Solution structure of the C1-subdomain of Bacillus stearothermophilus translation initiation factor IF2

Hans Wienk, Simona Tomaselli, Cédric Bernard, Roberto Spurio, Delia Picone, Claudio O. Gualerzi and Rolf Boelens

Protein Sci. 2005 14: 2461-2468; originally published online Aug 4, 2005;
doi:10.1110/ps.051531305

Solution structure of the C1-subdomain of *Bacillus stearothermophilus* translation initiation factor IF2

HANS WIENK,[1,4] SIMONA TOMASELLI,[1,3,4] CÉDRIC BERNARD,[1,5] ROBERTO SPURIO,[2] DELIA PICONE,[3] CLAUDIO O. GUALERZI,[2] AND ROLF BOELENS[1]

[1]Bijvoet Center for Biomolecular Research, Department of NMR Spectroscopy, Utrecht University, 3584 CH Utrecht, The Netherlands
[2]Laboratory of Genetics, Department of Biology MCA, University of Camerino, 62032 Camerino (MC), Italy
[3]Department of Chemistry, University of Napoli "Federico II," 80126 Napoli, Italy

(RECEIVED April 21, 2005; FINAL REVISION June 4, 2005; ACCEPTED June 9, 2005)

Abstract

IF2 is one of three bacterial translation initiation factors that are conserved through all kingdoms of life. It binds the 30S and 50S ribosomal subunits, as well as fMet-tRNA$_f^{Met}$. After these interactions, fMet-tRNA$_f^{Met}$ is oriented to the ribosomal P-site where the first amino acid of the nascent polypeptide, formylmethionine, is presented. The C-terminal domain of *Bacillus stearothermophilus* IF2, which is responsible for recognition and binding of fMet-tRNA$_f^{Met}$, contains two structured modules. Previously, the solution structure of the most C-terminal module, IF2-C2, has been elucidated by NMR spectroscopy and direct interactions between this subdomain and fMet-tRNA$_f^{Met}$ were reported. In the present NMR study we have obtained the spectral assignment of the other module of the C-terminal domain (IF2-C1) and determined its solution structure and backbone dynamics. The IF2-C1 core forms a flattened fold consisting of a central four-stranded parallel β-sheet flanked by three α-helices. Although its overall organization resembles that of subdomain III of the archaeal IF2-homolog eIF5B whose crystal structure had previously been reported, some differences of potential functional significance are evident.

Keywords: protein synthesis; translation initiation; IF2; fMet-tRNA; NMR structure

Initiation of protein biosynthesis in bacteria requires the three protein factors IF1, IF2, and IF3. These factors are crucial for the assembly of the ribosomal initiation complex and act on the kinetics of translation by speeding up several processes driving codon–anti-codon recognition. In the final steps of translation initiation a 70S ribosomal complex is formed that is composed of the 30S and 50S ribosomal subunits, GTP-associated IF2, mRNA, and fMet-tRNA$_f^{Met}$, the initiator tRNA unique for bacteria. IF2 plays a central role in the assembly of this complex by binding to fMet-tRNA$_f^{Met}$ and the 30S ribosomal subunit before recruiting the 50S subunit (for recent reviews, see Gualerzi et al. 2001 and Boelens and Gualerzi 2002). Thus, IF2 triggers the ribosomal initiation complex formation, brings the first amino acid of the polypeptide chain to the ribosome in the form of fMet-tRNA$_f^{Met}$, and facilitates the first transpeptidation reaction by positioning fMet-tRNA$_f^{Met}$ in the ribosomal P-site.

A current trend in the investigation of translation initiation is to obtain structures with the maximum resolution possible, to attribute distinct functions to individual re-

Reprint requests to: Rolf Boelens, Bijvoet Center for Biomolecular Research, Department of NMR Spectroscopy, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands; e-mail: r.boelens@chem.uu.nl; fax: +31-30-2537623.

[4]These authors contributed equally to the experiments reported in this work.

[5]Present address: Architecture et Fonction des Macromolécules Biologiques, Unité Mixte de Recherche (UMR) 6098, Centre National de la Recherche Scientifique, 31 chemin Joseph Aiguier, 13402 Marseille cedex 20, France.

Article published online ahead of print. Article and publication date are at http://www.proteinscience.org/cgi/doi/10.1110/ps.051531305.

gions of initiation factors and to correlate these functions with the elucidated structures. For what concerns IF2, the most detailed structure information available derives from the domain organizations of the free, GDP-bound, and GTP-bound forms of eIF5B from *Methanococcus thermoautotrophicum* (*Mth*) (Roll-Mecak et al. 2000). However, although it is the closest amino acid sequence homolog with structural information, this archaeal factor shares only 27% amino acid identity and 47% homology with bacterial IF2 from *Bacillus stearothermophilus* (*Bst*). In addition, although a wealth of biochemical and genetic data is available for ribosomes and translation initiation in bacteria, much less is known concerning the archaeal system. The biochemical data that are available for archaeal translation initiation indicate that the functional properties of eIF5B are significantly different from those of IF2 (Gualerzi et al. 2001; Boelens and Gualerzi 2002). Most importantly, direct binding of initiator tRNA, which represents a key feature of bacterial IF2, has never been demonstrated for eIF5B, and unlike that of bacterial IF2, the exact role of eIF5B in protein biosynthesis remains unclear.

