Solution Structure of C-Terminal *Escherichia coli* Translation Initiation Factor IF2 by Small-Angle X-ray Scattering[†]

Louise Carøe Vohlander Rasmussen,[‡] Cristiano Luis Pinto Oliveira,[§] Janni Mosgaard Jensen,[‡] Jan Skov Pedersen,[§] Hans Uffe Sperling-Petersen,[‡] and Kim Kusk Mortensen*[,‡]

Department of Molecular Biology, and Department of Chemistry, Centre for mRNP Biogenesis and Metabolism, and iNANO Interdisciplinary Nanoscience Center, University of Aarhus, DK-8000 Aarhus C, Denmark

Received January 11, 2008; Revised Manuscript Received April 1, 2008

ABSTRACT: Initiation of protein synthesis in bacteria involves the combined action of three translation initiation factors, including translation initiation factor IF2. Structural knowledge of this bacterial protein is scarce. A fragment consisting of the four C-terminal domains of IF2 from *Escherichia coli* was expressed, purified, and characterized by small-angle X-ray scattering (SAXS), and from the SAXS data, a radius of gyration of 43 ± 1 Å and a maximum dimension of ~145 Å were obtained for the molecule. Furthermore, the SAXS data revealed that *E. coli* IF2 in solution adopts a structure that is significantly different from the crystal structure of orthologous aIF5B from *Methanobacterium thermoautotrophicum*. This crystal structure constitutes the only atomic resolution structural knowledge of the full-length factor. Computer programs were applied to the SAXS data to provide an initial structural model for IF2 in solution. The low-resolution nature of SAXS prevents the elucidation of a complete and detailed structure, but the resulting model for C-terminal *E. coli* IF2 indicates important structural differences between the aIF5B crystal structure and IF2 in solution. The chalice-like structure with a highly exposed α-helical stretch observed for the aIF5B crystal structure was not found in the structural model of IF2 in solution, in which domain VI-2 is moved closer to the rest of the protein.

Initiation of protein synthesis in bacteria is carried out by the aid of three translation initiation factors, one of them being the essential protein, translation initiation factor IF2. Orthologues of this protein are found in both eukaryotes (eIF5B) and archaeabacteria (aIF5B) (*1–3*). Three isoforms termed IF2-1 (97.3 kDa), IF2-2 (79.7 kDa), and IF2-3 (78.8 kDa) are present in *Escherichia coli* as well as in other members of the *Enterobacteriaceae* family (*4, 5*), and *Bacillus subtilis* is the only organism outside the *Enterobacteriaceae* family experimentally determined to express more than one IF2 isoform, namely, IF2-1 and IF2-2 (*6*). The reason for the existence of the smaller isoforms of IF2-1 is not yet known, but they may have a function in cold-shock response (*7, 8*).

During translation initiation, IF2 mediates base pairing of the initiation tRNA anticodon to the mRNA initiation codon located in the ribosomal peptidyl-site (P-site)[1] and stimulates association of the 50S ribosomal subunit to the 30S initiation complex. The role of the IF2 GTPase activity during translation initiation is not yet fully understood, but it has been suggested to be important for IF2 release from the 70S complex (*9–11*), adjustment of initiator tRNA in the P-site (*12*), and cell viability (*13–16*). However, it has also been suggested that neither IF2 ejection from the ribosome nor IF2 recycling require GTP hydrolysis (*17, 18*); therefore, this IF2 activity needs to be further investigated.

IF2 is a multifunctional protein in that it also exhibits chaperone activity, promoting functional folding of proteins and forming stable complexes with unfolded proteins (*19*). In accordance with this, IF2 expression is upregulated during cold-shock conditions (*8, 20*). Another aspect of the role of IF2 in translation is its importance for the translation of leaderless mRNAs (*21–23*).

IF2 was first divided into six domains based on interspecies homology (*24*), and domain VI was subsequently divided into the subdomains VI-1 and VI-2 (Figure 1). The C-terminal part (domains IV−VI-2) is highly conserved among species, whereas the N-terminal region (domains I−III) varies in both length and amino acid composition but shows intraspecies homology in being completely conserved within the strains of *E. coli* (*25*). Domain IV is also known as the G domain, because it is responsible for GTP binding.

