

A model for the interaction of the G3-subdomain of *Geobacillus stearothermophilus* IF2 with the 30S ribosomal subunit

Ramachandra Dongre,¹ Gert E. Folkers,¹ Claudio O. Gualerzi,²
Rolf Boelens,^{1*} and Hans Wienk^{1*}

¹Department of Chemistry, NMR Spectroscopy, Bijvoet Center for Biomolecular Research, Utrecht University, The Netherlands

²Laboratory of Genetics, Department of Biosciences and Biotechnology, University of Camerino, Italy

Received 8 April 2016; Accepted 29 June 2016

DOI: 10.1002/pro.2977

Published online 1 July 2016 proteinscience.org

Abstract: Bacterial translation initiation factor IF2 complexed with GTP binds to the 30S ribosomal subunit, promotes ribosomal binding of fMet-tRNA, and favors the joining of the small and large ribosomal subunits yielding a 70S initiation complex ready to enter the translation elongation phase. Within the IF2 molecule subdomain G3, which is believed to play an important role in the IF2-30S interaction, is positioned between the GTP-binding G2 and the fMet-tRNA binding C-terminal subdomains. In this study the solution structure of subdomain G3 of *Geobacillus stearothermophilus* IF2 has been elucidated. G3 forms a core structure consisting of two β -sheets with each four anti-parallel strands, followed by a C-terminal α -helix. In line with its role as linker between G3 and subdomain C1, this helix has no well-defined orientation but is endowed with a dynamic nature. The structure of the G3 core is that of a typical OB-fold module, similar to that of the corresponding subdomain of *Thermus thermophilus* IF2, and to that of other known RNA-binding modules such as IF2-C2, IF1 and subdomains II of elongation factors EF-Tu and EF-G. Structural comparisons have resulted in a model that describes the interaction between IF2-G3 and the 30S ribosomal subunit.

Keywords: protein synthesis; translation initiation; protein structure; bacterial IF2; ribosomal subunits; NMR

Introduction

Three proteins play a crucial role in the early stages of protein biosynthesis on ribosomes in bacteria. These are the initiation factors IF1, IF2, and IF3 that act in concert as kinetic effectors of various steps of the translation initiation pathway and ensure both speed and accuracy to this process (for

reviews see Ref. 1–5). The largest of these factors is the multi-domain protein IF2 which is ~92 kDa in *Escherichia coli* (*Eco*)⁶ and ~82 kDa in *Geobacillus* (formerly *Bacillus*) *stearothermophilus*^{7,8} (*Gst*). IF2 is responsible for recruiting initiator fMet-tRNA to the 30S subunit and for promoting its decoding of the mRNA initiation triplet in the P-site of the 30S

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: The Netherlands Organisation for Scientific Research; Grant number: NWO/CW Echo project 700.56.010; Grant sponsor: European Commission; Grant number: EC FP7/2007-2013, project 261863, BioNMR.

*Correspondence to: Rolf Boelens, NMR Spectroscopy, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands. E-mail: r.boelens@uu.nl or Hans Wienk, NMR Spectroscopy, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands. E-mail: h.l.j.wienk@uu.nl

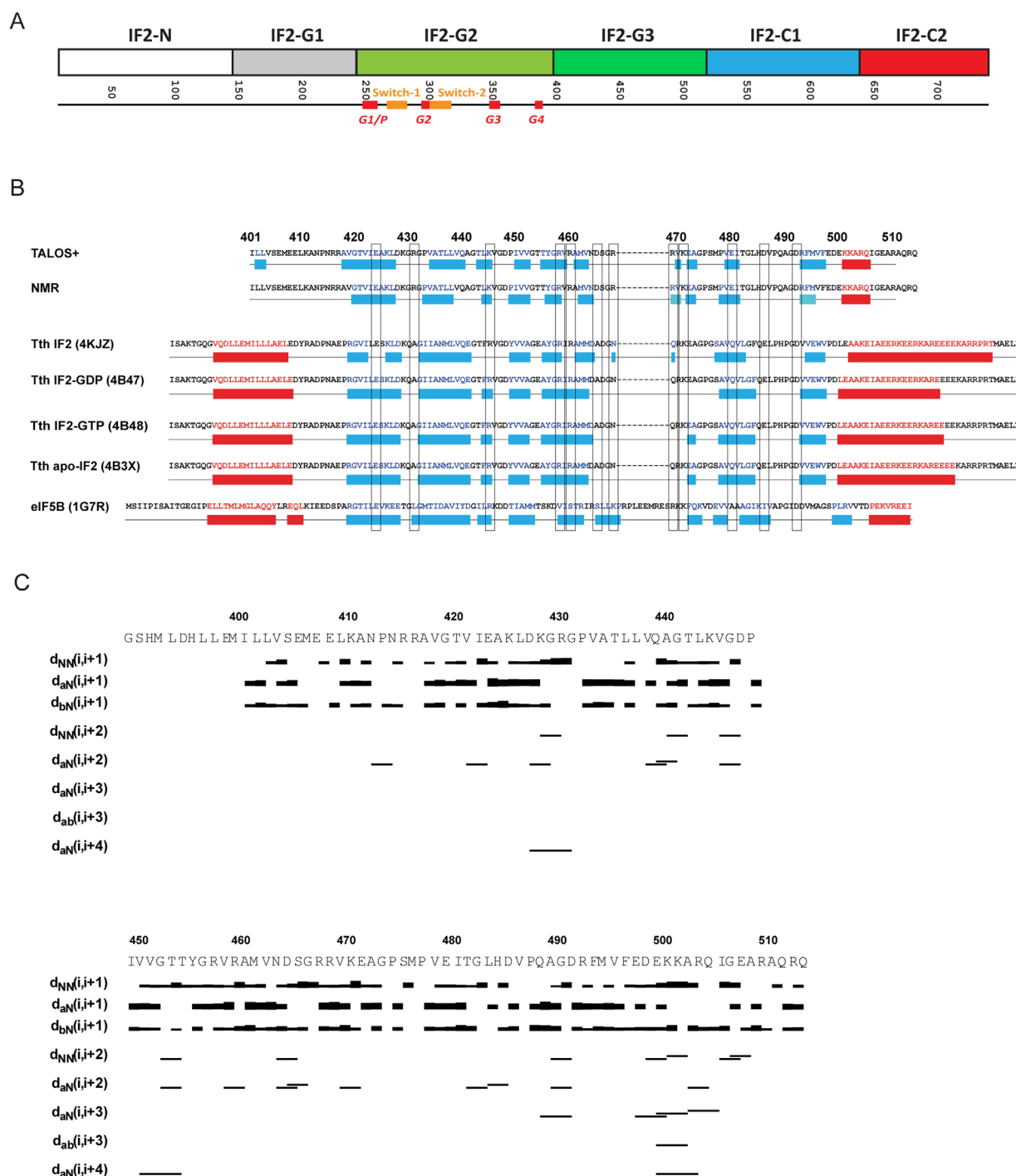


Figure 1. Domain organization of *Gst* IF2 and secondary structure elements of the IF2-G3 subdomain. (A) Domain organization of bacterial IF2 (*Gst*). Below the domain organization, in orange the regions Switch-1 and Switch-2 are indicated, in red the G1-G4 boxes and phosphate binding region. (B) Secondary structure elements of *Gst* IF2-G3 and amino acid sequence alignment with eIF5B subdomain II and *Tth* IF2-G3. Helices and β -strands from reported eIF5B-II and IF2-G3 structures (1G7R,¹⁴ 4KJZ,¹⁵ 4B47,¹⁶ 4B48,¹⁶ and 4B3X¹⁶) are colored red and blue, respectively; boxes show residues that are part of positively and negatively charged surface patches in *Gst* IF2-G3. (C) Overview of sequential and medium-range NOEs used to calculate the *Gst* IF2-G3 structure ensemble.

