Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structures of and interactions between domains of trigger factor from *Thermotoga maritima*

Erik Martinez-Hackert[a] and Wayne A. Hendrickson[a,b]*

[a]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA, and [b]Howard Hughes Medical Institute, Columbia University, New York, NY 10032, USA

Correspondence e-mail: wayne@convex.hhmi.columbia.edu

Received 13 November 2006
Accepted 28 February 2007

PDB References: $tmTF_C$, 2nsa, r2nsasf; $tmTF_N$, SeMet, 2nsb, r2nsbsf; native, 2nsc, r2nscsf.

Trigger factor (TF) is a eubacterial chaperone that associates with ribosomes at the peptide-exit tunnel and also occurs in excess free in the cytosol. TF is a three-domain protein that appears to exist in a dynamic equilibrium of oligomerization states and interdomain conformations. X-ray crystallography and chemical cross-linking were used to study the roles of the N- and C-terminal domains of *Thermotoga maritima* TF in TF oligomerization and chaperone activity. The structural conservation of both the N- and C-terminal TF domains was unambiguously established. The biochemical and crystallographic data reveal a tendency for these domains to partake in diverse and apparently nonspecific protein–protein interactions. It is found that the *T. maritima* and *Escherichia coli* TF surfaces lack evident exposed hydrophobic patches. Taken together, these data suggest that TF chaperones could interact with nascent proteins *via* hydrophilic surfaces.

1. Introduction

Trigger factor (TF) was originally identified as a soluble cytosolic protein in *Escherichia coli* that 'triggered' the folding of the outer membrane porin A (pro-OmpA) into a membrane-assembly competent form (Crooke & Wickner, 1987). Subsequent studies revealed that TF stably bound pro-OmpA with apparent 1:1 stoichiometry (Crooke *et al.*, 1988) and suggested that equimolar binding to the large subunit of the ribosome with 0.3 μM affinity positioned TF to contact nascent polypeptide chains (Lill *et al.*, 1988).

It is now well established that TF binds to the large subunit of the ribosome at proteins L23/L29 near the polypeptide-exit channel (Kramer *et al.*, 2002; Ullers *et al.*, 2003; Blaha *et al.*, 2003; Maier *et al.*, 2003; Ferbitz *et al.*, 2004; Baram *et al.*, 2005; Schlunzen *et al.*, 2005) and that it associates co-translationally with nascent polypeptides (Valent *et al.*, 1995; Hesterkamp *et al.*, 1996).

TF harbors chaperone and peptidyl-prolyl *cis*/*trans*-isomerase (PPIase) activities (Crooke & Wickner, 1987; Hesterkamp *et al.*, 1996; Stoller *et al.*, 1995; Callebaut & Mornon, 1995; Lecker *et al.*, 1989). It is generally accepted that TF interacts with short nascent polypeptides (Hesterkamp *et al.*, 1996; Valent *et al.*, 1995) independently of proline residues (Scholz *et al.*, 1998), presumably recognizing sequences enriched in hydrophobic amino-acid residues (Patzelt *et al.*, 2001). *In vivo*, TF chaperone activity partially overlaps with that of the *E. coli* Hsp70 homolog DnaK (Deuerling *et al.*, 1999; Teter *et al.*, 1999).

The trigger-factor gene (*tig*), which is universally present in bacteria, encodes a three-domain protein comprising an N-terminal ribosome-binding domain, an intermediate FKBP-like PPIase domain and a C-terminal domain of unknown function (Hesterkamp & Bukau, 1996; Hesterkamp *et al.*, 1997; Ferbitz *et al.*, 2004). The three domains are fully conserved.

A number of studies have described TF oligomerization involving the N- and C-terminal domains and have proposed a multi-state equilibrium in which TF binds as a monomer or dimer to the ribosome and exists in monomer–dimer equilibrium in solution, perhaps acting as a binding chaperone in its soluble dimeric form (Patzelt *et al.*, 2002; Blaha *et al.*, 2003; Liu *et al.*, 2005; Liu & Zhou, 2004).

Recently published structures of the ribosome-binding domain of *E. coli* TF (Kristensen & Gajhede, 2003), full-length *E. coli* TF (Ferbitz *et al.*, 2004), a C-terminally truncated *Vibrio cholerae* TF (Ludlam *et al.*, 2004), a ribosome-bound *Deinococcus radiodurans* TF N-terminal domain (Baram *et al.*, 2005; Schlunzen *et al.*, 2005) and a *Mycoplasma genitalium* TF PPIase domain (Vogtherr *et al.*, 2002) have detailed the individual TF domains, the interactions between TF and the ribosome (Ferbitz *et al.*, 2004; Baram *et al.*, 2005; Schlunzen *et al.*, 2005) and an unusually extended TF quaternary structure (Ferbitz *et al.*, 2004; Ludlam *et al.*, 2004). In addition, these studies have identified associations which could describe the proposed monomeric and dimeric species (Ludlam *et al.*, 2004; Kristensen & Gajhede, 2003).

While it is now generally accepted that TF exists in a variety of different states, each of which may well be involved in a distinct function, the specific activity and structural organization of some of these states is not well understood. Here, we have studied the oligomerization states and interactions of the N- and C-terminal domains of *Thermotoga maritima* TF ($tmTF_N$ and $tmTF_C$, respectively) using glutaraldehyde cross-linking and we have analyzed the crystal structures of $tmTF_N$ at 2.2 Å and $tmTF_C$ at 1.7 Å resolution.

2. Materials and methods

2.1. Cloning and purification

The *T. maritima* TF gene was PCR-amplified from genomic DNA purchased from the ATCC and fragments encoding specific portions of the TF protein (SwissProt Q9WZF8) were cloned into the pET24d plasmid (Novagen) by standard molecular-biology techniques. In particular, expression plasmids were prepared to produce $tmTF_N$ (amino-acid residues 1–116) and $tmTF_C$ (residues 243–404); several others were also prepared (Table 1). The resulting plasmids have an initiating ATG codon at the synthetic *Nco*I site and a C-terminal 6×His tag inserted following the synthetic *Sal*I or *Xho*I sites. In these plasmids, expression of the TF domains is controlled by a T7 promoter and a ribosome-binding site. Details of plasmid construction are available upon request.

The TF proteins were expressed from *E. coli* strain BL21-Codon Plus RIL into which appropriate plasmids (notably $ptmTF_C$ and $ptmTF_N$ for $tmTF_N$ and $tmTF_C$, respectively) had been introduced. Cells were grown at 310 K in Luria–Bertrani (LB) media supplemented with 100 mg l^{-1} kanamycin to a cell density corresponding to A_{600} = 0.6. Protein expression was induced by the addition of 0.5 µ*M* isopropyl β-D-thio-galactoside (IPTG) and cells were grown for another 3 h at 310 K. Cells were harvested by centrifugation, resuspended in 20 m*M* Tris pH 8.0, 200 m*M* NaCl and frozen. After thawing, cells were lysed by sonication and centrifuged at 20 000*g* for 30 min. The supernatants were heated to 338 K for 20 min and centrifuged at 5000*g* for 15 min. Selenomethionyl (SeMet) $tmTF_N$ and $tmTF_C$ were produced as above, but following a non-auxotrophic protein-expression protocol (Doublié, 1997).

Table 1
Construct design.

TF, trigger factor; N, N-terminal domain, residues 1–108; L, linker, residues 109–145; P, PPIase domain, residues 153–223; C, C-terminal domain, residues 243–425; ΔP, deleted PPIase domain, residues 146–236; CPA, carboxypeptidase A-digested; MS, mass spectrometry; ecTF, ecTF structure (Ferbitz *et al.*, 2004); primary structure, Swiss-Prot Q9WZF8 (Hesterkamp & Bukau, 1996; Hesterkamp *et al.*, 1997).

Construct	Residue range	Basis for residue choice	Crystals	d_{min} (Å)
TF_N	1–116	Primary structure	Yes	2.2
TF_{NLP}	1–243	ecTF	No	—
TF_{NLC}	1–410ΔP	ecTF	Yes	8/3.0
TF_P	153–243	Primary structure	Yes	8
TF_{LPC}	108–410	ecTF	No	—
TF_{LC}	108–410ΔP	ecTF	Yes	6
TF_C	243–404	Primary structure, CPA	Yes	1.7
TF_{365}	1–365	CPA, MS	No	—
TF_{404}	1–404	CPA, MS	Yes	7.5/3.5
TF_{410}	1–410	TF_C structure	Yes	7.5
TF_{425}	1–425	Full length	Yes	7.5

Supernatants from heat-treated cell lysates were loaded separately onto a HiTrap chelating column (Amersham Biosciences) equilibrated with column buffer (20 m*M* Tris pH 8.0, 200 m*M* NaCl and 5 m*M* imidazole pH 8.0). Each loaded column was washed with five column volumes of column buffer and the proteins were eluted with a linear gradient of 0.0–50 m*M* EDTA. Peak fractions were diluted with H_2O and loaded onto a HiTrapQ anion-exchange column (Amersham Biosciences) equilibrated with 20 m*M* Tris pH 8.0, 50 m*M* NaCl. The proteins were eluted with a linear gradient of 50–600 m*M* NaCl. Peak fractions were concentrated and loaded onto a Superdex 75 gel-filtration column (Amersham Biosciences) equilibrated with 20 m*M* Tris pH 8.0, 200 m*M* NaCl. Purified proteins were dialyzed against a buffer of 10 m*M* Tris pH 8.0, 50 m*M* NaCl and concentrated to a final concentration of approximately 30 mg ml^{-1}.