As mentioned, prokaryotic initiation factors have been extensively characterized by genetic and biochemical experiments. In the case of the 82 kDa *Bst* IF2, analysis of limited proteolysis allowed the identification of three defined fragments. These fragments were shown to possess not only distinct physical identities but also specific and characteristic functional properties, and were denominated N-domain (N-terminal), G-domain (GTPase), and C-domain (C-terminal; Gualerzi et al. 1991). Subsequent structure and function analyses have demonstrated that each of these domains contains defined structured modules, some of which are endowed with a clearly identifiable function. Throughout this work we shall continue to refer to the large proteolytic IF2 fragments as "domains" and to the smaller structural building blocks as "subdomains." The N-terminal domain is not conserved in length and sequence among bacterial IF2 molecules and its function is still not completely clear. The 50-residue long N-terminal subdomain of the N-domain of *Escherichia coli* IF2 was shown to form a folded module that is connected to the rest of the molecule via a very long flexible linker (Laursen et al. 2004). Although in vitro or in vivo deletion of the whole N-domain did not cause complete loss of function, it has been suggested that this domain enhances interactions between IF2 and both the 30S and 50S ribosomal subunits (Moreno et al. 1999).

The G-domain (amino acids 155–519 in *Bst* IF2) contains a guanosine nucleotide-binding motif homologous to those found in small GTPases (Vachon et al. 1990), which is probably responsible for both the GTPase activity of IF2 (Spurio et al. 1993) and the binding to the 50S ribosomal subunit (Marzi et al. 2003). C-terminally of the GTPase core is a β-barrel subdomain, involved in the functional interaction of IF2 with the 30S ribosomal subunit (Marzi et al. 2003; E. Caserta, J. Tomsic, R. Spurio, C.L. Pon, and C.O. Gualerzi, in prep.), and at the N-terminal end of the GTPase subdomain a polypeptide segment of about 50 residues is found, having unknown structure and function.

The C-terminal domain is composed of two subdomains of similar size, IF2-C1 and IF2-C2 (Misselwitz et al. 1997). It appears that all the structural determinants necessary for the binding of fMet-tRNA$_f^{Met}$ are embedded in the IF2-C2 subdomain, which is in accordance with the fact that the *Bst* IF2-C2–fMet-tRNA$_f^{Met}$ complex has a stability comparable to the complexes involving native IF2 or the entire IF2 C-domain (Krafft et al. 2000). The three-dimensional solution structure of *Bst* IF2-C2 has been solved by NMR spectroscopy (Meunier et al. 2000), and was shown to form a β-barrel structure with homology to the eIF5B subdomains II and IV (numbering as in Roll-Mecak et al. 2000). Furthermore, it has been shown that this subdomain recognizes directly the formylated α-NH_2-group of fMet-tRNA$_f^{Met}$ (Guenneugues et al. 2000).

So far, structure and function of the bacterial IF2-C1 subdomain are completely unknown. Subdomain III of *Mth* eIF5B, with amino acid sequence homology to the C1-subdomain of IF2, was suggested to play a role in transmitting and amplifying small structural changes occurring in the G-domain upon nucleotide binding to eIF5B subdomain IV (Roll-Mecak et al. 2000). However, whether or not this particular function of the archaeal initiation factor is preserved in the bacterial IF2 critically depends on the strict conservation of characteristic structure-related properties, such as the presence and restricted motional freedom of a long α-helix connecting the C1-subdomain with the C2-subdomain that would function as a rigid lever. To be able to focus on such structure-related properties, in the present study we have studied the solution structure of the *Bst* IF2-C1 subdomain. Our results show that although the overall fold of the bacterial subdomain displays a marked resemblance to the structure of the homologous eIF5B subdomain III, distinct differences exist between the two that could have important functional consequences.

Results

IF2-C1 resonance assignments

Although virtually all signal intensity in [^{1}H,^{15}N]-HSQC spectra could be assigned unambiguously (Fig. 1), the *Bst* IF2-C1 backbone amide assignment was only 76%

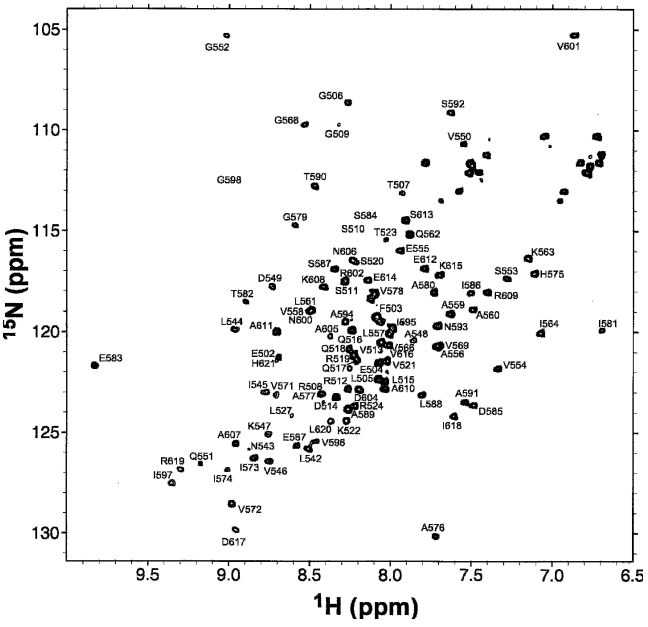

Figure 1. Assignment of *Bst* IF2-C1 backbone amide protons in 700-MHz 2D [^{1}H,^{15}N]-HSQC experiment.