ribosomal subunit, yielding a 30S initiation complex (30SIC).⁹ IF2 also favors the association of the 30SIC with the 50S ribosomal subunit^{10,11} and the adjustment of the fMet-tRNA orientation in the P-site of the peptidyl transferase center of the 50S

ribosomal subunit, where it can act as a donor in the formation of the initiation dipeptide.^{12,13}

Bacterial IF2 consists of three domains (N, G, and C), each composed of different subdomains with distinct structures and functions [Fig. 1(A)]. The N-

Table I. Predicted and Experimental Translational and Rotational Diffusion Parameters for the *Gst* IF2-G3 Subdomain

Translational diffusion	Full length ^a	Core+helix ^b	Core ^c
Theoretical M_w (kDa)	13.6	10.2	7.9
DOSY M_w (kDa)	11.9		

Rotational diffusion	Full length ^{a,b}	Core+helix ^{a,c}	Core ^{a,d}	FAST-MF ^{b,e}	TENSOR2 ^{e,f}
D_{xx} (s ⁻¹)	1.49×10^7	2.39×10^7	3.56×10^7		2.33×10^7
D_{yy} (s ⁻¹)	1.57×10^7	2.51×10^7	3.80×10^7		3.04×10^7
D_{zz} (s ⁻¹)	2.41×10^7	3.53×10^7	4.09×10^7		3.48×10^7
D_{ratio}	1.58	1.44	1.11	1.25 ± 0.05	1.30
τ_c (ns)	9.14	5.94	4.37	5.95 ± 0.06	5.63 ± 0.04

^a Predicted rotational diffusion parameters (HYDRONMR).^b All residues of the G3-construct, i.e. GSHM-Leu394–Gln514.^c Gly420–Ile506.^d Arg420–Phe495.^e Rotational diffusion parameters from relaxation data.^f Arg420–Phe495, residues in β -strands.

terminal domain (NTD), in bacteria not conserved in terms of sequence and length, is responsible for high-affinity yet non-specific anchoring of IF2 to the ribosome.^{17–19} Between the NTD and the C-terminal domain is the large (~42 kDa) G-domain, consisting of three subdomains (G1, G2, and G3).²⁰ The function of G1 is still elusive but it has been suggested that it may have a role in favoring subunit association.⁵ G2 contains a guanosine nucleotide binding motif and is able to bind 50S ribosomal subunits, albeit with low affinity.²⁹ It has been suggested that G3 facilitates the interaction of IF2 with the 30S and possibly also with the 50S ribosomal subunit.^{19,22–25} C-terminally from G3 are subdomains C1 and C2. The likely function of C1 is to establish contacts with G2/G3 and to mediate communication between these G subdomains and C2, the most C-terminal subdomain of IF2 that was shown to be necessary and sufficient to recognize and bind fMet-tRNA.^{26,27}

Overall, the function of bacterial IF2 relies to a large extent on its dynamic architecture that allows it to interact with the ribosomal subunits and to accompany the ribosome through subsequent steps of the translation initiation pathway. Therefore, atomic-resolution structures and detailed knowledge of intersubdomain dynamics are of utmost importance to understand how IF2 exactly functions. So far, the solution structures of the N-terminal fifty residues of the *Eco* NTD and of isolated *Gst* subdomains G2, C1, and C2 have been elucidated by NMR spectroscopy.^{27–30} In addition to these structures, recent studies reported cryo-EM and crystallographic structures of NTD-, G- and C1-subdomains¹⁶ of *Thermus thermophilus* (*Tth*) IF2 and full-length *Eco* IF2.¹⁵ As was already inferred from other data,^{14,31} the conformation of the G2/G3 core is guanosine nucleotide-dependent and may be responsible for transmitting conformational changes to distant

regions of the factor. The G3 subdomain may have an important role in communicating this allosteric response to the C-terminal parts of IF2.

So far, no structure of an isolated G3 subdomain has been reported. However, in order to understand the mechanism by which G3 may perform such communication function, it is important to be able to compare its structure within intact IF2 with that of G3 without interactions with other subdomains of the factor. Here, we present the solution structure, tumbling properties and internal dynamics of the isolated *Gst* G3 subdomain. Furthermore, to understand better the nature of potential cross-talk occurring between subdomains G2, G3, and possibly C1, we have undertaken the comparison of our NMR structure with similar structures reported before. This allowed us to build a model for the mechanism by which G3 binds to the small ribosomal subunit.

Results and Discussion

Resonance assignments and secondary structure of *Gst* IF2-G3

After overexpression and purification of *Gst* G3, Diffusion Ordered NMR Spectroscopy (DOSY) and ¹⁵N-relaxation experiments were carried out. These experiments indicated that isolated IF2-G3 in solution undergoes translational and rotational diffusion characteristic of a particle not bigger than a monomer (Table I). The analysis of triple resonance spectra yielded assignments for 102 amides of the 125 residues that constitute the IF2-G3 construct (Fig. 2). Six of the 23 residues with unassigned amides are prolines, whereas the others are the N-terminal region residues GSHML³⁹⁴DHLEMI⁴⁰¹ (i.e., including the GSHM tetrapeptide resulting from the cloning procedure), the region Asn415–Arg417, and the dipeptide Glu498–Asp499. Residues flanking the latter two regions showed reduced signal intensities in

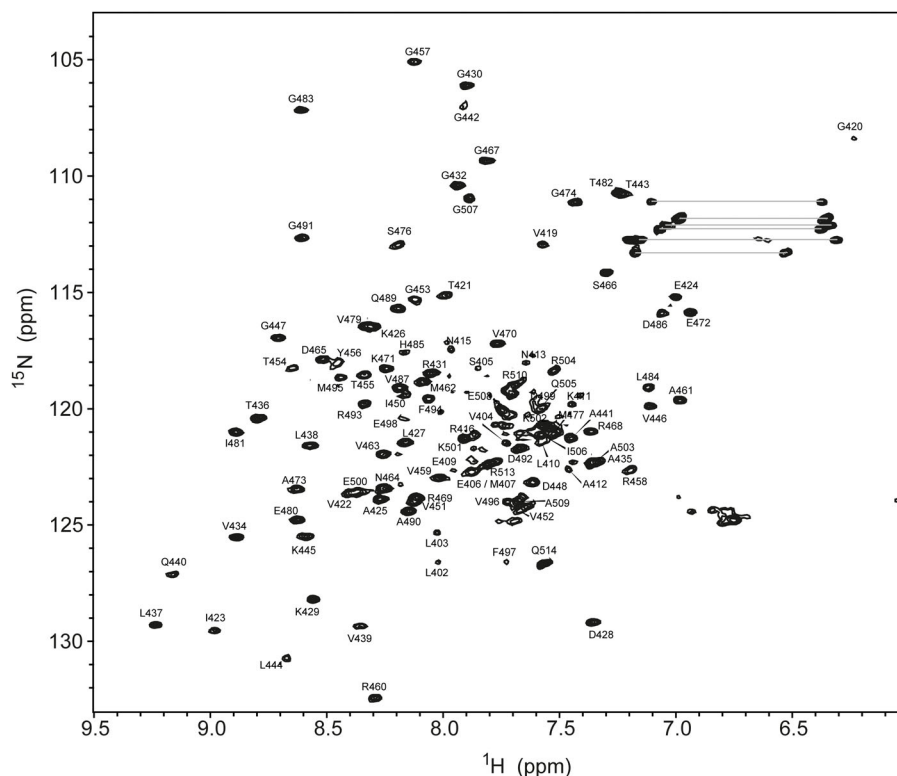


Figure 2. Assigned 2D [^{15}N ; ^1H]-HSQC spectrum of Gst IF2-G3. Asparagine and glutamine side chain NH_2 -groups are connected by a line.