2.2. Cross-linking and protein identification

Chemical cross-linking assays with purified components were performed to determine whether $tmTF_N$ and $tmTF_C$ form homo-oligomers or hetero-oligomers. Mixtures with different ratios of $tmTF_N$ to $tmTF_C$ and with the isolated components at various concentrations were incubated at 293 K for 5–10 min in the presence of 0.1% glutaraldehyde. Proteins were mixed to give $tmTF_N$:$tmTF_C$ mixtures of final concentrations 4:0, 16:2, 4:2, 4:8 and 0:2 mg ml^{-1}. Reactions

were stopped by adding excess Tris buffer pH 8.5. Samples were analyzed by SDS–PAGE and, to determine their composition, protein bands that appeared exclusively in the $tmTF_N$–$tmTF_C$ cross-linked mixture were analyzed using peptide mapping and liquid chromatography/mass spectrometry (LC-MS/MS). In addition, to determine the molecular weight of the cross-linked homomeric and heteromeric species, solutions of untreated and cross-linked $tmTF_N$, $tmTF_C$ and $tmTF_N$–$tmTF_C$ were analyzed by electrospray ionization (ESI) mass spectrometry.

Table 2
Crystallographic data.

Values in parentheses are for the outermost shell.

Crystal	$tmTF_C$	$tmTF_C$	$tmTF_{N1}$			$tmTF_{N2}$
Protein	SeMet	SeMet	SeMet	SeMet	SeMet	Native
X-ray source	X4A	RA†	X4A	X4A	X4A	X4A
Wavelength (Å)	0.97920	1.5418	0.97917	0.97884	0.96864	0.91994
Space group	$C222_1$	$C222_1$	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
Unit-cell parameters						
a (Å)	53.84	54.05	48.09	48.09	48.09	51.21
b (Å)	74.21	74.23	86.55	86.55	86.55	75.87
c (Å)	90.07	90.46	32.95	32.95	32.95	30.51
Z_a‡	1	1	1	1	1	1
Solvent content (%)	47	47	53	53	53	48
d_{min} (Å)	1.70	2.40	3.20	3.20	3.20	2.20
Phasing	SAD	SIRAS	MAD edge	MAD peak	MAD remote	MR
Derivative		NaI				
Merging R value	0.043 (0.157)	0.030 (0.054)	0.039 (0.147)	0.030 (0.137)	0.029 (0.155)	0.054 (0.269)
Overall $I/\sigma(I)$	40.0 (10.5)	36.5 (18.2)	49.9 (7.1)	37.2 (4.5)	24.6 (3.3)	34.1 (7.8)
Completeness	99.7 (99.8)	90.1 (82.4)	93.5 (74.0)	91.6 (68.5)	78.1 (47.1)	99.4 (99.2)
No. of reflections	19614	6645	2373			6440
R value (working set)	0.207		0.236			0.194
Free R value	0.224		0.270			0.260
Bond-length ideality (Å)	0.011		0.013			0.011
Bond-angle ideality (°)	1.183		1.137			1.164
Ramachandran analysis (%)						
Favored (%)	99.4					98.1
Outliers (%)	0.0					0.0
PDB code	2nsa		2nsb			2nsc

† RA, rotating anode. ‡ Z_a, number of protein chains per asymmetric unit.

2.3. Crystallization and data collection, $tmTF_N$

Crystallization of $tmTF_N$ was carried out at 293 K using the hanging-drop vapour-diffusion method. Crystals of both native and SeMet $tmTF_N$ were obtained in 5 d after mixing the protein solution and reservoir buffer (10–15% PEG 4000, 150–250 m*M* KCl and 100 m*M* Tris pH 7.5) in a 1:1 ratio. Crystals grew in space group $P2_12_12$, but had substantial variation in their unit-cell parameters despite their similar morphology. For cryoprotection, crystals were transferred for 10 min to cryobuffer (15% PEG 4000, 25% glycerol, 250 m*M* KCl, 100 m*M* Tris pH 7.5) and flash-frozen at 100 K in a nitrogen cryostream. All X-ray data sets were collected at 100 K at NSLS beamline X4A using a CCD Quantum 4 detector, including a three-wavelength data set from an SeMet crystal $tmTF_{N1}$. Diffraction data were processed and reduced with *DENZO* and *SCALEPACK* (Minor *et al.*, 2002). Statistics for refinement data sets are given in Table 2.

2.4. Structure determination and refinement, $tmTF_N$

The structure was solved by a combination of multiple-wavelength anomalous diffraction (MAD; Hendrickson *et al.*, 1990) and molecular-replacement methods. Coordinates for all three Se atoms were located from a 3.2 Å MAD data set using the program *SOLVE* (Terwilliger & Berendzen, 1999). The resulting electron-density map at 3.2 Å was interpretable but of very poor quality, even following density modification (Table 2).

An initial model containing 113 of the 126 residues was built using the program *RESOLVE* (Terwilliger & Berendzen, 1999) and rebuilt manually using the program *O* (Jones *et al.*, 1991). This preliminary model was refined against the $tmTF_{N1}$ data using *CNS* (Brünger *et al.*, 1998) and the partially refined model was used as input for a molecular-replacement search against the 2.2 Å data set from native crystal $tmTF_{N2}$ using the program *Phaser* (Read, 2001). The resulting model was subjected to two rounds of *ARP*/*wARP* (Perrakis *et al.*, 1999). Firstly, the *ARP*/*wARP* module 'improvement of maps by atoms update and refinement' was used to produce an improved model containing a large number of dummy atoms. Secondly, the improved model was input into the *ARP*/*wARP* module 'automated model building starting from existing model' to obtain a *de novo* trace of the molecule. Automated model building at 2.2 Å resolution produced an almost complete model and revealed several frame-shift errors in the low-resolution model. Additional residues were manually built with the program *O* and the model was subjected to iterative cycles of manual rebuilding and conjugated gradient minimization, simulated annealing, and individual *B*-factor refinement using the programs *O*, *CNS* and *REFMAC*. TLS (Winn *et al.*, 2001), anisotropic *B*-factor and bulk-solvent corrections were applied. The final model includes 109 amino-acid residues and 138 water molecules. Stereochemistry checks with the program *MolProbity* indicate that the refined model is in excellent agreement with expectations for models within this resolution range (Lovell *et al.*, 2003). Ultimately, the $tmTF_{N2}$ model was replaced back for further refinement of the $tmTF_{N1}$ structure. Statistics for the final models are given in Table 2.

2.5. Crystallization and data collection, $tmTF_C$

Crystallization of $tmTF_C$ was carried out at 293 K using the hanging-drop vapour-diffusion method. Both the native and

SeMet proteins crystallized in 3–5 d after mixing the protein solution and reservoir buffer [1.35 *M* $(NH_4)_2SO_4$, 10% glycerol and 100 m*M* Tris pH 8.5] in a 1:1 ratio. Crystals grew in space group $C222_1$ with one molecule per asymmetric unit. For cryoprotection, crystals were transferred for 1 min to cryobuffer [1.35 *M* $(NH_4)_2SO_4$, 15% glycerol, 15% ethylene glycol and 100 m*M* Tris pH 8.5] and flash-frozen at 100 K in a nitrogen cryostream. An iodide derivative was generated by soaking SeMet crystals for 30 s in cryobuffer solution plus 250 m*M* NaI. All X-ray data sets were collected at 100 K in-house on an R-AXIS IV image-plate detector using a Rigaku RU-H3R rotating-anode X-ray generator and on a CCD Quantum 4 detector at NSLS beamline X4A. Diffraction data were processed and reduced with *DENZO* and *SCALEPACK* (Minor *et al.*, 2002); statistics for all data sets are given in Table 2.

2.6. Structure determination and refinement, $tmTF_C$

The $tmTF_C$ structure was solved by a combination of single-wavelength anomalous diffraction (SAD) and single isomorphous replacement with anomalous scattering (SIRAS) data. The SAD data were measured from an SeMet crystal at the wavelength of peak Se *K*-edge absorption and data from an iodide-derivatized SeMet crystal were measured using Cu $K\alpha$ radiation. A clear solution of two Se atoms and three I atoms was obtained by the program *SOLVE* using a combination of SAD and SIRAS data. Phases modified with the program *RESOLVE* produced a readily interpretable map, which clearly revealed the greater part of the structure (Table 2).