complete. Besides the signals of residues N501, D565, V566, and F599, those of two longer amino acid stretches were not found, namely V525–E541 and R622–M635, with the exception of a very low intensity signal of L527, recognized by signals belonging to S526 in the CBCA(CO)NH spectrum. For these two regions signal intensity was probably lost as a result of line broadening, possibly due to protein aggregation or slow conformational exchange. Line broadening is also implied by visible differences in signal intensity, which is rather low for instance for G598 (9.1 × 112.5 ppm) and S584 (8.4 × 114.7 ppm)—in fact, no intensity is seen at the lowest contour level used to create Figure 1—but much higher for most other residues. From the assigned amino acids, 92% of the nonexchangeable protons were uniquely defined, as well as 95% of all carbon atoms. The final assignment includes the H^{α}, C^{α}, H^{β}, C^{β}, and C' chemical shifts for N501; the H^{α}, C^{α}, H^{β}, and C^{β} chemical shifts for S526; carbonyl, and complete side-chain information for E541; and the C^{α} and C^{β} frequencies for D565. The ^{1}H-, ^{15}N-, and ^{13}C-chemical shifts have been deposited in the BioMagResBank database under accession code 6577. Based on the chemical shift values, the programs CSI and TALOS identified the positions of secondary structure elements in IF2-C1 to coincide with the findings for the eIF5B subdomain III (vide infra). For the unique proline, P603, the C^{β} (31.7 ppm) and C^{γ} (27.7 ppm) chemical shift values indicate a *trans*-conformation (Schubert et al. 2002). This was confirmed by NOE analysis, which showed the absence of H^{α}–H^{α} contacts between R602 and P603 and the presence of R602-H^{α} × P603-H^{δ} NOEs.

Three-dimensional structure of IF2-C1

Input data for structure calculations and quality parameters for the final structure ensemble of IF2-C1 are summarized in Table 1. One thousand four hundred twenty-nine upper limit distance restraints as well as 95 TALOS restraints, 101 CSI restraints, and 53 $^{3}J_{HNH\alpha}$ coupling restraints were used to calculate the *Bst* IF2-C1 structure ensemble that is deposited in the Protein Data Bank (entry 1Z9B). The final structure ensemble shows no dihedral angle violations over 5°, and no distance violations over 0.5 Å. As anticipated, the lack of backbone assignments yielded disordered tails comprising

Table 1. *Statistics for the final IF2-C1 structure ensemble calculation*

		Assigned	Total
Manual peak assignment			
2D-NOESY		1165	2963
NOESY-[^{15}N]-HSQC		440	976
NOESY-[^{13}C]-HSQC		1043	2554
	Total	2648	6493
Unambiguous restraints (final.upl)			
Intraresidual		389	
Sequential		377	
Medium range[a]		273	
Long range		390	
Other restraints			
TALOS dihedral angles		95	
CSI dihedral angles		101	
$^{3}J_{HNH\alpha}$ couplings (ϕ-angles)		53	
Violations			
Distances >0.5 Å		0	
Torsion angles >5°		0	
RMS deviation from mean (Å)			
All core residues[b]			
All heavy atoms		1.24	
Backbone atoms (N, C^{α}, C')		0.75	
Secondary structure elements[c]			
All heavy atoms		0.97	
Backbone atoms(N, C^{α}, C')		0.50	
Ramachandran plot[b]			
Most favored region (%)		87.8	
Additionally allowed region (%)		9.5	
Generously allowed region (%)		1.2	
Disallowed region (%)		1.5	
WHAT CHECK structure Z-scores[b]			
Ramachandran plot appearance		–3.6	
2nd generation packing quality		–1.6	
χ^{1}–χ^{2} rotamer normality		–2.4	
Backbone conformation		–2.3	

[a] $1 < |i-j| < 5$.
[b] Residues E541–H621.
[c] Residues E541–A548, Q551–L561, R570–V578, E583–S592, I595–G598, A605–S613, I618–L620.

the N-terminal 40 residues and the C-terminal tail following residue H621 (not shown). In contrast, the IF2-C1 core region (residues E541 through H621) is structured and used for further analysis (Fig. 2). Overall, the ensemble adopts a flattened disk-shape, with a central sheet consisting of four parallel strands, surrounded by three α-helices (Fig. 3A). The consensus secondary structure elements that are found for *Bst* IF2-C1 are as follows: S20, E541–A548; H9, Q551–L561; S21, R570–V578; H10, E583–S592; S22, I595–G598; H11, A605–S613; and S23, I618–L620 (S, β-strand; H, α-helix; secondary structure element numbering follows Roll-Mecak et al. 2000). Based on amino acid sequence alignments and the crystal structure of free *Mth* eIF5B (PDB entry 1G7R), for the *Bst* IF2-C1 core the secondary structure arrangement was predicted to be S20–H9–S21–H10–S22–H11–S23 for residues L544–K547, G552–K563, H575–V578, E583–S592, A594–N600, A605–A610, and D617–H621, respectively. Clearly, all α-helices and β-strands in eIF5B subdomain III are conserved in the calculated structures of the ensemble, albeit with small differences in length and somewhat shifted positions (Fig. 3B). Using residues E541–H621, RMSD values for the ensemble were 0.75 Å for the backbone N-, C^{α}-, and C′ atoms and 1.24 Å for all heavy atoms. The regions with regular secondary structure showed RMSD values of 0.50 Å for the backbone atoms and of 0.97 Å for all heavy atoms. Overall, the RMSD could be improved by manual selection of twenty most similar structures out of an ensemble of 30 lowest energy structures (not shown).