2D [^{15}N ; ^1H]-HSQC (Heteronuclear Single Quantum Correlation) spectra, probably as the result of line broadening, indicating slow conformational dynamics. After analysis of amino acid side chains, assignments for 113 assignable (pseudo-)atoms were still missing, mostly of those localized in the same regions for which no backbone assignment was possible. The final completeness of assignment of non-labile and amide protons is 80.5%; 87.6% if the N-terminal eleven residues from the construct are not considered.

Based on the combined C' , $\text{C}\alpha$, $\text{C}\beta$, $\text{H}\alpha$, amide nitrogen and amide proton chemical shifts, the positions and lengths of secondary structure elements were predicted with TALOS+ [Fig. 1(B)]. Predicted β -strands involve residues Leu402-Leu403, Ala418-Leu427, Val434-Gln440, Thr443-Lys445, Ile450-Val452, Thr455-Val459, Ala461-Val463, Val470, Glu472-Ala473, Val479-Ile481, and Arg493-Phe497. Residues Lys501-Gln505 may form a single, C-terminal helical region. The overall positions of these secondary structure elements closely follow those of G3 structures that have been determined in the context of larger IF2 and eIF5B homologs, either in apo-form, or in GDP- or GTP-bound states.^{14–16}

Structure and dynamics of the Gst IF2-G3 subdomain

The solution structure of Gst IF2-G3 was calculated using NMR-based distance restraints and dihedral

angle restraints. For this, the combination of manual assignments and automatic assignment procedure from CYANA provided a total of 1453 distance restraints based on 3427 out of 5003 NOESY signals. Furthermore, TALOS+-predicted ϕ and ψ angles were converted into 149 dihedral angle restraints. Statistics for the final NMR ensemble are summarized in Table II. As seen in Figure 3, most amino acid residues (i.e., Gly420-Met495) form a well-defined core, with backbone and heavy atom RMSD values of 0.42 Å and 0.85 Å, respectively. An 89.1% of the core residues dihedral angles occur in the favorable region and 10.8% and 0.2% in the additionally and generously allowed regions of the Ramachandran plot, respectively. None of the residues forming the G3-core showed dihedral angles in disallowed regions. The final structure ensemble shows a distance violation >0.5 Å for a single, irrelevant contact involving the unstructured N-terminal tail. There are no dihedral angle violations $>5^\circ$.

The NMR structure shows β -strands in all conformers of the ensemble for Gly420-Leu427, Pro433-Leu438, Leu444-Lys445, Pro449-Val452, Tyr456-Arg458, Met462-Asn464, Glu472-Ala473, and Pro478-Ile481 [Fig. 1(B)]. An α -helix is present from Lys501 to Ile506. In the majority of the conformers forming the ensemble some of the secondary structure elements are extended towards the N- or C-terminal direction. These residues are Arg469-Val470 and Arg493-Met495 (β -strand), while residue

Table II. Statistics for the Structure Calculation of the NMR Ensemble of the *Gst* IF2-G3 Subdomain

Number of NOESY peaks	
(Automatically) assigned	3427
Total	5003
Number of experimental restraints after CYANA	
Total NOE	1453
Intra-residual	316
Sequential	455
Medium-range	179
Long-range	503
Dihedral angles	149
Consistent violations	
NOE > 0.5 Å	1 ^a
Dihedral > 5°	0
RMSD ^b	
bb/heavy (Å)	0.42/0.85
Ramachandran analysis ^{b,c}	
Favorable regions	89.1%
Additionally allowed regions	10.8%
Generously allowed regions	0.2%
Disallowed regions	0.0%

^a Asn413-QD × Thr421-MG.

^b Residues Gly420-Met495.

^c ICING webserver.

Glu500 shows helical propensity. Strikingly, although Ile423 was predicted by TALOS+ to be helical, in the ensemble this residue is located in a long twisted β -strand, albeit with dihedral angles that can be brought into agreement with a helix (i.e., Φ : -86° ; Ψ : -38°).

The core of *Gst* IF2-G3 (Fig. 3) shows three β -sheets, formed by anti-parallel β -strands connected by short inter-strand loops. The first β -sheet is formed by the four short strands Gly420-Thr421,

Arg493-Met495, Pro449-Val452, and Thr455-Arg458. The second β -sheet is formed by five β -strands: Val422-Asp428, Gly432-Leu438, Pro478-Thr482, Met462-Asn464, and Arg469-Val470. The third β -sheet is formed by two short strands Leu444-Lys445 and Glu472-Ala473. Overall, the *Gst* IF2-G3 core structure is well-packed due to hydrophobic interactions and stabilization by hydrogen bonds. Most of these hydrogen bonds support backbone-backbone interactions or are intra-helical. However, seven H-bonds (i.e., Asp428-O^{D1} × Gly430-H^N, Asp428-O^{D1} × Arg431-H^N, Ala461-H^N × Thr482-O^{G1}, Ser466-H^N × Asn464-O^{D1}, Arg468-H^{H11} × Ser466-O, Arg469-H^{H11} × Val470-O, Arg493-H^N × Asp492-O^{D2}) are formed between backbone and side chains. Comparison with other protein structures in the Protein Data Bank using the DALI server, indicates that the *Gst* IF2-G3 subdomain core adopts a typical OB fold.³² These are globular units built up from two anti-parallel β -sheets that share one β -strand. In the current case this shared β -strand is the one from Val419 to Asp428.

In the G3-ensemble, both the N-terminal (Ile401-Arg416) and C-terminal (Glu498-Gln514) regions are not uniformly positioned with respect to the β -barrel core. This is consistent with the fact that no long-range contacts were found between the core and either terminus. The N-terminal region is completely devoid of structure despite the fact that it was expected to be helical based on structures found for homologous G3-subdomains [Fig. 1(B)]. This could be due to poorly chosen boundaries that interfere with the formation of this helix in the construct. At the C-terminus residues Glu500-Ile506 form a short helix. This helix would probably be extended if extra amino acids were present, since some of the chemical shifts do show helical propensity further toward the C-terminus.

Model-free analysis of NMR relaxation data indicates that the isolated and structured *Gst* IF2-G3 construct tumbles in solution as a slightly

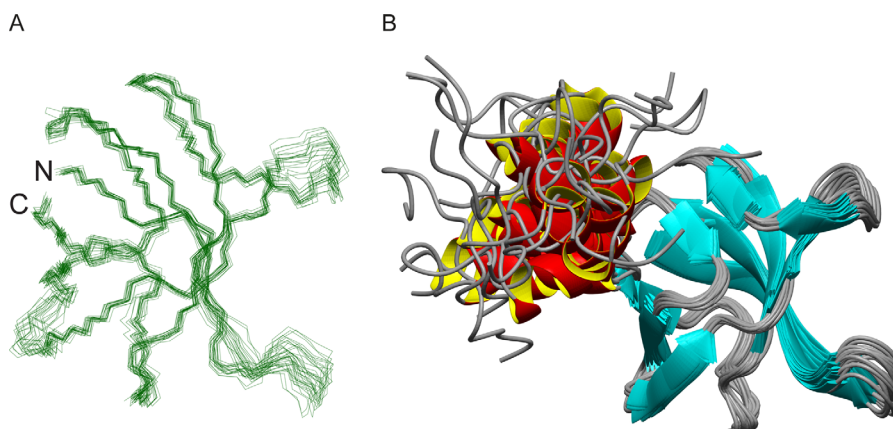


Figure 3. The NMR structure of the *Gst* IF2-G3 NMR subdomain (PDB-entry 2NBG). Left: organization of the OB core. Right: arrangement of *Gst* IF2-G3 OB fold core and C-terminal helix. Structures are fit on the core residues Gly420-Met495.