An initial model containing 137 of the 171 residues was built using the program *ARP*/*wARP* and 30 additional residues were manually built using the program *O*. The model was subjected to iterative cycles of manual rebuilding and conjugated-gradient minimization, simulated annealing and individual *B*-factor refinement using the programs *O*, *CNS* and *REFMAC*. TLS, anisotropic *B*-factor and bulk-solvent corrections were applied. The final model includes 167 amino acids, 208 water molecules and three sulfate ions. Stereochemistry checks indicate that the refined model is in excellent agreement with expectations for models within this resolution range. Statistics for the final model are given in Table 2.

3. Results

3.1. Domain dissection

To determine the structure of *T. maritima* trigger factor, we generated several constructs for protein expression, purification and crystallization (Table 1). Our original rationale in TF-construct design and domain definition was based on the assumption that TF domains are laid out in a linear sequence (Hesterkamp *et al.*, 1997; Hesterkamp & Bukau, 1996). In this scheme, the approximate boundaries of the *T. maritima* TF domains correspond to residues 1–116 for the N-terminal domain ($tmTF_N$), residues 158–231 for the PPIase domain ($tmTF_P$) and residues 243–425 for the C-terminal domain ($tmTF_C$) (Table 1). We generated, purified and attempted crystallization of constructs which included the single domains and combinations of neighboring domains, as well as the full-length TF. We obtained large single crystals for each of the single-domain constructs. Crystals of the N-terminal ($tmTF_N$) and C-terminal ($tmTF_C$) domain diffracted sufficiently to achieve atomic resolution; however, crystals corresponding to the PPIase-domain construct diffracted only poorly. Large single crystals of the full-length protein grew readily, but with one single exception these diffracted poorly.

Using the *E. coli* TF structure as a template (Ferbitz *et al.*, 2004), we redesigned some of our constructs. Residues 108–145, which link the N-terminal domain (N) to the PPIase domain (P), are an integral component of the C-terminal domain (C). The boundaries of the *T. maritima* TF C-terminal domain therefore correspond approximately to residues 108–425 and the PPIase domain could be considered as an inser-

Table 3
Mass-spectrometric identification of cross-linked proteins.

		Oligomerization state				
Construct	Gel lane†	1	2	3	4	5
$tmTF_N$ (kDa)	6	13.7	28.5	42.4	56.3	68.8 (trace)
$tmTF_C$ (kDa)	2	20.9	41.8	62.8 (trace)		

		Oligomerization state		
Construct	Gel lane†	1+1	2+1	2+2
$tmTF_N$ + $tmTF_C$ (kDa)	4	34.7	49.9	70.7

† Corresponding gel lane fraction used in mass-spectrometric experiment (see Fig. 1).

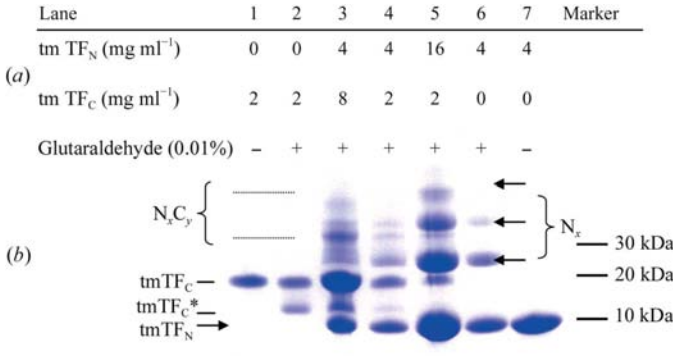

Figure 1
SDS–PAGE analysis of cross-linked *T. maritima* TF N- and C-terminal domains. (*a*) Compositions of solutions analyzed by electrophoresis in (*b*). The $tmTF_N$ and $tmTF_C$ proteins were mixed in varying ratios of 4:0, 16:2, 4:2, 4:8 and 0:2 mg ml^{-1}, respectively, for $tmTF_N$:$tmTF_C$ mixtures and cross-linked using glutaraldehyde to determine their *in vitro* aggregation state. (*b*) SDS–PAGE gel of the cross-linking experiment defined in (*a*). The $tmTF_N$ protein forms homomultimers (lanes 5 and 6). Additionally, $tmTF_N$ binds $tmTF_C$ in a concentration-dependent manner, producing several higher molecular-weight heteromeric complexes (lanes 3 and 4). Formation of the heteromeric complexes results in loss of the multimeric $tmTF_N$ species (lanes 3 and 4) and in the loss of a cross-linked $tmTF_C$ species ($tmTF_C$*, lanes 2, 4 and 5), which is dimeric by mass spectrometry (Table 3) but which migrates with anomalously high mobility as cross-linked. The molecular weights of $tmTF_N$ and $tmTF_C$ are 12.0 and 20.0 kDa, respectively.

tion in the C-terminal domain. We generated a second set of domain constructs which took into consideration the expanded C-terminal domain definition. We also produced constructs that lacked the N-terminal domain, the PPIase domain or both of these domains ($tmTF_{LPC}$, $tmTF_{NLC}$ and $tmTF_{LC}$, respectively). Large single crystals of $tmTF_{LC}$ and $tmTF_{NLC}$ grew readily but diffracted poorly (Table 1).

3.2. Association of $tmTF_N$ and $tmTF_C$

Published crystallographic and cross-linking data have indicated that the *E. coli* and *V. cholerae* TF N- and C-terminal domains may participate in homomeric and/or heteromeric associations (Patzelt *et al.*, 2002; Kristensen & Gajhede, 2003; Ludlam *et al.*, 2004). To test for the possibility of homo- and hetero-oligomerization of $tmTF_N$ and $tmTF_C$ *in vitro*, we used glutaraldehyde cross-linking. Cross-linked samples were separated by SDS–PAGE (Fig. 1) and subjected to peptide mapping; molecular weights were determined by MALDI–TOF mass spectrometry (Table 3).

Our data show that recombinant $tmTF_N$ forms homo-multimers and that $tmTF_C$ forms homodimers in solution. Additionally, $tmTF_N$ binds $tmTF_C$ in a concentration-dependent manner, producing several higher molecular-weight heteromeric complexes, which in all likelihood correspond to N_1C_1, N_2C_1, N_2C_2, N_3C_2 *etc*. Interestingly, formation of the heteromeric complexes appears to result in the loss of the multimeric $tmTF_N$ species, indicating that the multimeric $tmTF_N$ complexes and heteromeric complexes may form in an exclusive way (Fig. 1).

3.3. Structure-based sequence alignments

We generated structure-based sequence alignments of the TF N- and C-terminal domains (TF_N and TF_C) using the program *T-Coffee* (Notredame *et al.*, 2000). We performed a combination of pairwise structure–structure and structure–sequence alignments using the $tmTF_N$, $ecTF_N$, $vcTF_N$ and $drTF_N$ structures for the N-terminal TF domain and the $tmTF_C$ and $ecTF_C$ structures for the C-terminal TF domain. The resulting multiple sequence alignments are shown in Fig. 2. We find that the *V. cholerae* and *E. coli* TF_N proteins share a strong sequence identity of 61%. The identity between the *T. maritima* and *E. coli* TF_N domains is significantly lower: only 22% of the residues are identical. Similarly, the *T. maritima* and *E. coli* TF_N domains have low sequence identities of 20 and 27%, respectively, when aligned with *D. radiodurans* TF_N. The *T. maritima* and *V. cholerae* TF_N domains appear to be the most removed, with a sequence identity of 15% over the entire TF_N domain.

The C-terminal TF domains (TF_C) of *T. maritima* and *E. coli* share almost negligible sequence identities of 12% over the C-terminal 170 amino acids; however, the structure-based sequence alignment reveals that the secondary-structure elements match up with surprising precision, indicating that the TF_C domain is conserved in all the organisms that contain the *tig* gene.