Attempts to crystallize full-length *E. coli* IF2-1 have thus far been unfruitful, and the only structural knowledge of IF2 has come from X-ray crystallography of orthologous proteins (*26*), circular dichroism (CD) and nuclear magnetic resonance (NMR) studies of IF2 and IF2 fragments (*27, 28*), and cryo-electron microscopy (cryo-EM) of ribosomal complexes (*29, 30*).

[†] This research was supported by Danish Natural Science Research Council Grants 21-03-0592 and 21-03-0465 and Carlsberg Grants 2005-1-126, 2006-1-167 (to H.U.S.-P. and K.K.M.), and 2007-1-261 (to L.C.V.R.).

* To whom correspondence should be addressed: Department of Molecular Biology, Gustav Wieds Vej 10, DK-8000 Aarhus C, Denmark. Telephone: +45-89425047. Fax: +45-86182812. E-mail: kkm@mb.au.dk.

[‡] Department of Molecular Biology.

[§] Department of Chemistry.

[1] Abbreviations: IF, translation initiation factor; cryo-EM, cryo-electron microscopy; PDB, Protein Data Bank; P-site, peptidyl-site; SAXS, small-angle X-ray scattering; IF2C, domains IV−VI-2 of *E. coli* IF2; DLS, dynamic light scattering.

10.1021/bi8000598 CCC: $40.75
Published on Web 04/29/2008

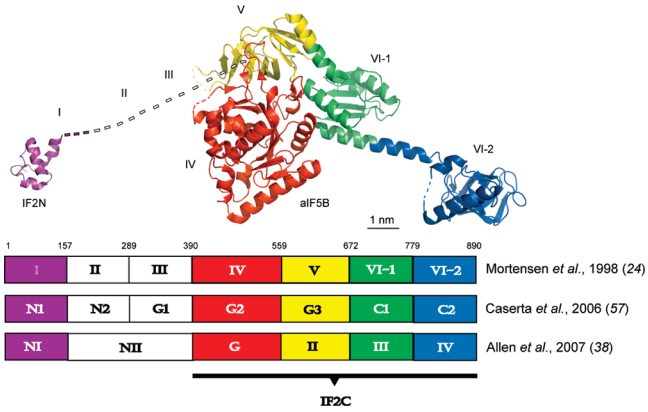

FIGURE 1: Translation initiation factor IF2 and structural homologues. The structure of *E. coli* IF2N as determined by NMR spectroscopy is shown in purple on the left [Protein Data Bank (PDB) entry 1ND9]. The structure of IF2 orthologue aIF5B from *M. thermoautotrophicum* as determined by X-ray crystallography is shown on the right (PDB entry 1G7T), and this protein corresponds to IF2C in *E. coli*. The orientation of the two structures relative to each other is unknown and is depicted here only for illustrative purposes. The domains of aIF5B are color-coded according to the chart, which displays the roughly equivalent domain nomenclatures used for IF2 in the literature (*24, 38, 57*). The color coding of the domains will be used throughout this paper.

The extremely soluble domain I of *E. coli* IF2 (*31*) contains a subdomain of approximately 50 residues termed IF2N, the structure of which has been solved by NMR spectroscopy (*28*) (Figure 1). This small subdomain shows structural homology to folds and domains of different tRNA synthetases. IF2N is internally well-ordered and tumbles in solution independently of the other domains (*27*). This is possible, because IF2N is connected with domain IV through a highly flexible linker region of significant helical content (domains I−III) (*27*).

The crystal structure of the orthologous aIF5B from archaeabacterium *M. thermoautotrophicum* has been solved (*26*) (Figure 1). This protein corresponds to domains IV−VI-2 of *E. coli* IF2 (throughout this paper termed IF2C), and sequence homology is most significant for domain IV (*32*) (Figure 2A). It consists of an eight-stranded β sheet (*26*) including a GTP-binding motif, which is found in at least four other proteins involved in translation [EF1A (formerly EF-Tu), EF2 (formerly EF-G), RF3, and SelB (*33–35*)].

The crystal structure of *M. thermoautotrophicum* aIF5B also reveals a β-barrel structure for domain V and an $\alpha\beta\alpha$ sandwich for domain VI-1, which is connected with another β-barrel structure in domain VI-2 by a long α helix (*26*). The protrusion of this single and highly exposed α helix seems puzzling in that it must confer an element of instability to the protein in addition to a high degree of flexibility.