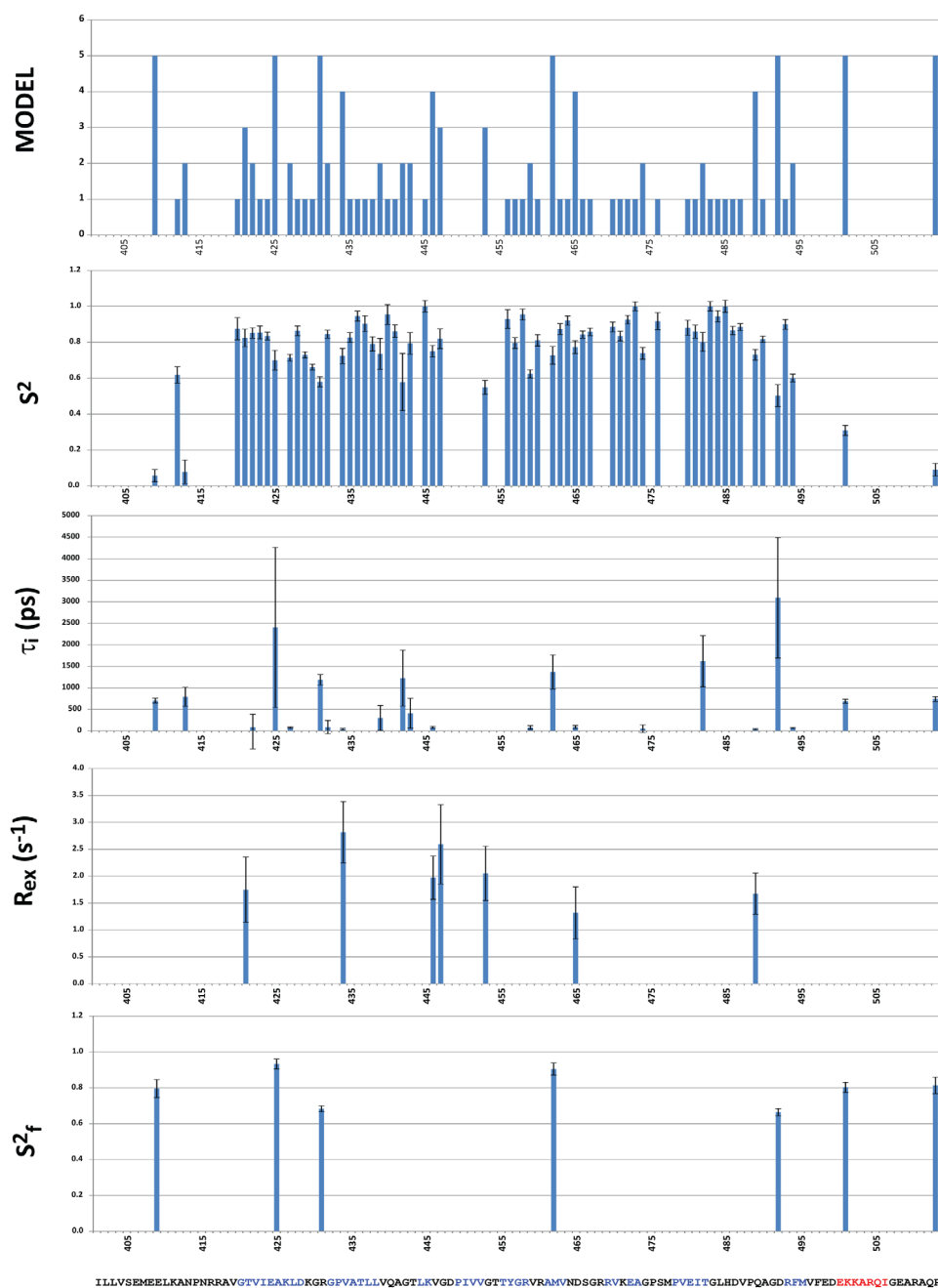


Figure 4. Internal dynamics following Model-free analysis of NMR relaxation data for Gst IF2-G3. From top to bottom: model number used to describe the dynamics, squared general order parameter S^2 , fast internal dynamics τ_i (nsec), slow conformational exchange R_{ex} (s^{-1}) and the fast general order parameter S^2_f .

flattened monomer, with the ratio between long and short rotational diffusion tensor axes (D_{ratio}) of about 1.3 (Table I). The rotational diffusion tensor, as well as the tumbling time of IF2-G3 closely resembles values expected for the NMR-determined G3 β -barrel core plus C-terminal helix (i.e., Arg420-Ile506). The fact that the anisotropy does not completely agree with what is predicted supports the premise that the C-terminal helix does not have a uniform orientation with respect to the β -sheet

core. Analysis of internal dynamics confirms that IF2-G3 mainly forms a rigid unit. On the other hand, reduced NOE and R_2/R_1 ratios show that N- and C-terminal extensions from the core undergo primarily fast internal dynamics (Fig. 4, Supporting Information Fig. 1). It is possible that the helical region at the C-terminus influences the tumbling behavior of the G3 β -sheet core, but is unstable and lacks a structural scaffold that would restrict it in a single conformation.

Table III. Statistics for Amino Acid and Three-dimensional Structure Comparison Between the Structured Core of *Gst* IF2-G3 and Other Domains from ClustalW and DALI

	<i>Tth</i> IF2-G3	eIF5B-"G3" ^a
Identity	42%	20%
Homology	81%	54%

	Z score	RMSD (Å)	Equiv. res.	PDB-entry
<i>Tth</i> IF2-G3	12.8	1.3	75	4B43
<i>Tth</i> IF2-G3	12.5	1.3	75	4KJZ
aEF-1A ^a	11.5	1.4	74	3AGJ
eIF5B-"G3"	10.5	1.9	75	1G7S
<i>Tth</i> EF-Tu	11.2	1.5	74	4LBY
eEF-Tu	11.0	1.4	74	1D2E
<i>Gst</i> IF2-C2	8.6	2.0	74	1D1N
eIF5B-"C2"	8.6	2.3	72	1G7S

^a e: eukaryotic; a: archaeal domain.

Towards *Gst* IF2-G3 functionality

Table III shows statistics for ClustalW and DALI alignments of the *Gst* IF2-G3 OB-fold core elucidated here, with subdomains G3 and C2 of IF2 from other bacterial species and of archaeal eIF5B. Interestingly, the DALI search revealed that the core *Gst* IF2-G3 fold is more closely related to domains belonging to other factors involved in protein synthesis, such as Elongation Factors EF-1 α ³³ and EF-Tu,³⁴ than to the archaeal/eukaryotic IF2-homolog a/eIF5B¹⁴ (Fig. 5). The amino acid homology among bacterial IF2s is anyway much higher than between bacterial IF2 and archaeal eIF5B, as seen also from Figure 1(B). The fact that the structural homology

with *Thermus thermophilus* (*Tth*) IF2-G3 is much higher than with subdomain III (corresponding to G3) of a/eIF5B, points toward a general difference between bacterial and archaeal/eukaryotic forms of this subdomain. This most likely relates to the overall different functions of IF2 and a/eIF5B in translation initiation.