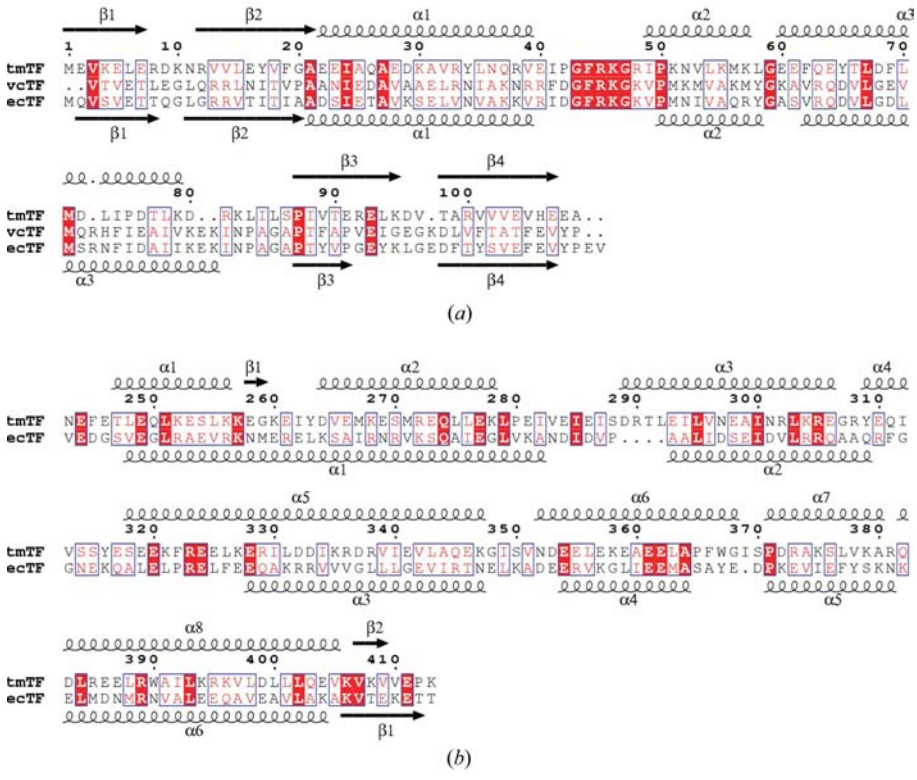

Figure 2
Structure-based sequence alignments. (*a*) Alignment of the TF N-terminal domains of *T. maritima*, *V. cholerae* and *E. coli*. The ribosome-binding loop contains the most highly conserved amino acids. When TF is not associated with the ribosome, this structure is highly variable. (*b*) Alignment of the TF C-terminal domains of *T. maritima* and *E. coli*. The structure-based sequence alignment reveals a strong conservation of the secondary structure. Only 20 out of 169 amino acids are identical.

3.4. Structure of $tmTF_N$

We have solved crystal structures of the *T. maritima* TF N-terminal domain ($tmTF_N$) using a combination of multiple-wavelength anomalous dispersion (MAD) and molecular-replacement (MR) phasing. We used two different crystals in this analysis: SeMet $tmTF_{N1}$ and native $tmTF_{N2}$ (Table 2). Both crystals belong to the orthorhombic space group $P2_12_12$, with one $tmTF_N$ molecule per asymmetric unit, but they are non-isomorphous, with the unit-cell volume of TF_{N2} reduced 14% from that of TF_{N1}. The structure was partially refined at 3.2 Å resolution from the SeMet MAD phasing of crystal TF_{N1} and then further refined at 2.2 Å resolution after replacement into crystal TF_{N2}, where crystal packing is similar. The 2.2 Å crystal structure is well defined throughout with an average coordinate uncertainty, estimated by the

cross-validated Luzzati method, of 0.24 Å. All residues lie in the most favored or allowed regions of the Ramachandran diagram.

The $tmTF_N$ domain has an elongated $\alpha+\beta$ structure delimited by a four-stranded antiparallel β-sheet on one face of the molecule and by one short and two long α-helices on the opposite face (Fig. 3*a*). Overall, the structure is very similar in its organization to the structures of the *E. coli*, *V. cholerae* and *D. radiodurans* TF N-terminal domains [$ecTF_N$ (Ferbitz *et al.*, 2004; Kristensen & Gajhede, 2003), $vcTF_N$ (Ludlam *et al.*, 2004) and $drTF_N$ (Baram *et al.*, 2005; Schlunzen *et al.*, 2005), respectively] (Fig. 3*c*). However, the relative orientation of the first α-helix and the ribosome-binding loop vary significantly between species. We used the program *ESCET* (Schneider, 2002) to determine the conformationally flexible and invariant regions of TF_N and found that the residues forming the helical face of the molecule, including the ribosome-binding loop, are among the most flexible in the structure. Using *LSQKAB* (Kabsch, 1976), we calculated an

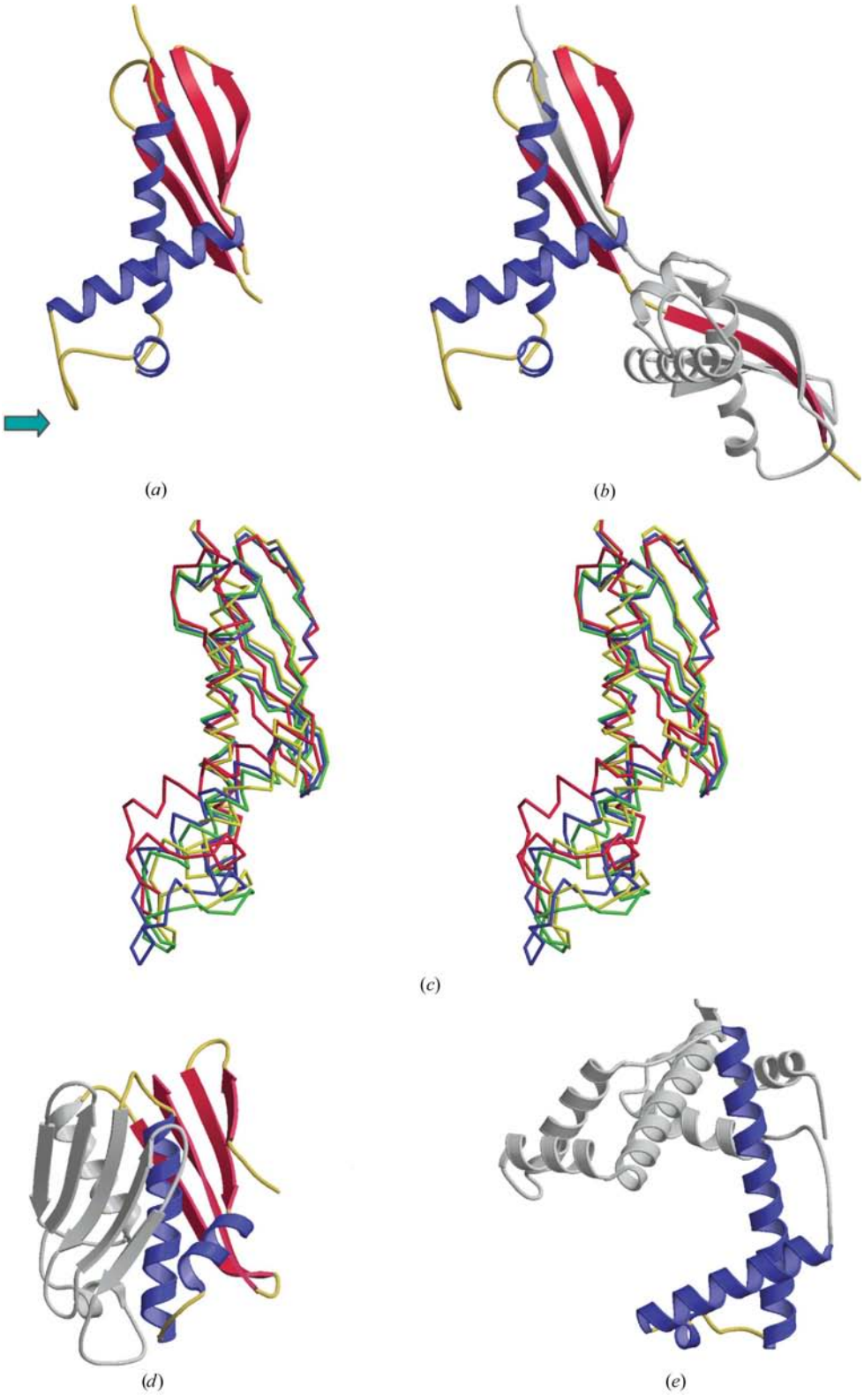

Figure 3
Structural analysis of the N-terminal domain of *T. maritima* TF ($tmTF_N$). (*a*) Ribbon diagram of a $tmTF_N$ protomer based on the 2.2 Å resolution crystal structure. The domain has an elongated $\alpha+\beta$ structure, which includes a four-stranded antiparallel β-sheet (red) on one face of the molecule, two large α-helices (blue) on the opposing face and nonregular segments (yellow). The cyan arrow marks the ribosome-binding loop. This model was built by replacing the C-terminal β-strand with that of a symmetry-related $tmTF_N$. We refer to this pseudo-protomer in our discussions of monomeric $tmTF_N$. (*b*) The C-terminal β-strand (red) from one $tmTF_N$ chain is swapped with the corresponding β-strand of a symmetry-related molecule (grey). (*c*) Stereo diagram showing a superposition of C^{α}-backbone traces of four TF N-terminal domain structures: $tmTF_N$ (red), $ecTF_N$ (blue), $vcTF_N$ (green) and $drTF_N$ (yellow). The orientations of the first TF α-helix and the ribosome-binding loop account for the most pronounced differences among the four TF_N structures. (*d*) Ribbon diagram of *E. coli* Hsp33 oriented to have portions in common with $tmTF_N$ (strands, red; helices, blue; nonregular segments, yellow) placed as in (*a*). Elements not included in $tmTF_N$ are colored grey. The $tmTF_N$ and *E. coli* Hsp33 structures are surprisingly similar, with an r.m.s.d. of 2.8 Å. (*e*) Ribbon diagram of the $tmTF_C$ structure oriented with one of its helical protrusions (blue) placed as in the helical portion of $tmTF_N$ in (*a*). The $tmTF_N$ α-helical structure superimposes very well with the helical protrusions that form the core of the $tmTF_C$ structure.