The solution structures of C-terminal domains IF2-C1 and IF2-C2 of *Bacillus stearothermophilus* IF2 have been elucidated by NMR spectroscopy. IF2-C1 structurally resembles domain VI-1 of *M. thermoautotrophicum* IF2, although there are distinct differences that could have important functional consequences (*36*). IF2-C2 (PDB entry 1D1N) contains a β-barrel structure with homology to both domains V and VI-2 of *M. thermoautotrophicum* IF2 (*37*). The alignment of the C2 sequence to *E. coli* domain VI-2 shows great homology between these two sequences (Figure 2B); therefore, the IF2-C2 structure can be useful in the modeling of the *E. coli* IF2C solution structure.

The cryo-EM reconstruction of the *E. coli* 70S initiation complex has been obtained using nonhydrolyzable guanylyliminodiphosphate (GDPNP) to stall the progression of initiation (*29*). The complex included 30S and 50S subunits, mRNA, initiator tRNA, and all three initiation factors. The different components were located in the density map by the aid of individually determined X-ray structures. *E. coli* IF2-2 was found to be located in the intersubunit space in contact with both 30S and 50S subunits as well as initiator tRNA. However, the crystal structure of free *M. thermoautotrophicum* IF2 in its GTP form did not fit the density expected to be IF2-2 in the cryo-EM map; therefore, the domains of the IF2 crystal structure were modeled to fit this density (PDB entry 1ZO1).

Cryo-EM structures were also obtained on *T. thermophilus* 70S complexes comprising mRNA, initiator tRNA, and IF2 with either GDP or nonhydrolyzable guanylyl 5′-(β,γ-methylenediphosphonate) (GMPPCP) (*30*). In this cryo-EM map, the crystal structure of *M. thermoautotrophicum* aIF5B was also used to model IF2 and was docked onto the ribosome without domain rearrangements. Transition from the GTP analogue bound to the GDP-bound state was found to involve substantial conformational changes of both IF2 and the entire ribosome. This conformational change resulted in a reduced interaction of IF2 with the ribosome, and the contact of IF2 domain VI-2 with 16S rRNA and initiator tRNA in the P-site was lost, suggesting a "ready-to-leave" state. IF2 in this structure contacts the D loop of initiator tRNA and not the fMet moiety as found in other studies (*38*). However, because this complex is lacking IF1 and IF3, it might correspond to a later stage in initiation.

Recent studies on the ribosome interaction of eIF5B indicate that the overall orientation of eIF5B on 80S ribosomes is similar to that of IF2 on 70S ribosomes (*39*). In fact, these studies provide indirect evidence in favor of eIF5B domain arrangement as in the cryo-EM reconstruction of ribosome-bound IF2 mentioned above (*29*). Because the 80S complex investigated in this study does not contain IF1 or IF3 orthologues, the eIF5B structure would be expected to resemble the IF2 crystal structure as docked in the bacterial 70S complex also lacking IF1 and IF3. Because this is not the case, the domain arrangement performed by Allen et al. (*29*) may prove a better model for ribosome-bound IF2 than the crystal structure.

To expand the structural knowledge of *E. coli* IF2 and investigate whether the chalice-like X-ray crystallographic structure of *M. thermoautotrophicum* aIF5B is also present in *E. coli* IF2, we decided to apply small-angle X-ray scattering (SAXS) to the C-terminal part of *E. coli* IF2. When the requirements of good crystals for X-ray crystallography and low molecular mass for NMR spectroscopy can not be met by a particular protein, SAXS is a useful alternative, which allows the study of the protein in solution close to its physiological environment. Hence, the aim of this study has been to apply SAXS to domains IV−VI-2 of *E. coli* IF2 to elucidate structural information and compare this to the existing structural knowledge of IF2 and its homologues (*26*).