Gst IF2-G3 is structurally homologous to other OB-fold proteins, including ribosomal protein S10, aspartyl tRNA synthetase, the C2-subdomain of *Gst* IF2, and cold shock protein CspA. A common property to many of these proteins is their capacity to bind RNA.^{21–23} Results obtained for hydroxyl radical cleavage by Fe(II)-EDTA tethered to cysteine residues in IF2,³⁵ for amino acid substitutions (Gly420 and Glu424),¹⁹ for cryo-EM reconstructions,^{21–23} as well as for titrations with ribosomal subunits (RD, unpublished results) indicate that IF2-G3 is indeed involved in binding of 30S ribosomal subunits and participates in joining the small and large ribosomal subunits during the transition from 30S_{IC} to 70S_{IC}.

In full-length bacterial IF2, N- and C-terminal helices connect G3 to subdomains G2 and C1, respectively. The present relaxation data agree with functional models that indicate a positional reorganization of G3 during protein synthesis with respect to the other subdomains. In turn, this could be due to internal destabilization and/or due to a reorientation of the connector helices. The rigidity of the *Gst* IF2-G3 core and the lack of stable interactions between core and the N- and C-terminal region suggest that IF2-G3 can mediate inter-subdomain surface interactions. Concerning this point it is interesting that patches with non-neutral potential are present on the surface of G3 (Fig. 6). Two small negatively charged regions are formed by Glu424,

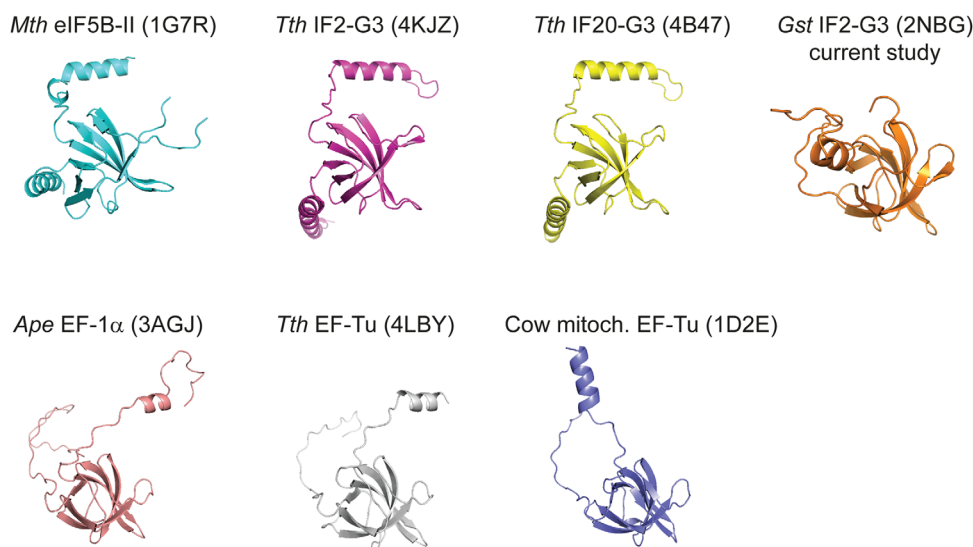


Figure 5. Structures of G3-subdomains from other proteins aligned with the current structure of *Gst* IF2-G3. The linker to the preceding G2-subdomain is the higher free terminus, the linker to the C1-subdomain the lower. *Mth*: *Methanothermobacter thermautotrophicus*; *Ape*: *Aeropyrum pernix*. PDB IDs are indicated.

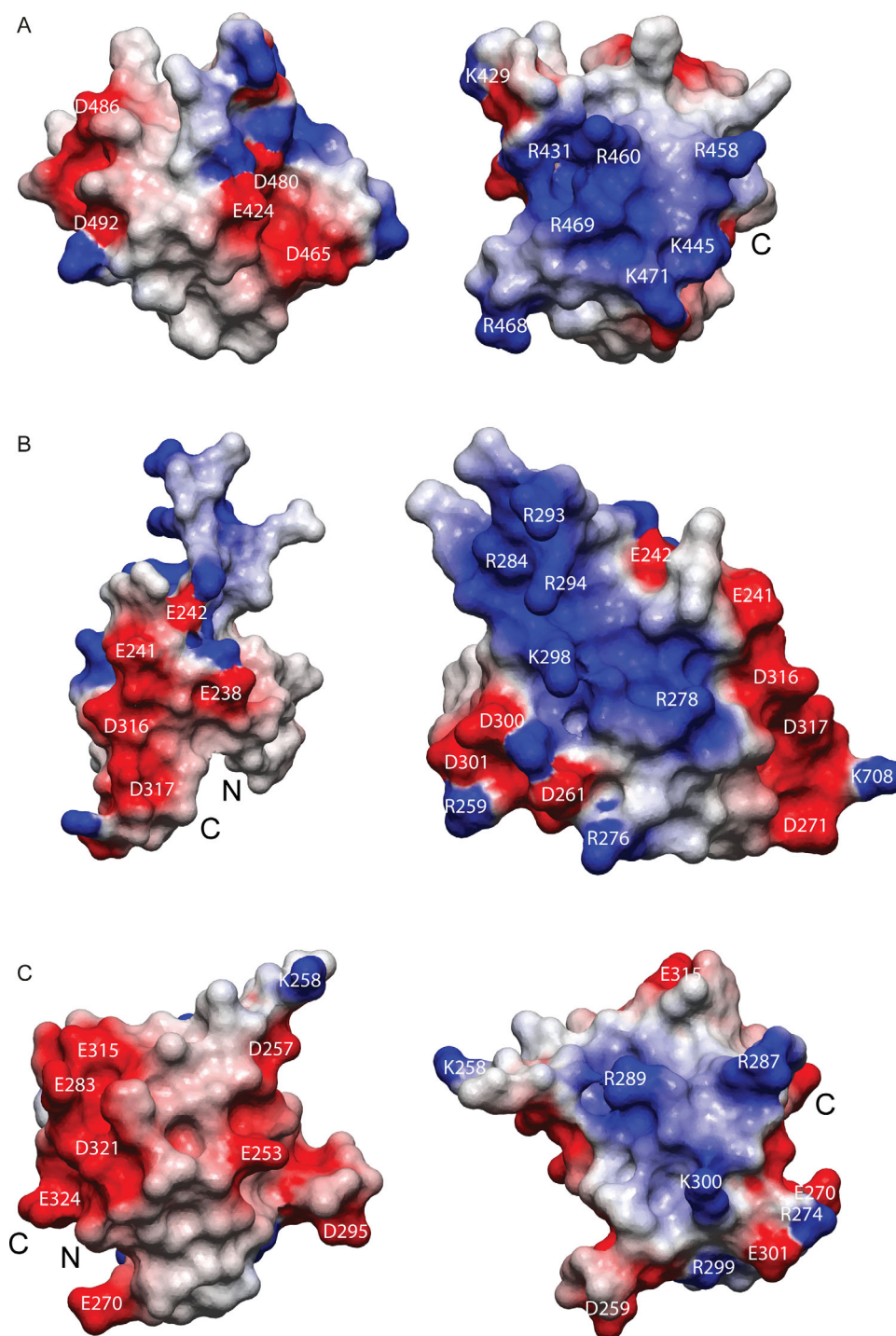


Figure 6. Surface potential for selected G3-subdomains. (A) The closest to average NMR structure of Gst IF2-G3 (PDB-entry 2NBG; model 12). To go from left to right orientation the model was rotated: x,y,z : $+40^\circ, -140^\circ, 0^\circ$. (B) The crystal structure of eIF5B G3-homologous subdomain-II.¹⁴ To go from left to right orientation the model was rotated: x,y,z : $+60^\circ, +120^\circ, 0^\circ$. (C) The crystal structure of *Tth* IF2-G3.¹⁶ To go from left to right orientation the model was rotated: x,y,z : $+30^\circ, -170^\circ, 10^\circ$. Red: overall negatively charged surface potential. Blue: overall positively charged surface potential. Structures are manually aligned on charged residues that are conserved among the three proteins [Fig. 1(B)].