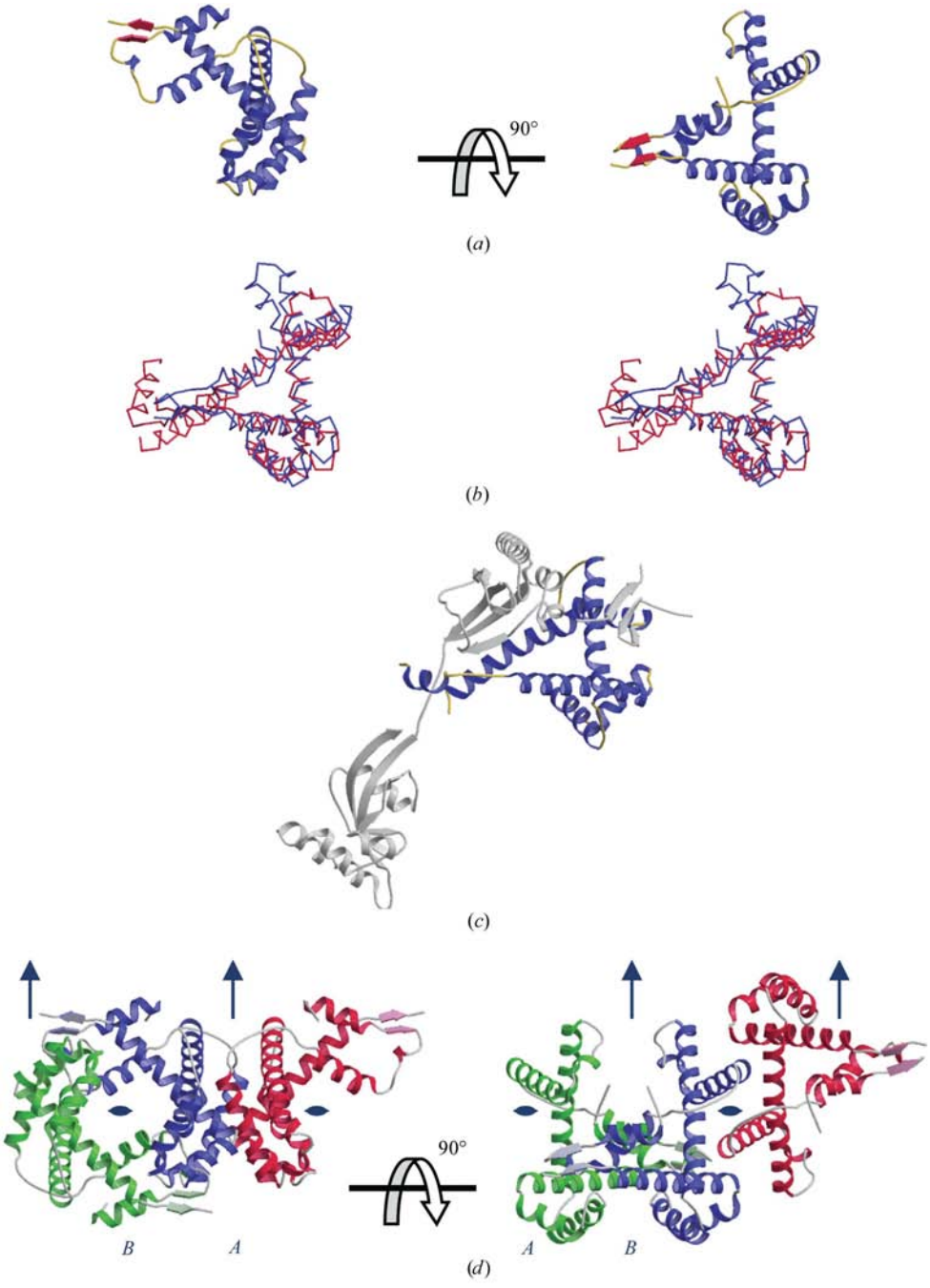

Figure 4
Structural analysis of the C-terminal domain of *T. maritima* TF ($tmTF_C$). (*a*) Ribbon diagram of $tmTF_C$ based on the 1.7 Å resolution crystal structure. The $tmTF_C$ domain consists of several α-helices (blue), a two-stranded parallel β-sheet (red) and nonregular segments (yellow). Two orthogonal views are shown rotated about the horizontal axis. The view on the right is approximately down the quasi-threefold axis. (*b*) Stereo diagram showing a superposition of TF C-terminal domain structures: $tmTF_C$ (blue) and $ecTF_C$ (red). The view is along the quasi-threefold axis of the trivet-like structure, with its prefoldin-like cavity underneath the three crossing helices at the apex. The 24 N-terminal $tmTF_C$ residues attain a different conformation in the truncated $tmTF_C$ domain. (*c*) Ribbon diagram of SurA oriented as in (*a*) after superposition of common helical elements (blue) and nonregular segments (yellow). The structural similarity between $tmTF_C$ and SurA comprises all of $tmTF_C$ and the N and C domains of SurA. Two intervening parvulin domains of SurA are colored grey. (*d*) Intimate lattice contacts. The central molecule (blue) interacts both through interface *A* with one symmetry mate (red) and through interface *B* with another symmetry mate (green). The orientations of the central molecules are exactly as in (*a*) above. Crystallographic dyad axes are represented by black eye-shaped symbols for dyads perpendicular to the page and by vertical arrow symbols for dyads in the plane of the page.

r.m.s.d. of 2.07 Å between the $tmTF_N$ and $ecTF_N$ invariant regions determined by *ESCET* and an r.m.s.d. of 3.13 Å for the entire TF_N domain.

The C-terminal β-strand in this structure of $tmTF_N$ is swapped with the corresponding β-strand of a symmetry-related molecule (Fig. 3*b*). This most striking difference between the $tmTF_N$ structure and other TF_N structures is perhaps an artefact of the particular $tmTF_N$ construct and is likely to be one reason for homomultimer formation *in vitro*. Other examples of artefactual strand swapping have been observed, including pH-induced strand swapping in bovine RNAse A (Liu & Eisenberg, 2002). The $tmTF_N$ protein exists in solution as a mixture of monomers and dimers, which can be separated cleanly by size-exclusion chromatography. We crystallized from the predominant dimer fraction. This is consistent with presumably artefactual strand-swapped dimers having been preformed in solution. We built a model of monomeric $tmTF_N$ by replacing the C-terminal β-strand with that of the symmetry-related $tmTF_N$ (Figs. 3*a* and 3*b*). The structure for this pseudo-protomer is the same as those of other TF N-terminal domains (Fig. 3*c*) and we use this model in all subsequent analyses and discussions.

We searched the Protein Data Bank for structures that are similar to $tmTF_N$ using the *DALI* server (Holm & Sander, 1994) and found the recurring selection of Hsp33s and small heat-shock proteins (Graumann *et al.*, 2001; Vijayalakshmi *et al.*, 2001; Kim *et al.*, 2001; Kim, Kim & Kim, 1998). The HSP33 and $tmTF_N$ structures are surprisingly similar, with an r.m.s.d. of 2.8 Å across 76 aligned residues and an 8% sequence identity (Kim *et al.*, 2001; Fig. 3*d*).

The α-helical structure which protrudes from the core of the TF_N domain strongly resembles

the helical protrusions that form the core of the $tmTF_C$ structure. For example, we superimposed the two $tmTF_N$ α-helices (residues 23–39 and 58–72) with helices αc3 and αc5 from the second $tmTF_C$ protrusion (residues 288–304 and 319–333). The overlapping segments superimpose strikingly well with an r.m.s.d. of 1.17 Å (Fig. 3*e*).

3.5. Structure of $tmTF_C$

We have solved the crystallographic structure of the *T. maritima* TF C-terminal domain ($tmTF_C$) using a combination of single-wavelength anomalous dispersion (SAD) and single isomorphous replacement and anomalous scattering (SIRAS) phasing. SIRAS data were collected to 2.4 Å Bragg spacing in-house from SeMet-$tmTF_C$ crystals that had been soaked with NaI-containing buffers. SAD data from crystals of the SeMet protein were collected at NSLS beamline X4A and the structure was refined to 1.7 Å using these data (Table 2). The 1.7 Å structure is well defined throughout and the stereochemistry is in excellent agreement with expected values; for example, the backbone conformations of all residues lie in the most favored or allowed regions of the Ramachandran diagram. The average coordinate uncertainty, estimated by the cross-validated Luzzati method, is 0.21 Å.