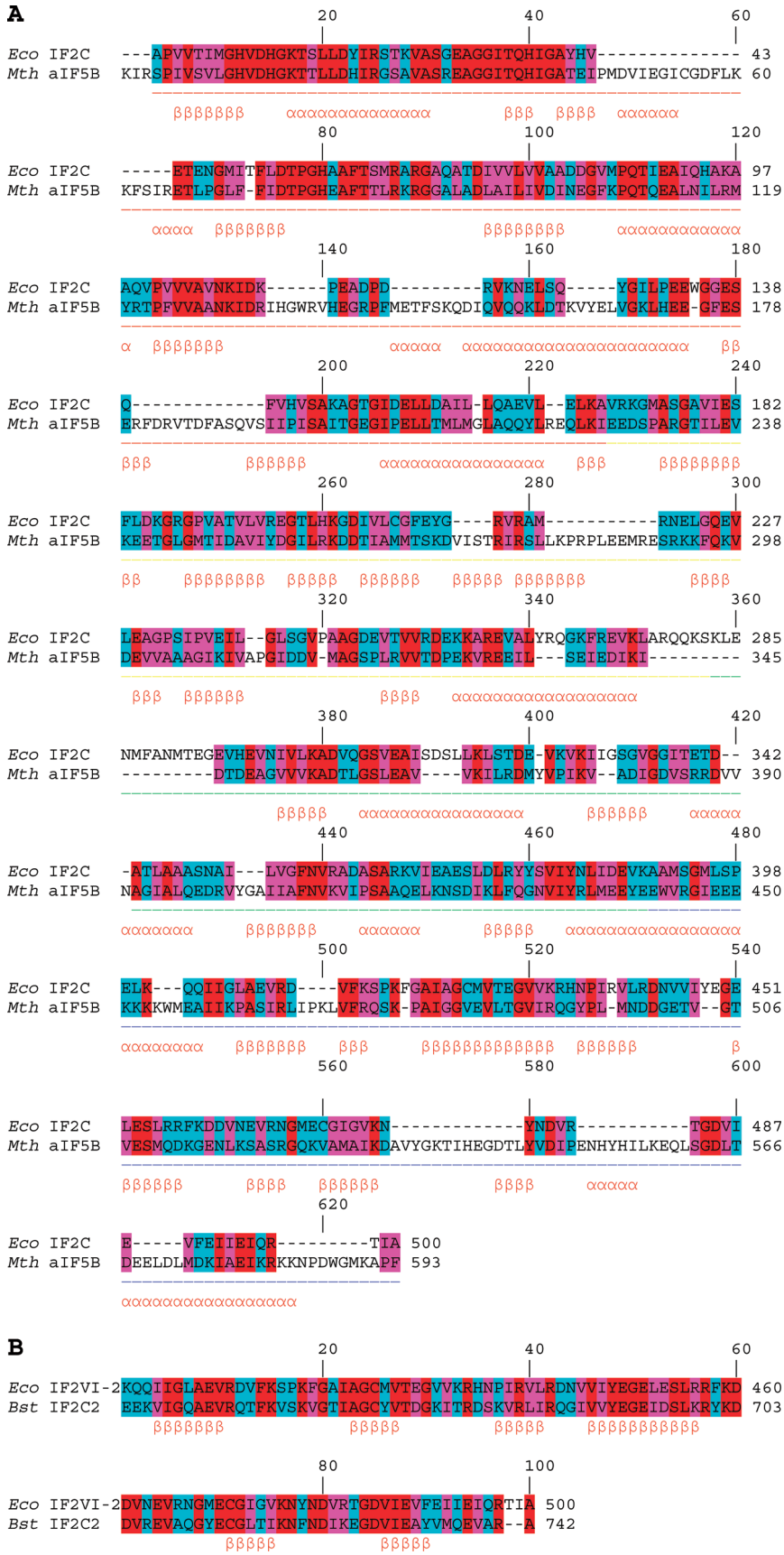

FIGURE 2: Alignment of *E. coli* IF2C to *M. thermoautotrophicum* aIF5B and *B. stearothermophilus* IF2. (A) Sequences of *E. coli* IF2C and *M. thermoautotrophicum* aIF5B were aligned and colored according to the following color code: identical residues, red; similar residues, magenta; remaining residues, cyan; gaps, not colored. The domains of IF2C are indicated below the sequences according to the color code of Figure 1. The α helices and *β* strands of aIF5B are also indicated below the sequences (*Eco*, *E. coli*; *Mth*, *M. thermoautotrophicum*). (B) Sequences of domain VI-2 of *E. coli* IF2 and the corresponding domain C2 in *B. stearothermophilus* IF2 were aligned and colored according to the color code used in A. Numbering of IF2-C2 is according to Meunier et al. (*37*). The *β* strands of the IF2-C2 NMR structure are indicated below the sequences. Numbering of IF2-VI-2 is the same as in A (*Eco*, *E. coli*; *Bst*, *B. stearothermophilus*).