The IF2-C1 core shows good Ramachandran statistics (cf. Table 1). Only for the residues between D565 and G568, and for N593, N600, and R602, for which few experimental restraints were found and which are located in regions with considerable internal dynamics, in some of the structures ϕ- and ψ-angle outliers are found.

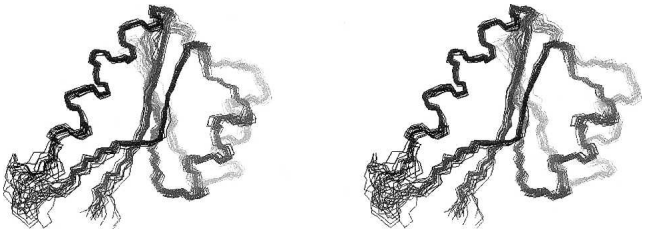

Figure 2. Stereo view for the backbone of the solution structure ensemble calculated for the *Bst* IF2-C1 core (PDB entry 1Z9B).

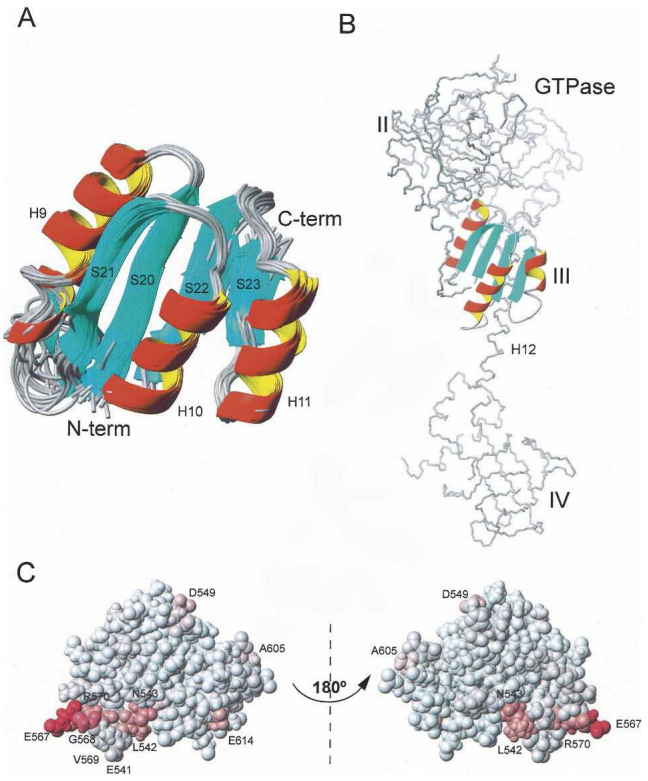

Figure 3. Structure information for the *Bst* IF2-C1 core structure. (*A*) Ribbon representation for the NMR ensemble of *Bst* IF2-C1 with secondary structure element numbering as for *Mth* eIF5B (Roll-Mecak et al. 2000). (*B*) Ribbon representation of subdomain III in the *Mth* eIF5B crystal structure (PDB entry 1G7R). Orientations of the IF2-C1 subdomain and the eIF5B subdomain III are manually aligned. (*C*) Observed slow conformational exchange R_{ex} projected on a space-filling model for the core of the lowest energy structure. The *left* panel is aligned with Figure 2. Although data were obtained for backbone amides, for clarity whole amino acids are color coded.

Internal dynamics of IF2-C1

In Figure 4, A and B, the heteronuclear ^{15}N relaxation data are summarized. The unstructured N-terminal region of the IF2-C1 construct shows fast internal dynamics, as indicated by reduced $^{15}N\{^{1}H\}$-heteronuclear NOE values (Fig. 4A) and the reduced R2/R1 ratios (Fig. 4B) for the residues that were assigned. Throughout the folded IF2-C1 core the heteronuclear NOEs are distributed over a small range, roughly between 0.6 and 0.8, which is in agreement with a relatively rigid structure devoid of motions on the nanosecond–picosecond (nsec–psec) timescale. The R2/R1 ratios for the IF2-C1 core are less homogeneous than the heteronuclear NOEs. Increases in R2/R1 may result from slow conformational exchange within the backbone. Alternatively, for each residue the R2/R1 ratio may be affected differently by the anisotropy of the overall rotational diffusion induced by flattening of the molecule.