The three-dimensional structure of $tmTF_C$ is illustrated in Fig. 4(*a*). This is a mostly α-helical protein with two short parallel β-strands inserted after the first and last helices. The domain consists of one short and seven long α-helices: six long α-helices create a framework composed of three helical protrusions, referred to elsewhere as 'arms' and 'back' (Ferbitz *et al.*, 2004) or lobes (Schulze-Gahmen *et al.*, 2005). The three helical protrusions assume quasi-threefold symmetry to produce a trivet-like structure; they combine to harbor a large cavity in their center. Herein, the $tmTF_C$ structure vaguely resembles a miniature form of prefoldin, a chaperone that harbors a large central cavity bordered by six large α-helical coiled coils (Martin-Benito *et al.*, 2002; Siegert *et al.*, 2000).

The structures of $tmTF_C$ and $ecTF_C$ are practically identical in the layout of their secondary-structural elements. A superposition using the program *TOP* (Lu, 2000) gives an r.m.s.d. of 2.0 Å across 89 matching residues. The superposition does not include the first 24 $tmTF_C$ amino-acid residues, which belong to a structural element in the full-length structure that includes a linker that is distal in primary sequence (Fig. 4*b*). The linker connects the N-terminal and PPIase domains. Not surprisingly, these 24 residues attain a different conformation in the truncated $tmTF_C$ domain, folding backwards and taking a position occupied by the linker present in the full-length structure.

The structural similarity between $tmTF_C$ and the *V. cholerae* TF C-terminal domain ($vcTF_C$) is marginal. The $vcTF_C$ protein was truncated at the C-terminus, thereby essentially removing the last helical protrusion (Ludlam *et al.*, 2004; Ferbitz *et al.*, 2004; Schulze-Gahmen *et al.*, 2005). This truncation probably leads to misfolding of the $vcTF_C$ domain.

A search of the Protein Data Bank for structures that are similar to $tmTF_C$ using the *DALI* server yielded two credible hits with *Z* scores of 7.0 and 8.8 and r.m.s.d.s of 4.1 and 3.7 Å, respectively: the periplasmic chaperone SurA (Bitto & McKay, 2002) and a protein of unknown function called mpn555 (Schulze-Gahmen *et al.*, 2005). The structural similarity between $tmTF_C$ and SurA comprises all of $tmTF_C$ and the N and C domains of SurA (Fig. 4*c*).

The $tmTF_C$ crystal packing is very intimate (Fig. 4*d*), such that each TF_C domain interacts extensively with two symmetry

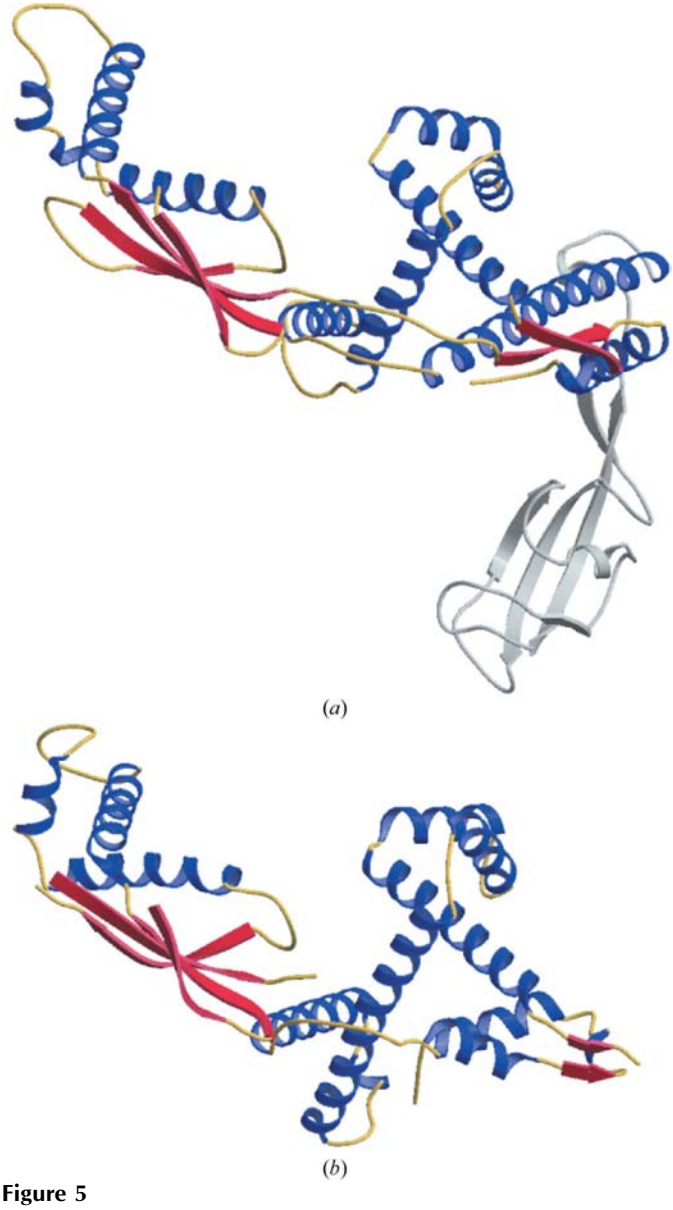

Figure 5
Comparison of N- and C-terminal domains of *T. maritima* TF with those in intact *E. coli* TF. (*a*) Ribbon diagram of intact *E. coli* TF. The coloring code is α-helices, blue; β-strands, red; non-regular segments, yellow; elements not included in *T. maritima* domains are shown in grey. (*b*) Ribbon diagrams of $tmTF_N$ (left) and $tmTF_C$ (right); coloring is as in (*a*). Each domain is positioned as it superimposes onto the corresponding domain of *E. coli* TF as shown in (*a*). A tmTF PPIase-domain structure is not available and has thus not been included in the superposition of individual *T. maritima* TF domains onto the *E. coli* TF structure.

mates in the $C222_1$ lattice. One contact is about a crystallographic dyad parallel to the *a* axis (interface *A*) and the other is about an orthogonal dyad parallel to the *b* axis (interface *B*). Totals of 2450 and 2500 Å^2 of molecular surface area are buried in interfaces *A* and *B*, respectively. These buried areas exceed those in many authentic protein–protein associations (Jones & Thornton, 1996) and might be taken to correspond to relevant dimers. Dimers have been observed in solution for $ecTF_C$ (Patzelt *et al.*, 2002) and also here by cross-linking for $tmTF_C$. Portions of the surfaces of both $tmTF_C$ interfaces are also found at the contacts between domains in intact ecTF (Ferbitz *et al.*, 2004; Fig. 5), however, so the main relevance of these crystal interfaces may concern the general proclivity for protein association by TF domains.

4. Discussion

4.1. Structural relationships to other proteins

TF is a modular three-domain protein which is universally distributed in eubacteria. The intermediate PPIase domain of TFs is a commonplace module, being present in all forms of life except viruses. The N- and C-terminal domains, on the other hand, appear to be rarely utilized structural modules that are replicated only in three known structures, all of which are eubacterial: Hsp33 for the N-terminal domain and SurA and mpN555 for the C-terminal domain. Interestingly, Hsp33 and SurA are both bona fide molecular chaperones, suggesting that perhaps both the N- and C-terminal domains contribute independently to chaperone activity. Experiments on the *in vivo* and *in vitro* activities of isolated TF domains support this view (Kramer, Rutkowska *et al.*, 2004; Genevaux *et al.*, 2004; Kramer, Patzelt *et al.*, 2004; Merz *et al.*, 2006).

4.2. Structural variability

Several TF structures have been published recently (Ludlam *et al.*, 2004; Ferbitz *et al.*, 2004; Kristensen & Gajhede, 2003; Baram *et al.*, 2005; Schlunzen *et al.*, 2005; Vogtherr *et al.*, 2002). We analyzed the various TF structures and found clear structural similarities of the N- and C-terminal domains of TF from different organisms, despite often limited sequence conservation. The N-terminal domains compare with r.m.s.d.s of 2.0–3.0 Å (Fig. 3*c*) and the C-terminal domains of tmTF and ecTF have an r.m.s.d. of 2.0 Å over 89 residues (Fig. 4*b*). Notably, the individual *T. maritima* domains described here compare well with the corresponding domains in intact *E. coli* TF (Fig. 5). The overall sequence identities between TFs from distant organisms range from 21% between *T. maritima* and *E. coli* or *V. cholerae* TFs to 27% for the more closely related *D. radiodurans* and *E. coli* TFs. The close TF orthologs from *V. cholerae* and *E. coli* have a much higher sequence identity of 70%.

A comparison of TF N-terminal domain sequences identifies *T. maritima* and *V. cholerae* as the most distant orthologs, sharing only 15% sequence identity. Not surprisingly, the *T. maritima* and *V. cholerae* TF N-terminal domains appear at both ends of a structural spectrum defined by the relative orientation of the first helix and the ribosome-binding loop (Fig. 3). The *V. cholerae* and *E. coli* TF_N structures virtually overlap, as would be expected for structures of proteins that share significant sequence identities.