Asp465, Glu480, Asp486, and Asp492, while a rather extended, positively charged region is formed by Arg431, Lys445, Arg458, Arg460, Arg468, Arg469, and Lys471. It has been suggested that the negatively charged residue Glu424 is involved in an

interaction with IF2-G2.³⁶ Furthermore, the interaction between the negatively charged patch formed by Glu424, Asp465, and Asp480 and IF2-G2 would be compatible with the crystal structures of *T. thermophilus* IF2.¹⁶ As a matter of fact these negatively

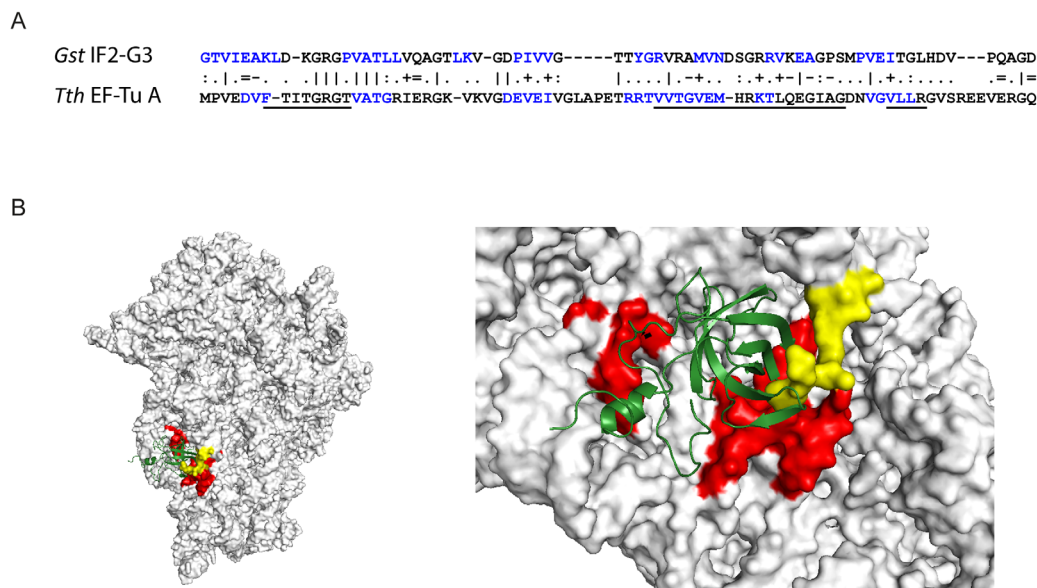


Figure 7. Model for the interaction between *Gst* IF2-G3 and the 30S ribosomal subunit based on structural homology between *Gst* IF2-G3 and *Tth* EF-Tu A, and the complex of the latter with the 70S ribosome (pdb-entry 4V5G³⁸). (A) Amino acid conservation between *Gst* IF2-G3 and EF-Tu A OB-folds. (B) Model for the interaction between *Gst* IF2-G3 and the 30S ribosomal subunit. White: 30S ribosomal subunit; green: *Gst* IF2-G3 subdomain; red: known interaction sites for IF2-G3; yellow: part of the ribosomal A-site.

charged residues are extremely conserved. In addition, substitution of Glu424 to lysine within the G3 subdomain is not lethal, but does abolish completely the GTPase activity of G2. This provides clear evidence that there is cross-talk between *Gst* IF2 subdomains G2 and G3. Although the guanosine nucleotide binding capacity is only marginally altered compared to the wild-type protein, this point mutation interferes with the functional interplay between G3 and G2. It is possible that the Glu424Lys mutation prevents interactions between G3 and residues Gly300 and His301 of Switch-II that are involved in GTP-binding by IF2-G2.

Comparison with structures of larger IF2-systems suggests that the negatively charged residues Asp486 and Asp492 can interact with the partially positively charged C-terminal helical linker (H8) that connects G3 with subdomain C1. Such interactions would be bacteria-specific because these negative charges are not conserved in archaeal eIF5B.¹⁴

Comparison of the determined G3 subdomain structure with existing models of larger IF2 systems¹⁵ indicates that the positively charged region comprising residues Arg431, Lys445, Arg458, Arg460, Arg468, Arg469, and Lys471 is exposed to the solvent and amenable for interactions with negatively charged regions of other compounds or IF2-regions. We suggest that interactions between OB-fold and 30S ribosomal subunits are overall very similar for IF2-G3 and domain II of EF-Tu.³⁷ Contacts between the OB-fold of EF-Tu and helices h5 and h14 of the 16S rRNA are present in the crystal structure of the ribosomal complex.³⁸ Similarly,

cryo-EM reconstitution studies have positioned IF2-G3 in contact with rRNA helices h5 and h14.^{23,24} Furthermore, in the context of 30S-bound IF2, the G3-residue V451 has been shown to bind the nearby 16S rRNA helices h3/h4 and at the base of h17.³⁶ A model for *Gst* IF2-G3 binding to the 30S subunit, based on amino acid conservation between *Gst* IF2-G3 and *Tth* EF-Tu, and on the structural details of the EF-Tu interaction with the 70S ribosome³⁸ is presented in Figure 7. In the model presented, IF2-G3 binds to the 30S subunit where other subdomains of IF2 may interfere with the binding of the 50S ribosomal subunit. This interaction is in overall agreement with data obtained for a eukaryotic system consisting of eIF5B and 80S ribosomes.³⁹ The eIF5B G3-domain is localized to the same region on the 40S ribosomal subunit, involving similar helices of the rRNA. Although the interacting regions of IF2, EF-Tu and eIF5B with the small ribosomal subunit may be similar, details may be different. For instance, comparison of the surface potentials of the OB-folds suggests that binding of *Gst* IF2-G3 to the 30S ribosomal subunit may rely more on electrostatic interactions than that of the EF-Tu OB-fold.³⁸ The G3-subdomain blocks the 30S A-site, thereby preventing premature tRNA-binding to this site before the 70SIC is formed.

Materials and Methods

NMR sample preparation

The *Gst* IF2-G3 subdomain was expressed from a pEV-expression vector in *E. coli* strain BL21. The

construct encodes all 114 IF2-G3 amino acid residues preceded by an N-terminal cleavable His-tag. Bacteria were grown at 35°C and induced at OD₆₀₀ = 0.6 with a final concentration of 0.1 mM isopropyl β-D-1-thiogalactopyranoside. Cells were harvested at OD₆₀₀ = 1.2 and resuspended in 150 mM NaCl, 50 mM NaPO₄ (pH 8.0), 10 mM imidazole, 1 mg mL⁻¹ lysozyme. After freezing and thawing, cells were disrupted by sonication on ice and insoluble material was pelleted by high-speed centrifugation. The protein was purified from the supernatant using Ni-NTA affinity chromatography and the bound His-tagged protein was eluted with 500 mM imidazole, 150 mM NaCl, and 50 mM NaPO₄ (pH 8.0). Thrombin (Novagen, 2 U/20 mL of 200 μM protein) was used to cleave off the His-tag in 150 mM NaCl, 50 mM NaPO₄ (pH 8.0), and 1 mM CaCl₂. Because the Gst IF2-G3 subdomain appeared prone to aggregation during purification, concentrations of protein above 180 μM were avoided and small amounts of glycerol appeared beneficial to stabilize the protein. Final G3 subdomain purification was achieved by FPLC gel filtration (Pharmacia) in 50 mM NaPO₄ (pH 6.5), 300 mM KCl, 2% glycerol. Fractions containing the G3 subdomain were collected and buffer-exchanged to 50 mM NaPO₄ (pH 6.5), 100 mM KCl, and 2% glycerol.