To dissect the overall tumbling anisotropy properties of the IF2-C1 subdomain from its internal dynamics a modelfree analysis was performed with FAST-Model-

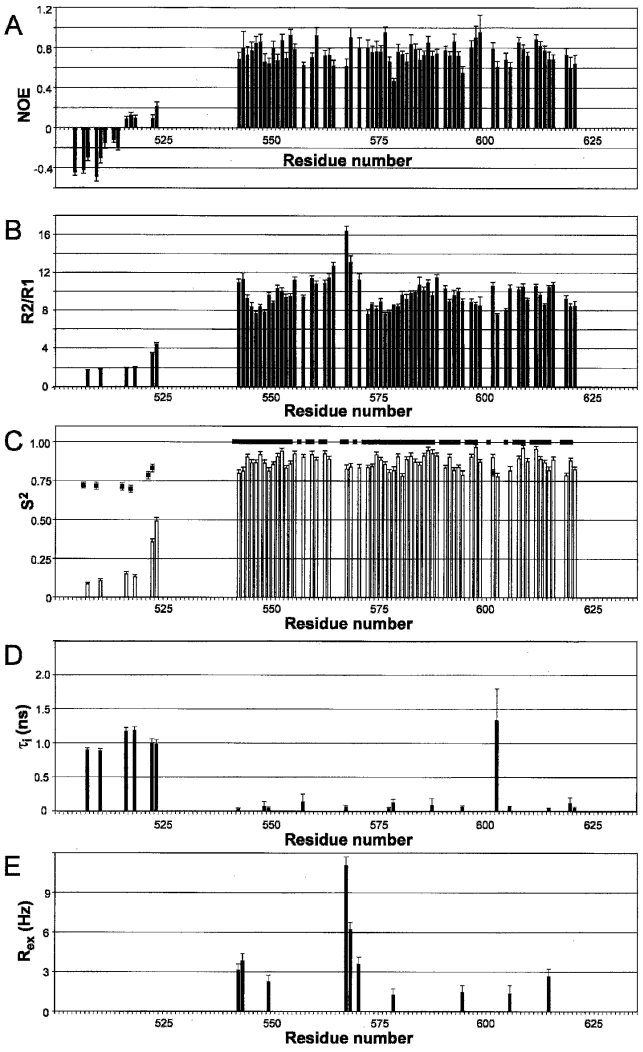

Figure 4. Parameters describing internal dynamics for each residue of *Bst* IF2-C1. (*A*) $^{15}N\{^1H\}$-heteronuclear NOE values; (*B*) ^{15}N-R2/R1 ratios; (*C*) squared general order parameters in the slow (bars) and the fast timescales (filled squares); (*D*) fast (nsec–psec) internal motions expressed as internal correlation times τ_i (nsec); (*E*) slow (msec–μsec) conformational exchange expressed as excess transverse relaxation rates R_{ex}.

free (Cole and Loria 2003). For all C1-core amino acids except for the loop-residue I564 and D604 and some of the residues constituting the N-terminal linker dynamics could be described using one of the five standard combinations of contributions to internal motions [i.e., (1) squared generalized order parameter S^2, (2) S^2 and internal correlation time τ_i, (3) S^2 and slow conformational exchange term R_{ex}, (4) S^2 and τ_i and R_{ex}, and (5) S^2 and τ_i and extra order parameter for fast internal motions S^2_f]. Concerning the overall tumbling characteristics, the total correlation time ($\tau_c = 9.02 \pm 0.04$ nsec) and the anisotropy of the axial symmetric rotational diffusion tensor ($D^{//}/D^{\perp} = 1.37 \pm 0.04$) are consistent with the molecular weight and the nonspherical shape found for the IF2-C1 core structure, although an aggregated state cannot be excluded. As was indicated by the heteronuclear NOEs and R2/R1 ratios, the residues in the extreme N terminus exhibit complex fast internal dynamics on the nanosecond timescale, in agreement with this region being unstructured (Fig. 4). Concerning the internal motion behavior of the IF2-C1 core, the squared general order parameters all have values between and 0.78 and 0.97 with an average $S^2 = 0.87 \pm 0.05$ (Fig. 4C, filled bars). These confirm an overall rigid fold lacking internal dynamics on the nsec–psec timescale, which is further supported by the fact that only for few residues model-free parameters are required that consider fast internal dynamics. For the loop-residue R602 complex, fast internal dynamics is detected with an internal correlation time τ_i above 1 nsec as well as a term considering an additional general order parameter for fast internal motions ($\tau_i = 1330 \pm 467$ psec; $S^2_f = 0.80 \pm 0.02$). Furthermore, although fast internal vibrations could also be present for L542, A548, D549, L557, E567, A577, V578, S587, A594, A605, E614, R619, and L620, the obtained τ_i values and errors suggest that most of them are too small to be analyzed accurately. More interesting internal dynamics appears to occur on the msec–μsec timescale, as indicated by slow conformational exchange contributions (R_{ex}) over 3 Hz shown (1) for residues E567, G568, and R570 belonging to the loop connecting α-helix H9 and β-strand S21 and (2) residues L542 and N543 in the N-terminal strand S20 (Fig. 4E). In Figure 3C, it is shown that these residues cluster in the folded C1-subdomain. Also, E541 and V569 could take part in this region; however, for E541 no backbone ^{15}N assignment was found, and for V569 no reliable relaxation data was obtained because of overlap with A556. Other residues for which smaller msec–μsec motions are indicated are D549, V578, A594, A605, and E614, but these may be of no significance.