EXPERIMENTAL PROCEDURES

Bacterial Strains and Plasmids. Construction of an expression vector encoding domains IV−VI-2 of IF2 with a C-terminal His tag was performed by polymerase chain reaction (PCR) amplification of the *infB* sequence using the following primers: 5′-TTTCC**ATGG**CGCCGGTTGTGACCATCATGGGTCAC (forward primer domain IV; underlined, restriction site; bold, start codon) and 5′-GGGGAATTCATTAGTGGTGGTGGTGGTGAGCAATGGTACGTTGGATCTC (reverse primer domain VI-2). Genomic DNA from *E. coli* strain JM109 was used as a template. The sequences were inserted into a pET24d vector using the restriction sites *Nco*I and *Eco*RI. The JM109 transformants obtained were subjected to sequence analysis to verify the presence of the expected sequences. The constructed plasmid pET24d-IF2C was purified and transformed into BL21(DE3) cells for protein expression.

Protein Expression and Purification. A total of 2 × 1.6 L of 2xTY medium (16 g/L peptone from casein (AppliChem), 10 g/L yeast extract, and 5 g/L NaCl) containing 50 μg/mL kanamycin was inoculated using a 2 × 16 mL pelleted and resuspended overnight culture of BL21(DE3) cells carrying the pET24d-IF2C expression vector. The cultures were incubated at 37 °C with constant shaking, until the optical density at 550 nm reached 0.8. Isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.1 mM to induce protein expression, and the cultures were incubated at 37 °C with constant shaking for 5 h. The cells were harvested by centrifugation at 9500*g* for 10 min followed by a wash in 0.9% NaCl and centrifugation at 4700*g* for 25 min. The pelleted cells were resuspended in 2 mL of buffer A [50 mM Tris-HCl at pH 7.6, 500 mM NaCl, 0.1 mM phenylmethylsulphonyl fluoride (PMSF), and 5 mM imidazole] per 1 g of cells. The cells were opened by sonication and ultracentrifuged at 250000*g* and 4 °C for 75 min. The filtered supernatant was loaded onto a nickel-charged HisBind IMAC column, and the His-tagged IF2C protein was eluted using an increasing gradient of imidazole (5−500 mM). The eluted protein was applied to a hydrophobic Phe-Sepharose-FF column and eluted with H_2O. The protein was concentrated on a Q-Sepharose-XL column and finally dialyzed against SAXS buffer [20 mM Tris-HCl at pH 7.5, 100 mM NaCl, 1 mM $MgCl_2$, 0.1 mM PMSF, and 1 mM dithiothreitol (DTT)].

The protein concentration was determined by measuring the absorbance at 280 nm.

SAXS. The experiments were performed in the laboratory-based SAXS instrument at the Department of Chemistry, University of Aarhus (*40*). Three concentrations were measured to control for possible particle aggregation and other concentration effects. All data reduction was performed using home-written software. The SAXS intensity is displayed as a function of the momentum transfer modulus q [$q = 4\pi/\lambda \sin(\theta)$, where λ is the radiation wavelength and 2θ is the scattering angle]. Pure water was used as the primary standard for absolute scale normalization (*40*). The characteristic real-space distance distribution functions $p(r)$ was determined from the scattering data using indirect Fourier transformation (*41–43*). This function corresponds to a histogram over all distances between pairs of points within the particle, and it gives direct insight into the particle shape and size. *Ab initio* model calculations were performed using GASBOR (*44, 45*). Rigid body and dummy chain modeling were performed using BUNCH (*46*). Model alignment and average was performed using SUBCOMP and DAMAVER (*47*). The calculation of theoretical solution SAXS intensity curves from atomic coordinates of protein structure and its comparison to experimental data was performed using CRYSOL (*48*). Model visualization and the generation of the model figures were performed using MOLMOL (*49*).