The TF C-terminal domain includes the least conserved sequences of the *tig* gene, leading to the original speculation that this domain may be absent in some species (Kristensen & Gajhede, 2003). In addition, the structure of a C-terminally truncated *V. cholerae* TF appeared to support the notion of a divergent structure for this domain (Ludlam *et al.*, 2004). However, superposition of the TF C-terminal domains from *T. maritima* and *E. coli* reveal a clear structural kinship between the two proteins (Fig. 4), despite only 12% sequence identity (Fig. 2), suggesting that this domain is a fully conserved component of all TFs.

4.3. Surface properties

It is commonly assumed that molecular chaperones recognize and bind exposed hydrophobic residues in non-native proteins by way of a hydrophobic surface or patch (Craig *et al.*, 1994; Kim, Kim, Yokota *et al.*, 1998; Zhu *et al.*, 1996; Flynn *et al.*, 1991). The surface properties of TF therefore pose an intriguing dilemma. While some have suggested that the TF surface includes hydrophobic patches with which it may interact with non-native proteins (Ferbitz *et al.*, 2004; Baram *et al.*, 2005), others have failed to detect the presence of such a surface (Kristensen & Gajhede, 2003).

We analyzed the surface properties of *T. maritima* and *E. coli* TF using the program *GRASP* (Nicholls *et al.*, 1993) and failed to find convincing evidence of an exposed hydrophobic patch or surface (Fig. 6). Although small hydrophobic surface patches clearly exist, the role of these surfaces in binding non-native proteins seems questionable given their apparently random distribution throughout the various TF structures. This could mean that TF might interact with proteins *via* hydrophilic surfaces, as suggested for the tubulin cofactor A *Saccharomyces cerevisiae* ortholog Rbl2p (Steinbacher, 1999).

4.4. Oligomerization

E. coli TF exists in monomer–dimer equilibrium in solution. The approximate K_d of TF dimerization is 18 μM (Patzelt *et al.*, 2002). Cross-linking data suggest that TF binds to the ribosome in the form of a monomer (Patzelt *et al.*, 2002), while neutron scattering data indicate that TF binds to the ribosome in the form of a homodimer (Blaha *et al.*, 2003). It has been proposed that the TF monomer–dimer equilibrium is physiologically relevant and that the monomeric and dimeric forms have separate functions (Patzelt *et al.*, 2002). The nature and role of TF dimerization, however, remain elusive.

We attempted to shed light on the role of the *T. maritima* TF N- and C-terminal domains in TF oligomerization by using chemical cross-linking in addition to X-ray crystallography. Our cross-linking data reveal a succession of self-associations of TF N-terminal domains, N_2–N_5. We also observe a cross-linked homodimer of C-terminal domains. The cross-linking of

$tmTF_N$–$tmTF_C$ mixtures leads to the formation of heteromeric species corresponding to N_1C_1, N_2C_1 and N_2C_2 and the disruption of higher order associations of N-terminal domains (Fig. 1; Table 3). The N_2 species that was crystallized proved to be a strand-swapped dimer, presumably an artefact of truncation of no biological consequence. Strand swapping alone, however, cannot explain the higher order cross-linked species. Additional contacts, perhaps also revealed in crystal packing, may also preform in solution.

Crystal-packing analysis of various TF structures has led to conflicting conclusions regarding the architecture of a possible TF dimer (Kristensen & Gajhede, 2003; Ludlam *et al.*, 2004). Our data provide evidence for other different dimerization models. Which dimerization model is correct? We believe that all the published dimerization models are plausible; however, these dimerization models may not necessarily reflect TF dimerization *in vivo*, but rather may reveal nonspecific protein–protein interactions, analogous to TF–substrate

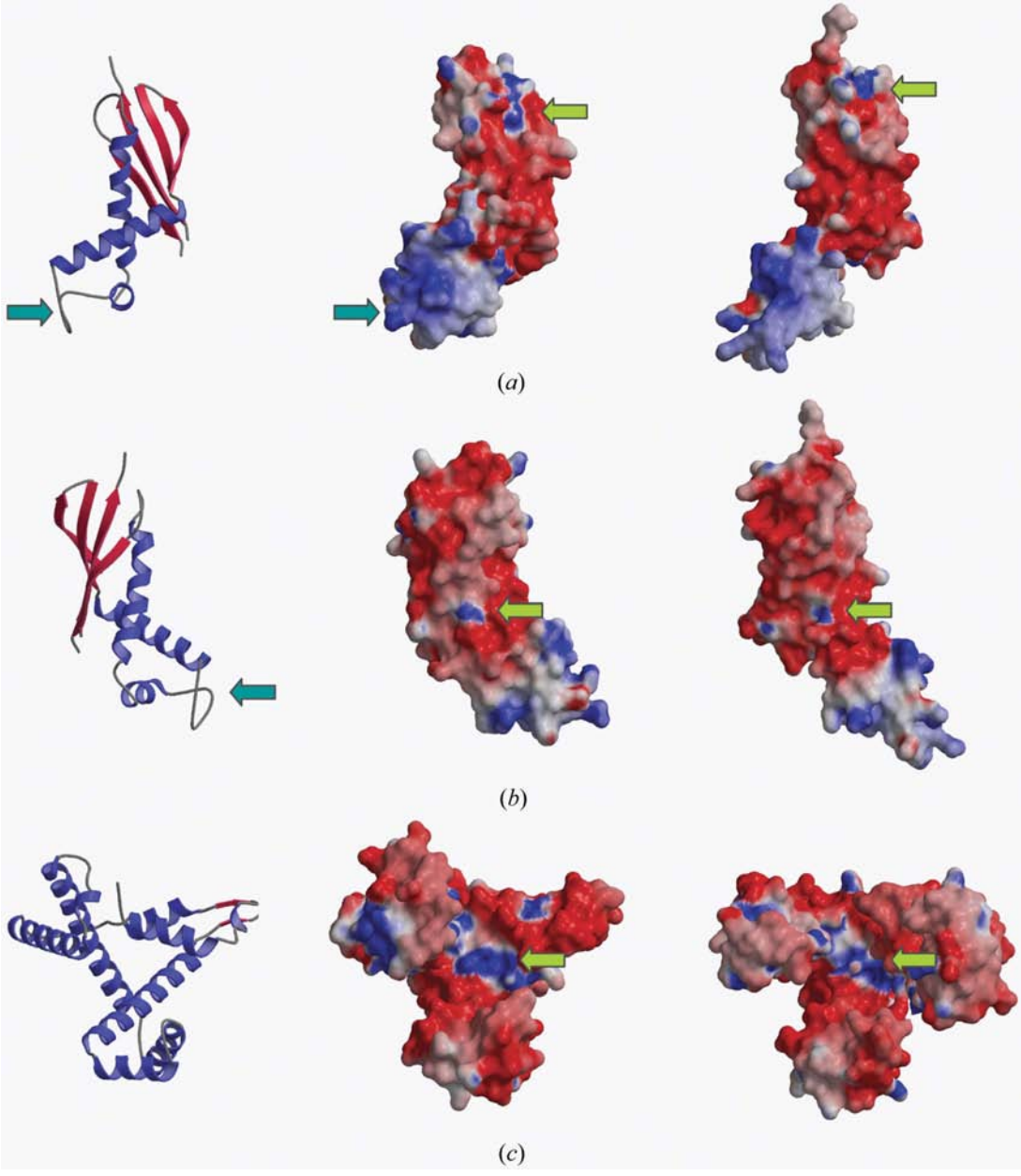

Figure 6
Electrostatic surface potential of the TF N- and C-terminal domains calculated with the program *GRASP*. (*a*) TF N-terminal domains. A ribbon diagram of $tmTF_N$ is shown for reference (left) and electrostatic surface potentials are shown for $tmTF_N$ (middle) and for $ecTF_N$ (right). (*b*) Orthogonal view of (*a*). The ribosome-binding loop is marked by a cyan arrow and characterized by a positively charged surface (blue). The overall TF N-terminal domain structure appears positively charged (red), with few exposed hydrophobic surfaces or patches (light grey). (*c*) TF C-terminal domains. A ribbon diagram of $tmTF_C$ is shown for reference (left) and electrostatic surface potentials are shown for $tmTF_C$ (middle) and for $ecTF_C$ (right). This view highlights the concave surface in the C-terminal domain. The mostly negatively charged surface (red) is interspersed with small positive (blue) and hydrophobic (light grey) patches. Some positive patches, identified by green arrows, appear to be conserved.

interactions *via* TF hydrophilic surfaces. In fact, the domain interactions between $vcTF_N$ and $vcTF_C$ seen in the C-terminally truncated *V. cholerae* trigger-factor structure appear to involve a misfolded or partially folded TF C-terminal domain (Ludlam *et al.*, 2004), possibly reflecting interactions between TF and a folding substrate.

We thank M. A. Gawinowicz for mass spectrometry, C. J. Lusty and M. Floer for discussions and M. Collins, Q. Fan, G. Gregorio, J. Lidestri, Q. Liu, A. Marina, C. Min, J. Moore, L. You, H. Xie, R. Xu and Z. Zhang for help. This work was supported in part by NIH grant GM34102; EM-H was a fellow of the Leukemia and Lymphoma Society of America and the Jane Elissa/Charlotte Meyers Endowment Fund. Beamline X4A at the National Synchrotron Light Source, a Department of Energy facility, is supported by the New York Structural Biology Center.