Discussion

The overall fold of *Bst* IF2-C1 presented in Figure 3A shows that two helices (i.e., H11 and H10) flank the central β-sheet on one side, whereas the third helix (H9) contacts the β-sheet on the opposite side. From the comparison of the three-dimensional structure of the IF2-C1 core (Fig. 3A) with that of subdomain III of *Mth* eIF5B (Fig. 3B) it is evident that, in addition to the established amino acid homology, also the three-dimensional arrangement for the secondary structure elements of the two protein modules are closely related. Indeed, using the DALI Web server (http://www.ebi.ac.uk/dali),

we found that the closest structural homolog to the lowest energy NMR structure for the IF2-C1 core (residues E541–H621) was by far this eIF5B region, for which a Z-score (i.e., the compatibility of a sequence with a given fold) of 9.8, and a pairwise positional RMSD for C^{α}-atoms of 2.6 Å were retrieved. For *Mth* eIF5B the subdomain III structure has been presented as a novel fold (Roll-Mecak et al. 2000); therefore, the finding of the homologous structure for the *Bst* IF2-C1 core in the current study doubles the number of family members for this fold and suggests that conservation of this domain is an essential feature for the functionality of bacterial IF2s as well as their nonbacterial homologs.

In spite of this structure homology, some clear-cut differences between the two structures are evident. The most remarkable difference between the two homologous subdomains concerns the number of α-helices flanking the central β-sheet: Whereas there are only three in *Bst* IF2-C1 (Fig. 3A), there are four (two on each side) in *Mth* eIF5B. The region corresponding to helix H12 (Fig. 3B) was not assigned in the bacterial protein. As seen in Figure 3B, the overall structure of eIF5B resembles a pendulum with helix H12 forming a stiff rod connecting subdomain III to subdomain IV, the IF2-C2 subdomain homolog that represents the bob. In the *Mth* eIF5B domain organization, helix H12 is folded back over the dorsal face of the central β-sheet of subdomain III, and could be fixed by mutual direct intramodular contacts, thus contributing to the rigidity assumed to be important for the reorientation of subdomain IV. The section corresponding to helix H12 in *Bst* IF2-C1, which would extend from Y625 through E645, is truncated at M635 in the construct and is undefined in the final structure ensemble due to lacking signal intensity. We have strong indications that this trimming is the cause of the absence of helix H12, because a construct of the complete IF2 C-domain does show helical propensity in this region (H. Wienk, unpubl.).

Another difference between the two homologous subdomains concerns the central β-sheet. It appears that S22 and S23 are shorter in *Bst* IF2-C1 than in the homologous subdomain III of *Mth* eIF5B. This may be related to the previous discussion about helix H12, since in eIF5B this helix is amphipathic in the region covering the β-sheet, with charged residues exposed to the solvent, while hydrophobic and aromatic side chains are in direct contact with hydrophobic residues of the central β-sheet. These properties of helix H12 are not evident in the *Bst* IF2 amino acid sequence, suggesting that the amphipathic character of this helix is less important for IF2 functioning. The absence of helix H12 in *Bst* IF2 would expose hydrophobic residues present in β-strands S22 and S23 to the solvent, and to relieve this energetically unfavorable state their side chains may be somewhat reoriented. As a result, this could prohibit the formation of typical β-strand dihedral angles, at the same time maintaining the C1-subdomain integrity as indicated by the relaxation data and the overall fold similarity with eIF5B subdomain III.

Finally, the linker connecting the C1-subdomain N-terminally to the G-domain appears to be completely unstructured and seems at least partly subject to fast internal motions. In addition, in the *Bst* IF2-C1 fold the two loops supporting this linker have different length compared to the homologous *Mth* eIF5B subdomain III structure; the one between H9 and S21 is longer in IF2-C1, while the reverse is seen for the loop between H10 and S22. Clearly, in IF2-C1 the former loop displays structural divergence (Figs. 2, 3A), which indicates that this region contains internal mobility. This agrees with our relaxation analysis implying slow conformational exchange for this loop that could be transferred directly to the linker to the GTPase domain (Fig. 3B). The overall reduction in structural support for this linker between the G-domain and the C-domain could be biologically relevant, since it would cast additional doubts on the general validity of the "signal transmission" model suggested for *Mth* eIF5B, which is almost exclusively based on protein rigidity.

Materials and methods

Sample preparation and NMR conditions

Details of the cDNA cloning and overexpression of IF2-constructs have been described previously (Spurio et al. 1993, 2000). The IF2-C1 protein was heterologously expressed from a pEV-expression vector in *Escherichia coli* strain JM109 that harbors the plasmid pCI to render protein overproduction temperature-inducible. The IF2-C1 clone encodes a 135-residue protein consisting of a 15-residue N-terminal linker, from which the N-terminal formyl-methionine is removed after expression, followed by the IF2 wild-type C1 region Leu515–Met635. For stable isotope-labeling bacteria were grown at 30°C in M9-based minimal medium containing either 0.8 g/L $^{15}NH_4Cl$ and 5 g/L unlabeled glucose, or 0.8 g/L $^{15}NH_4Cl$ and 3 g/L [^{13}C]-glucose. After overexpression for 1 h at 42°C, followed by harvesting and sonication, IF2-C1 was isolated from the lysate at 4°C in a procedure with three chromatography steps: (1) a CM Sepharose column (Amersham) equilibrated with 20 mM Tris-HCl (pH 7.1), 5% glycerol, 1 mM EDTA, 20 mM ammonium chloride, and 5 mM β-mercaptoethanol; (2) a DEAE-Sepharose column (Amersham) in the same buffer, eluted with a linear NH_4Cl gradient from 20 mM to 400 mM; and (3) a Sephadex G-75 gel filtration column (Amersham) equilibrated in 20 mM Tris-HCl (pH 7.1), 1 mM EDTA, 5% glycerol, 150 mM NH_4Cl, 5 mM β-mercaptoethanol, and 0.2 mM phenylmethylsulfonyl fluoride.