Dynamic Light Scattering (DLS). The experiments were performed in a light scattering instrument from ALV (Langen, Germany), which is a CGS-8F goniometer system equipped with an ALV-6010/EPP multitau digital correlator and a helium−neon diode laser (JDS Uniphase). Measurements were performed at 90°. The software delivered with the system was used for deriving the hydrodynamic (Stokes) radius using cumulant analysis. The Stokes radius is related to the diffusion of a particle in a certain liquid and is a function of the temperature and the fluid characteristics itself (Stokes−Einstein equation for a sphere in a liquid: $D = k_B T/6\pi\eta R_{Stokes}$, where D is the translational diffusion constant, k_B is Boltzmann's constant, T is the temperature, η is the liquid viscosity, and R_{Stokes} is the Stokes radius). For clarification, the radius of gyration R_g is related to the rotational properties of a particle, and it is given by the distribution of electron density around the center of mass of the particle. R_g is thus given only by particle size and shape; therefore, for a given particle in solution, the Stokes radius and R_g does not necessarily have to be equal but of the same order of magnitude.

Two sample concentrations were measured, 1 and 0.1 mg/mL, giving practically the same results. A Stokes radius of 39 ± 3 Å was found for the IF2C protein in solution. To compare this value with the three-dimensional models obtained from the SAXS analysis and also with the crystallographic structures for the homologous aIF5B protein, we have used HYDROPRO (*50*). The values obtained from the calculation with this program agree very well with the experimental value: for the SAXS model, we obtained 41 ± 1 Å; for the aIF5B protein structure 1G7T, we obtained 38.5Å; and for 1ZO1, we obtained 40 Å. Although the Stokes radii are rather similar for these protein structures (principally because of the low resolution of the DLS experiments), the SAXS data indicated that the proteins should have important differences in shape as shown in Figure 3.

RESULTS AND DISCUSSION

SAXS Data Collection and Comparisons to Orthologous Structures. The C-terminal *E. coli* IF2 fragment (IF2C) containing domains IV−VI-2 was cloned and expressed in BL21(DE3) and purified by three different column chromatography methods. After the buffer change to SAXS buffer, IF2C was subjected to SAXS analysis. For the first measurements, protein concentrations of 2.5−10 mg/mL resulted in aggregation effects, indicated by the calculated molecular weight and particle dimension (data not shown). A more controlled and stable situation was obtained at 1−2 mg/mL, because aggregation is absent at such low protein concentrations. The SAXS curves normalized by the concentration in these two cases were practically identical, indicating also

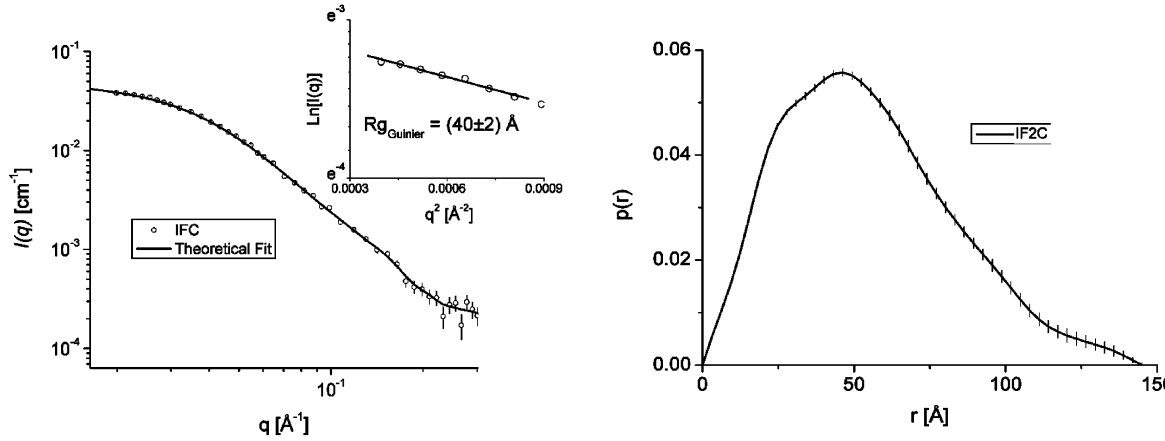

FIGURE 3: SAXS data on IF2C. To the left, experimental data for IF2C (2 mg/mL in all graphs) is plotted as a function of the scattering vector q and compared to the theoretical fit. (Inset) Guinier Plot for IF2C SAXS data. To the right, the distance distribution function $p(r)$ indicates the shape of the particle. A skewed distribution with a maximum at small distances is characteristic of elongated particles (*44*).

the absence of structure factor effects. The monodispersity of the samples were also confirmed by the Guinier plot [ln($I(q)$) versus q^2], which gave a reasonable linearity (Figure 3), and light scattering measurements.