NMR spectroscopy

NMR experiments on ¹⁵N-labelled IF2-G3 and ¹³C/¹⁵N-labelled IF2-G3 samples were carried out with 0.15 mM protein in 50 mM NaPO₄ (pH 6.5), 100 mM KCl, 2% glycerol, and 5% D₂O. Double and triple resonance experiments were recorded at 303 K on a 600 MHz Bruker Avance III spectrometer equipped with TXI probe. After processing with NMRPipe,⁴⁰ backbone assignments were performed with 2D [¹⁵N;¹H]-HSQC, 2D [¹³C;¹H]-HSQC and 3D HN(CO)CACB, HNCACB, HNCA, HNCO, and HN(CA)CO experiments⁴¹ using the program Sparky (Goddard and Kneller, UCSF). For side chain assignments additional 3D TOCSY-^[15N;1H]-HSQC, 3D HBHA(CO)NH and 3D H(C)CH-TOCSY experiments were used. A 900-MHz Bruker Avance III spectrometer equipped with cryogenic probe was used to obtain 2D NOE, and 3D NOESY-^[15N;1H]-HSQC and NOESY-^[13C;1H]-HSQC experiments. Referencing was performed on internal water.⁴² Secondary structure elements were predicted using TALOS +.⁴³

To determine translational diffusion, Diffusion Ordered NMR Spectroscopy (DOSY) was used with a stimulated echo bipolar pulsed field gradient with single spoil gradient and 3-9-19 watergate (Bruker pulseprogram stebpgp1s19). For relaxation analysis, ¹⁵N-R₁ values were determined from experiments with relaxation delays of 10, 20, 40, 80, 160, 400, 1000, 2000, 4000, 6000, and 7000 ms. ¹⁵N-R₂ values were obtained from Carl-Purcell-Meiboom-Gill

(CPMG) experiments performed with relaxation delays of 32 (2×), 63, 95, 126 (2×), 159, 190, 222, and 286 ms. ¹⁵N{¹H}-heteronuclear NOEs were determined as described in literature.²¹ After two-parameter exponential fitting of ¹⁵N-R₁ and ¹⁵N-R₂ data, and calculation of heteronuclear NOE values, experimental rotational diffusion parameters and internal dynamics were determined by model-free analyses⁴⁴ with FAST-MF^{45,46} and TENSOR2.⁴⁷ Based on the calculated NMR structure, rotational diffusion rate and tensor were predicted with HYDRONMR.⁴⁸

Structure calculations and analysis

For restraint-based structure calculations, TALOS+-predicted backbone φ- and ψ-angles were converted into dihedral angle restraints, and distance restraints were derived from NOESY spectra. Structures were calculated with the program CYANA 3.0⁴⁹ with 20,000 steps of simulated annealing. After each iteration the 20 lowest-energy structures of 100 calculated structures were used for automatic NOE assignments. After eight iterations, water-refinement was performed using the RECOORD protocol.⁵⁰ Here, the final CYANA distance and dihedral angle restraints were used in a single simulated annealing step that was followed by refinement in explicit water, both with the program CNS.⁵¹ NMR structure validation was done using the ICING server (nmr.le.ac.uk/icing). Hydrogen bond and surface potential analyses were performed with the program MolMol⁵²; structures were drawn with MolMol or PyMol (pymol.org). ClustalW (www.ebi.ac.uk/Tools/msa/clustalw2/) was used for pairwise amino acid sequence alignments. On-line DALI servers were used to search homologous structures in the PDB and for pairwise structure alignments (http://ekhidna.biocenter.helsinki.fi/dali_server/start).

Acknowledgments

The authors thank Dr. Suat Özdirekcan for his help during the early stage of the project.

References

1. Gualerzi CO, Brandi L, Caserta E, Garofalo C, Lammi M, La Teana A, Petrelli D, Spurio R, Tomsic J, Pon CL (2001) Initiation factors in the early events of mRNA translation in bacteria. Cold Spring Harbor Symp Quant Biol 66:363–376.
2. Boelens R, Gualerzi CO (2002) Structure and function of bacterial initiation factors. Curr Prot Prot Sci 3:107–119.
3. Laursen BS, Sørensen HP, Mortensen KK, Sperling-Petersen HU (2005) Initiation of protein synthesis in bacteria. Microbiol Mol Biol Rev 69:101–123.
4. Milón P, Rodnina MV (2012) Kinetic control of translation initiation in bacteria. Crit Rev Biochem Mol Biol 47:334–348.