References

Baram, D., Pyetan, E., Sittner, A., Auerbach-Nevo, T., Bashan, A. & Yonath, A. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 12017–12022.
Bitto, E. & McKay, D. B. (2002). *Structure*, **10**, 1489–1498.
Blaha, G., Wilson, D. N., Stoller, G., Fischer, G., Willumeit, R. & Nierhaus, K. H. (2003). *J. Mol. Biol.* **326**, 887–897.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Callebaut, I. & Mornon, J. P. (1995). *FEBS Lett.* **374**, 211–215.
Craig, E. A., Weissman, J. S. & Horwich, A. L. (1994). *Cell*, **78**, 365–372.
Crooke, E., Guthrie, B., Lecker, S., Lill, R. & Wickner, W. (1988). *Cell*, **54**, 1003–1011.
Crooke, E. & Wickner, W. (1987). *Proc. Natl Acad. Sci. USA*, **84**, 5216–5220.
Deuerling, E., Schulze-Specking, A., Tomoyasu, T., Mogk, A. & Bukau, B. (1999). *Nature (London)*, **400**, 693–696.
Doublié, S. (1997). *Methods Enzymol.* **276**, 523–530.
Ferbitz, L., Maier, T., Patzelt, H., Bukau, B., Deuerling, E. & Ban, N. (2004). *Nature (London)*, **431**, 590–596.
Flynn, G. C., Pohl, J., Flocco, M. T. & Rothman, J. E. (1991). *Nature (London)*, **353**, 726–730.
Genevaux, P., Keppel, F., Schwager, F., Langendijk-Genevaux, P. S., Hartl, F. U. & Georgopoulos, C. (2004). *EMBO Rep.* **5**, 195–200.
Graumann, J., Lilie, H., Tang, X., Tucker, K. A., Hoffmann, J. H., Vijayalakshmi, J., Saper, M., Bardwell, J. C. & Jakob, U. (2001). *Structure*, **9**, 377–387.
Hendrickson, W. A., Horton, J. R. & Lemaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.
Hesterkamp, T. & Bukau, B. (1996). *FEBS Lett.* **385**, 67–71.
Hesterkamp, T., Deuerling, E. & Bukau, B. (1997). *J. Biol. Chem.* **272**, 21865–21871.
Hesterkamp, T., Hauser, S., Lutcke, H. & Bukau, B. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 4437–4441.
Holm, L. & Sander, C. (1994). *Nucleic Acids Res.* **22**, 3600–3609.
Jones, S. & Thornton, J. M. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 13–20.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kabsch, W. (1976). *Acta Cryst.* A**32**, 922–923.
Kim, K. K., Kim, R. & Kim, S.-H. (1998). *Nature (London)*, **394**, 595–599.
Kim, R., Kim, K. K., Yokota, H. & Kim, S.-H. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 9129–9133.
Kim, S. J., Jeong, D. G., Chi, S. W., Lee, J. S. & Ryu, S. E. (2001). *Nature Struct. Biol.* **8**, 459–466.
Kramer, G., Patzelt, H., Rauch, T., Kurz, T. A., Vorderwulbecke, S., Bukau, B. & Deuerling, E. (2004). *J. Biol. Chem.* **279**, 14165–14170.
Kramer, G., Rauch, T., Rist, W., Vorderwulbecke, S., Patzelt, H., Schulze-Specking, A., Ban, N., Deuerling, E. & Bukau, B. (2002). *Nature (London)*, **419**, 171–174.
Kramer, G., Rutkowska, A., Wegrzyn, R. D., Patzelt, H., Kurz, T. A., Merz, F., Rauch, T., Vorderwulbecke, S., Deuerling, E. & Bukau, B. (2004). *J. Bacteriol.* **186**, 3777–3784.
Kristensen, O. & Gajhede, M. (2003). *Structure*, **11**, 1547–1556.
Lecker, S., Lill, R., Ziegelhoffer, T., Georgopoulos, C., Bassford, P. J. Jr, Kumamoto, C. A. & Wickner, W. (1989). *EMBO J.* **8**, 2703–2709.
Lill, R., Crooke, E., Guthrie, B. & Wickner, W. (1988). *Cell*, **54**, 1013–1018.
Liu, C. P., Perrett, S. & Zhou, J. M. (2005). *J. Biol. Chem.* **280**, 13315–13320.
Liu, C. P. & Zhou, J. M. (2004). *Biochem. Biophys. Res. Commun.* **313**, 509–515.
Liu, Y. & Eisenberg, D. (2002). *Protein Sci.* **11**, 1285–1299.
Lovell, S. C., Davis, I. W., Arendall, W. B. III, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S. & Richardson, D. C. (2003). *Proteins*, **50**, 437–450.
Lu, G. (2000). *J. Appl. Cryst.* **33**, 176–183.
Ludlam, A. V., Moore, B. A. & Xu, Z. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 13436–13441.
Maier, R., Eckert, B., Scholz, C., Lilie, H. & Schmid, F. X. (2003). *J. Mol. Biol.* **326**, 585–592.
Martin-Benito, J., Boskovic, J., Gomez-Puertas, P., Carrascosa, J. L., Simons, C. T., Lewis, S. A., Bartolini, F., Cowan, N. J. & Valpuesta, J. M. (2002). *EMBO J.* **21**, 6377–6386.
Merz, F., Hoffmann, A., Rutkowska, A., Zachmann-Brand, B., Bukau, B. & Deuerling, E. (2006). *J. Biol. Chem.* **281**, 31963–31971.
Minor, W., Cymborowski, M. & Otwinowski, Z. (2002). *Acta Phys. Pol. A*, **101**, 613–619.
Nicholls, A., Bharadwaj, R. & Honig, B. (1993). *Biophys. J.* **64**, A166.
Notredame, C., Higgins, D. G. & Heringa, J. (2000). *J. Mol. Biol.* **302**, 205–217.
Patzelt, H., Kramer, G., Rauch, T., Schonfeld, H. J., Bukau, B. & Deuerling, E. (2002). *Biol. Chem.* **383**, 1611–1619.
Patzelt, H., Rudiger, S., Brehmer, D., Kramer, G., Vorderwulbecke, S., Schaffitzel, E., Waitz, A., Hesterkamp, T., Dong, L., Schneider-Mergener, J., Bukau, B. & Deuerling, E. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 14244–14249.
Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Read, R. J. (2001). *Acta Cryst.* D**57**, 1373–1382.
Schlunzen, F., Wilson, D. N., Tian, P., Harms, J. M., McInnes, S. J., Hansen, H. A., Albrecht, R., Buerger, J., Wilbanks, S. M. & Fucini, P. (2005). *Structure*, **13**, 1685–1694.
Schneider, T. R. (2002). *Acta Cryst.* D**58**, 195–208.
Scholz, C., Mucke, M., Rape, M., Pecht, A., Pahl, A., Bang, H. & Schmid, F. X. (1998). *J. Mol. Biol.* **277**, 723–732.
Schulze-Gahmen, U., Aono, S., Chen, S., Yokota, H., Kim, R. & Kim, S.-H. (2005). *Acta Cryst.* D**61**, 1343–1347.
Siegert, R., Leroux, M. R., Scheufler, C., Hartl, F. U. & Moarefi, I. (2000). *Cell*, **103**, 621–632.
Steinbacher, S. (1999). *Nature Struct. Biol.* **6**, 1029–1032.
Stoller, G., Rucknagel, K. P., Nierhaus, K. H., Schmid, F. X., Fischer, G. & Rahfeld, J. U. (1995). *EMBO J.* **14**, 4939–4948.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.

Teter, S. A., Houry, W. A., Ang, D., Tradler, T., Rockabrand, D., Fischer, G., Blum, P., Georgopoulos, C. & Hartl, F. U. (1999). *Cell*, **97**, 755–765.
Ullers, R. S., Houben, E. N., Raine, A., ten Hagen-Jongman, C. M., Ehrenberg, M., Brunner, J., Oudega, B., Harms, N. & Luirink, J. (2003). *J. Cell Biol.* **161**, 679–684.
Valent, Q. A., Kendall, D. A., High, S., Kusters, R., Oudega, B. & Luirink, J. (1995). *EMBO J.* **14**, 5494–5505.
Vijayalakshmi, J., Mukhergee, M. K., Graumann, J., Jakob, U. & Saper, M. A. (2001). *Structure*, **9**, 367–375.
Vogtherr, M., Jacobs, D. M., Parac, T. N., Maurer, M., Pahl, A., Saxena, K., Ruterjans, H., Griesinger, C. & Fiebig, K. M. (2002). *J. Mol. Biol.* **318**, 1097–1115.
Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). *Acta Cryst.* D**57**, 122–133.
Zhu, X., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E. & Hendrickson, W. A. (1996). *Science*, **272**, 1606–1614.