[^{15}N]-labeled or [^{15}N,^{13}C]-labeled IF2-C1 NMR-samples (0.5 mM) were prepared in a buffer containing 20 mM KPi (pH 6.4),

200 mM KCl, 5% glycerol-*d*8 and 5% D_2O. NMR spectra were acquired at 34°C on Bruker Avance spectrometers equipped with 5-mm triple resonance probes, were processed using NMRPipe (Delaglio et al. 1995) and were analyzed with NMRView5.0.3 (Johnson and Blevins 1994). Protein structures are visualized using MOLMOL (Koradi et al. 1996).

Spectral assignments

Sequence-specific resonance assignments were performed using three-dimensional HNCA, HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO, and HBHA(CO)NH triple resonance experiments in combination with a 3D NOESY-[^{1}H,^{15}N]-HSQC at 700 MHz. Subsequently, side-chain carbons were assigned from a 700-MHz (H)CCH-TOCSY spectrum while additional side-chain protons were assigned using a 700-MHz H(C)CH-TOCSY spectrum.

NOE-based distance restraints were extracted from 900-MHz 2D NOE spectra, the 700-MHz 3D NOESY-[^{1}H,^{15}N]-HSQC spectrum, and a 900-MHz 3D NOESY-[^{1}H,^{13}C]-HSQC spectrum. Over 40% of the total NOESY signals was assigned manually (Table 1). Dihedral angle constraints were derived from chemical shift values using both TALOS (Cornilescu et al. 1999), giving ϕ- and ψ-angles with an uncertainty to which an extra uncertainty of $\pm 20°$ was added, and the chemical shift index output from the program CSI (Wishart and Sykes 1994) with an error of $\pm 40°$. Additional ϕ-angle restraints were obtained from $^3J_{HNH\alpha}$ coupling constants extracted from a 700-MHz quantitative *J*-correlated HNHA experiment. All restraints used to calculate the final structure ensemble are added to BMRB accession number 6577.

Structure calculations

Initial structure calculations involved eight iterations on an extended input structure, using CNS protocols in ARIA version 1.2, in which each iteration consisted of an automatic NOE assignment followed by a simulated annealing procedure (Linge et al. 2001). Besides the similarity in secondary structure positions as implied by CSI and TALOS analyses, for the IF2-C1 core a fold similar to that of the *Mth* eIF5B subdomain III was obtained, albeit with low-quality parameters, that were difficult to improve. Better results were obtained when the preliminary ARIA-ensemble was used in a second round of structure calculations using the routine CANDID (Herrmann et al. 2002) followed by 10,000-step simulated annealing calculations using torsion angle dynamics (Güntert et al. 1997), as is embedded in CYANA version 2.0. Here, in each iteration, 100 IF2-C1 structures were calculated, from which the 20 with lowest target function were used in the next. Prior to water refinement with the CNS protocol re_h2o.inp, upper limit restraints used for the final iteration in CYANA were used in an annealing round using ARIA. Finally, the 20 structures with lowest energy were selected as the ensemble to be analyzed.

Internal dynamics

^{15}N-R1, ^{15}N-R2 and ^{15}N{^{1}H}-heteronuclear NOE relaxation values were obtained from experiments performed at 600 MHz essentially as described (Houben et al. 2004). ^{15}N-R1 data were obtained from decaying peak intensities in a randomized series of spectra recorded with relaxation delays of 100, 200, 400, 500, 600, 700, 800, 1000, and 1200 msec. R2-values were extracted from CPMG spectra recorded with relaxation delays of 8, 16, 24, 36, 48, 76, 96, 120, 148, and 192 msec. Signal intensities were fitted using the procedure Curvefit (A.G. Palmer III) and error values of 3% were estimated. Heteronuclear NOEs were obtained from normalized signal intensity differences for experiments recorded with proton saturation for 2.5 sec and without proton saturation; their error values were fixed at 0.05. Relaxation rates and heteronuclear NOE values are added to BMRB accession number 6577. Prior to model-free analysis (Lipari and Szabo 1982) with the program FAST-Modelfree (Cole and Loria 2003), assuming axial symmetric rotational diffusion, the IF2-C1 core of the lowest energy structure was reoriented in its frame of inertia using PDBINERTIA (A.G. Palmer III), after which the full-length IF2-C1 was fit to this orientation.

Acknowledgments

S.T. acknowledges receipt of a Marie Curie Fellowship from the European Community. This work was supported by the NMR Large Scale Facility scheme of the European Union and the Council for Chemical Research of the Netherlands Foundation for Scientific Research (NWO-CW subsidy no. 97016 and a program subsidy), and in part by a MIUR-PRIN 2003 grant to C.O.G. We express our gratitude to Dr. S. Meunier and Dr. S. Rothkrantz-Kos for their contributions at early stages of the project, and Dr. E. AB and Dr. A. Nederveen for providing tools for structure calculations.

References

Boelens, R. and Gualerzi, C.O. 2002. Structure and function of bacterial initiation factors. *Curr. Protein Pept. Sci.* **3:** 107–119.

Cole, R. and Loria, P. 2003. FAST-Modelfree: A program for rapid automated analysis of solution NMR spin-relaxation data. *J. Biomol. NMR* **26:** 203–213.

Cornilescu, G., Delaglio, F., and Bax, A. 1999. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR* **13:** 289–302.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. 1995. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6:** 277–293.

