the three domains constituting the native EF-Tu play a role in maintaining the heat stability of the molecule. This is supported by the evidence that the isolated G domain of T. maritima EF-Tu is slightly less stable than the full-length protein, in accordance with the situation in E. coli (Jensen et al., 1989). (2) Sequence element(s) that are crucial for heat stability appear to lie within the N-terminal portion of the T. maritima molecule between residues 1 and 90. This is borne out by the fact that replacement of the first 90 amino acids of both the isolated G domain and the full-length EF-Tu with the corresponding sequence of a mesophile (E. coli) EF-Tu gives rise to heat-labile proteins that do not tolerate heating at temperatures above 65 °C (EF-Tu) and 60 °C (G domain). (3) The N-terminal residues involved in thermal stability defy detection, as no amino acid preferences (or unique sequence motifs) appear to distinguish the T. maritima EF-Tu from the corresponding EF-Tu of mesophilic and other thermophilic (Thermus thermophilus and Aquifex pyrophilus) bacteria. Regardless of which residue(s) are involved, however, these appear to act through interaction with distal residues of the molecule. This interpretation is suggested by the fact that hybrid proteins comprising the first 90 residues of T. maritima EF-Tu and the remaining residues of E. coli EF-Tu are thermally labile, as would be expected if the E. coli moiety lacks putative residues involved in (stabilizing) interactions with residues situated in the T. maritima N-terminal portion of the molecule.

The data encourage random mutagenesis experiments focused on the initial portion of the T. maritima tuf gene. Combined functional and structural studies of the modified EF-Tus may cast light on the mechanism(s) that enable the thermophile factor to retain catalytic activity near the temperature of boiling water.

ACKNOWLEDGEMENTS

This work was supported by grants from MURST and by a CNR grant Progetto Finalizzato Biotecnologie e Biostromentazione, sottoprogetto Bioingegneria Molecolare e Cellulare.
REFERENCES


Received 6 December 1995; revised 17 April 1996; accepted 19 April 1996.