5. Gualerzi CO, Pon CL (2015) Initiation of mRNA translation in bacteria: structural and dynamic aspects. *Cell Mol Life Sci* 72:4341–4367.
6. Steffensen SA, Poulsen AB, Mortensen KK, Sperling-Petersen HU (1997) *E. coli* translation initiation factor IF2—an extremely conserved protein. Comparative sequence analysis of the *infB* gene in clinical isolates of *E. coli*. *FEBS Lett* 419:281–284.
7. Brombach M, Gualerzi CO, Nakamura Y, Pon CL (1986) Molecular cloning and sequence of the *Bacillus stearothermophilus* translational initiation factor IF2 gene. *Mol Gen Genet* 205:97–102.
8. Pon CL, Paci M, Pawlik RT, Gualerzi CO (1985) Structure–function relationship in *Escherichia coli* initiation factors. Biochemical and biophysical characterization of the interaction between IF-2 and guanosine nucleotides. *J Biol Chem* 260:8918–8924.
9. Milon P, Carotti M, Konevega AL, Wintermeyer W, Rodnina MV, Gualerzi CO (2010) The ribosome-bound initiation factor 2 recruits initiator tRNA to the 30S initiation complex. *EMBO Rep* 11:312–316.
10. Antoun A, Pavlov MY, Tenson T, Ehrenberg M (2004) Ribosome formation from subunits studied by stopped-flow and Rayleigh light scattering. *Biol Proced Online* 6:35–54.
11. Qin H, Grigoriadou C, Cooperman BS (2009) Interaction of IF2 with the ribosomal GTPase-associated center during 70S initiation complex formation. *Biochemistry* 48:4699–4706.
12. La Teana A, Pon CL, Gualerzi CO (1996) Late events in translation initiation. Adjustment of fMet-tRNA in the ribosomal P-site. *J Mol Biol* 256:667–675.
13. Tomsic J, Vitali LA, Daviter T, Savelsbergh A, Spurio R, Striebeck P, Wintermeyer W, Rodnina MV, Gualerzi CO (2000) Late events of translation initiation in bacteria: a kinetic analysis. *Embo J* 19:2127–2136.
14. Roll-Mecak A, Cao C, Dever TE, Burley SK (2000) X-ray structures of the universal translation initiation factor IF2/eIF5B: conformational changes on GDP and GTP binding. *Cell* 103:781–792.
15. Sprink T, Ramrath DJ, Yamamoto H, Yamamoto K, Loerke J, Ismer J, Hildebrand PW, Scheerer P, Bürger J, Mielke T, Spahn CM (2016) Structures of ribosome-bound initiation factor 2 reveal the mechanism of subunit association. *Sci Adv* 2:e1501502.
16. Eiler D, Lin J, Simonetti A, Klaholz BP, Steitz TA (2013) Initiation factor 2 crystal structure reveals a different domain organization from eukaryotic initiation factor 5B and mechanism among translational GTPases. *Proc Natl Acad Sci USA* 110:15662–15667.
17. Moreno JM, Drskjotersen L, Kristensen JE, Mortensen KK, Sperling-Petersen HU (1999) Characterization of the domains of *E. coli* initiation factor IF2 responsible for recognition of the ribosome. *FEBS Lett* 455:130–134.
18. Caserta E, Tomsic J, Spurio R, La Teana A, Pon CL, Gualerzi CO (2006) Translation initiation factor IF2 interacts with the 30S ribosomal subunit via two separate binding sites. *J Mol Biol* 362:787–799.
19. Caserta E, Ferrara C, Milon P, Fabbretti A, Rocchetti A, Tomsic J, Pon CL, Gualerzi CO, La Teana A (2010) Ribosomal interaction of *Bacillus stearothermophilus* translation initiation factor IF2: characterization of the active sites. *J Mol Biol* 396:118–129.
20. Gualerzi CO, Severini M, Spurio R, La Teana A, Pon CL (1991) Molecular dissection of translation initiation factor IF2. Evidence for two structural and functional domains. *J Biol Chem* 266:16356–16362.
21. Myasnikov AG, Marzi S, Simonetti A, Giuliodori AM, Gualerzi C, Yusupova G, Yusupov M, Klaholz BP (2005) Conformational transition of initiation factor 2 from the GTP- to GDP-bound state visualized on the ribosome. *Nat Struct Mol Biol* 12:1145–1149.
22. Allen GS, Zavialov A, Gursky R, Ehrenberg M, Frank J (2005) The cryo-EM structure of a translation initiation complex from *Escherichia coli*. *Cell* 121:703–712.
23. Simonetti A, Marzi S, Myasnikov AG, Fabbretti A, Yusupov M, Gualerzi CO, Klaholz BP (2008) Structure of the 30S translation initiation complex. *Nature* 455:416–420.
24. Julián P, Milon P, Agirrezabala X, Lasso G, Gil D, Rodnina MV, Valle M (2011) The Cryo-EM structure of a complete 30S translation initiation complex from *Escherichia coli*. *PLoS Biol* 9:e1001095.
25. Spurio R, Brandi L, Caserta E, Pon CL, Gualerzi CO, Misselwitz R, Krafft C, Welfle K, Welfle H (2000) The C-terminal subdomain (IF2 C-2) contains the entire fMet-tRNA binding site of initiation factor IF2. *J Biol Chem* 275:2447–2454.
26. Guenneugues M, Caserta E, Brandi L, Spurio R, Meunier S, Pon CL, Boelens R, Gualerzi CO (2000) Mapping the fMet-tRNA^{fMet} binding site of initiation factor IF2. *Embo J* 19:5233–5240.
27. Wienk H, Tomaselli S, Bernard C, Spurio R, Picone D, Gualerzi CO, Boelens R (2005) Solution structure of the C1-subdomain of *Bacillus stearothermophilus* translation initiation factor IF2. *Protein Sci* 14:2461–2468.
28. Laursen BS, Mortensen KK, Sperling-Petersen HU, Hoffman DW (2003) A conserved structural motif at the N terminus of bacterial translation initiation factor IF2. *J Biol Chem* 278:16320–16328.
29. Wienk H, Tishchenko E, Belardinelli R, Tomaselli S, Dongre R, Spurio R, Folkers GE, Gualerzi CO, Boelens R (2012) Structural dynamics of bacterial translation initiation factor IF2. *J Biol Chem* 287:10922–10932.
30. Meunier S, Spurio R, Czisch M, Wechselberger R, Guenneugues M, Gualerzi CO, Boelens R (2000) Structure of the fMet-tRNA^{fMet}-binding domain of *B. stearothermophilus* initiation factor IF2. *EMBO J* 19:1918–1926.
31. Simonetti A, Marzi S, Billas IM, Tsai A, Fabbretti A, Myasnikov AG, Roblin P, Vaiana AC, Hazemann I, Eiler D, Steitz TA, Puglisi JD, Gualerzi CO, Klaholz BP (2013) Involvement of protein IF2 N domain in ribosomal subunit joining revealed from architecture and function of the full-length initiation factor. *Proc Natl Acad Sci USA* 110:15656–15661.
32. Theobald DL, Mitton-Fry R, Wuttke S (2003) Nucleic acid recognition by OB-fold proteins. *Annu Rev Biophys Biomol Struct* 32:115–133.
33. Kobayashi K, Kikuno I, Kuroha K, Saito K, Ito K, Ishitani R, Inada T, Nureki O (2010) Structural basis for mRNA surveillance by archaeal Pelota and GTP-bound EF1 α complex. *Proc Natl Acad Sci USA* 107:17575–17579.
34. Grøftehaug MK, Therkelsen MØ, Taaning R, Skrydstrup T, Morth JP, Nissen P (2013) Identifying ligand-binding hot spots in proteins using brominated fragments. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 69:1060–1065.
35. Marzi S, Knight W, Brandi L, Caserta E, Soboleva N, Hill W, Gualerzi CO, Lodmell JS (2003) Ribosomal localization of translation initiation factor IF2. *RNA* 9:958–969.

36. Fabbretti A, Brandi L, Milon P, Spurio R, Pon CL, Gualerzi CO (2012) Translation initiation without IF2-dependent GTP hydrolysis. *Nucleic Acids Res* 40:7946–7955.
37. Fischer N, Neumann P, Konevega AL, Bock LV, Ficner R, Rodnina MV, Stark H (2015) Structure of the *E. coli* ribosome-EF-Tu complex at <3 Å resolution by Cs-corrected cryo-EM. *Nature* 520:567–570.
38. Schmeing TM, Voorhees RM, Kelley AC, Gao YG, Murphy FV, IV, Weir JR, Ramakrishnan V (2009) The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science* 326:688–694.
39. Unbehaun A, Marintchev A, Lomakin IB, Didenko T, Wagner G, Hellen CU, Pestova TV (2007) Position of eukaryotic initiation factor eIF5B on the 80S ribosome mapped by directed hydroxyl radical probing. *EMBO J* 26:3109–3123.
40. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–293.
41. Sattler M, Schleucher J, Griesinger C (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog Nuclear Magn Reson Spectrosc* 34:93–158.
42. Wishart DS, Bigam CG, Yao J, Abildgaard F, Dyson HJ, Oldfield E, Markley JL, Sykes BD (1995) ¹H, ¹³C and ¹⁵N chemical shift referencing in biomolecular NMR. *J Biomol NMR* 6:135–140.
43. Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* 44:213–223.
44. Lipari G, 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.
45. Mandel AM, Akke M, Palmer AG, III (1995) Backbone dynamics of *Escherichia coli* ribonuclease HI: correlations with structure and function in an active enzyme. *J Mol Biol* 246:144–163.
46. Cole R, Loria JP (2003) FAST-Modelfree: a program for rapid automated analysis of solution NMR spin-relaxation data. *J Biomol NMR* 26:203–213.
47. Dosset P, Hus JC, Blackledge M, Marion D (2000) Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J Biomol NMR* 16: 23–28.
48. Garcia de la Torre J, Huertas ML, Carrasco B (2000) HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J Magn Reson* 147:138–146.
49. Herrmann T, Güntert P, 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.
50. Nederveen AJ, Doreleijers JF, Vranken W, Miller Z, Spronk CA, Nabuurs SB, Güntert P, Livny M, Markley JL, Nilges M, Ulrich EL, Kaptein R, Bonvin AM (2005) RECOORD: a recalculated coordinate database of 500+ proteins from the PDB using restraints from the BioMagResBank. *Proteins* 59:662–672.
51. Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998) Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 54:905–921.
52. Koradi R, Billeter M, Wüthrich K (1996) MOLMOL. A program for display and analysis of macromolecular structures. *J Mol Graph* 14:51–55; 29–32.