Gualerzi, C.O., Severini, M., Spurio, R., La Teana, A., and Pon, C.L. 1991. Molecular dissection of translation initiation factor IF2. Evidence for two structural and functional domains. *J. Biol. Chem.* **266:** 16356–16362.

Gualerzi, C.O., Brandi, L., Caserta, E., Garofalo, C., Lammi, M., La Teana, A., Petrelli, D., Spurio, R., Tomsic, J., and Pon, C.L. 2001. Initiation factors in the early events of mRNA translation in bacteria. *Cold Spring Harbor Symp. Quant. Biol.* **66:** 363–376.

Guenneugues, M., Caserta, E., Brandi, L., Spurio, R., Meunier, S., Pon, C.L., Boelens, R., and Gualerzi, C.O. 2000. Mapping the fMet-tRNA$_f^{Met}$ binding site of initiation factor IF2. *EMBO J.* **19:** 5233–5240.

Güntert, P., Mumenthaler, C., and Wüthrich, K. 1997. Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* **273:** 283–298.

Herrmann, T., Güntert, P., and Wüthrich, K. 2002. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.* **319:** 209–227.

Houben, K., Dominguez, C., Van Schaik, F.M.A., Timmers, M., Bonvin, A.M.J.J., and Boelens, R. 2004. Solution structure of the ubiquitin-conjugating enzyme UbcH5B. *J. Mol. Biol.* **344:** 513–526.

Johnson, B.A. and Blevins, R. 1994. NMRView: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR* **4:** 603–614.

Koradi, R., Billeter, M., and Wüthrich, K. 1996. MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14:** 51–55, 29–32.

Krafft, C., Diehl, A., Laettig, S., Behlke, J., Heinemann, U., Pon, C.L., Gualerzi, C.O., and Welfle, H. 2000. Interaction of fMet-tRNA^{fMet} with the C-terminal domain of translational initiation factor IF2 from *Bacillus stearothermophilus*. *FEBS Lett.* **471:** 128–132.

Laursen, B.S., Kjærgaard, A.C., Mortensen, K.K., Hoffman, D.W., and Sperling-Petersen, H.U. 2004. The N-terminal domain (IF2N) of bacterial translation initiation factor IF2 is connected to the conserved C-terminal domains by a flexible linker. *Protein Sci.* **13:** 230–239.

Linge, J.P., O'Donoghue, S.I., and Nilges, M. 2001. Automated assignment of ambiguous nuclear overhauser effects with ARIA. *Methods Enzymol.* **339:** 71–90.

Lipari, G. and Szabo, A. 1982. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.* **104:** 4546–4559.

Marzi, S., Knight, W., Brandi, L., Caserta, E., Soboleva, N., Hill, W.E., Gualerzi, C.O., and Lodmell, J.S. 2003. Ribosomal localization of translation initiation factor IF2. *RNA* **9:** 958–969.

Meunier, S., Spurio, R., Czisch, M., Wechselberger, R., Geunneugues, M., Gualerzi, C.O., and Boelens, R. 2000. Structure of the fMet-tRNA^{fMet}-binding domain of *B. stearothermophilus* initiation factor IF2. *EMBO J.* **19:** 1918–1926.

Misselwitz, R., Welfle, K., Krafft, C., Gualerzi, C.O., and Welfle, H. 1997. Translational initiation factor IF2 from *Bacillus stearothermophilus*: A spectroscopic and microcalorimetric study of the C-domain. *Biochemistry* **36:** 3170–3178.

Moreno, J.M.P., Dyrskjøtersen, L., Kristensen, J.E., Mortensen, K.K., and Sperling-Petersen, H.U. 1999. Characterization of the domains of *E. coli* initiation factor IF2 responsible for recognition of the ribosome. *FEBS Lett.* **455:** 130–134.

Roll-Mecak, A., Cao, C., Dever, T.E., and Burley, S.K. 2000. X-ray structures of the universal translation initiation factor IF2/eIF5B: Conformational changes on GDP and GTP binding. *Cell* **103:** 781–792.

Schubert, M., Labudde, D., Oschkinat, H., and Schmieder, P. 2002. A software tool for the prediction of Xaa-Pro peptide bond conformations in proteins based on ^{13}C chemical shift statistics. *J. Biomol. NMR* **24:** 149–154.

Spurio, R., Severini, M., La Teana, A., Canonaco, M.A., Pawlik, R.T., Gualerzi, C.O., and Pon, C.L. 1993. Novel structural and functional aspects of translation initiation factor IF2. In *The translational apparatus* (eds. K.H. Nierhaus et al.), pp. 241–252. Plenum Press, New York.

Spurio, R., Brandi, L., Caserta, E., Pon, C.L., Gualerzi, C.O., Misselwitz, R., Krafft, C., Welfle, K., and Welfle, H. 2000. The C-terminal subdomain (IF2 C2) contains the entire fMet-tRNA binding site of initiation factor IF2. *J. Biol. Chem.* **275:** 2447–2454.

Vachon, G., Laalami, S., Grunberg-Manago, M., Julien, R., and Cenatiempo, Y. 1990. Purified internal G-domain of translational initiation factor IF-2 displays guanine nucleotide binding properties. *Biochemistry* **29:** 9728–9733.

Wishart, D.S. and Sykes, B.D. 1994. The ^{13}C chemical-shift index: A simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. *J. Biomol. NMR* **4:** 171–180.