The radius of gyration R_g, which characterizes the particle size, was derived from the SAXS data using two approaches. From the Guinier plot, we obtained a value of 40 ± 2 Å. From the indirect Fourier transformation method (see the Experimental Procedures), a value of 43 ± 1 Å was obtained, in very good agreement with the Guinier analysis. In the same analysis, the maximum dimension of the particle was estimated to be ~145 Å. For comparison, the long α helix protruding from the "cup" of the aIF5B chalice-like structure has a length of 40 Å. Furthermore, the skewed appearance of the distance distribution function $p(r)$ (Figure 3), which was derived from the data by the indirect Fourier transformation, is indicative of an elongated shape. The forward scattering $I(0)$ is related to the molecular weight of the protein, and it was determined to be 45 ± 10 kDa. The expected molecular weight is 55 kDa, which is just within the uncertainties of the measurement. Note that imprecision in the determined protein concentration and also in the calibration with the primary standard on SAXS experiments (in the present work, water at 20 °C) might contribute to the discrepancy between expected and experimentally determined molecular weights.

To compare the structure of aIF5B (PDB entry 1G7T) to the SAXS data on IF2C, it is possible to calculate the theoretical SAXS intensity from the atomic coordinates and use it to fit the experimental data (*48*). The CRYSOL program (*48*) was applied for this calculation. R_g for aIF5B was found to be 36.49 Å, with a maximum dimension of 121.2 Å; therefore, the dimension differences between the two proteins might indicate differences in the structures of the two proteins. As will be shown later, this is supported by the poor fit of the experimental SAXS data by the aIF5B atomic structure.

As mentioned in the introduction, a slightly different domain arrangement for aIF5B (PDB entry 1ZO1) on the ribosome was suggested by Allen et al. (*29*). The CRYSOL program was again used to predict the solution scattering curve of the model from the atomic coordinates. As can be seen in Figure 4, for both structures, we have a poor fit,

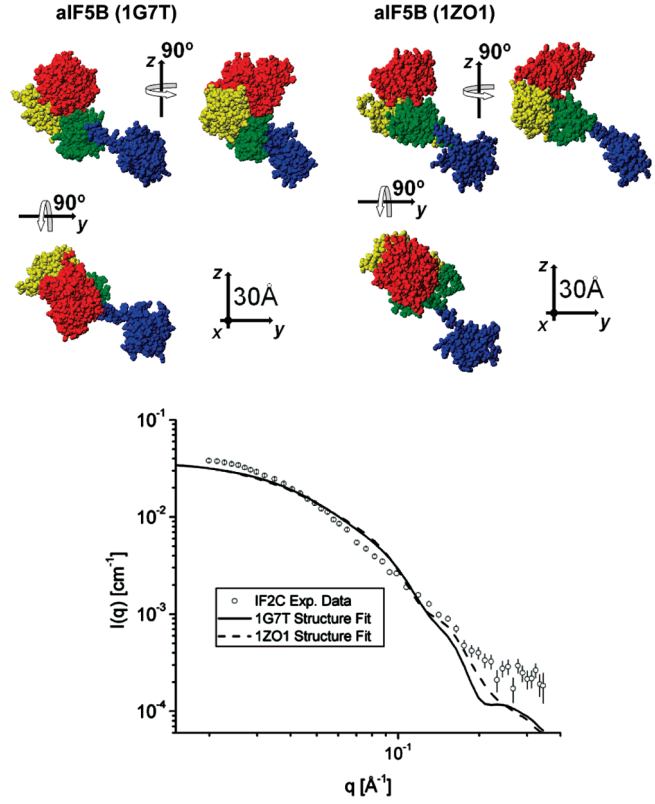

FIGURE 4: Comparison between aIF5B and IF2C SAXS data. The crystal structures of aIF5B (PDB entry 1G7T) and aIF5B as modeled onto the ribosome (PDB entry 1ZO1) are shown above. The color coding used throughout this paper is the same as in Figure 1. Below, the predicted solution scattering curves of the aIF5B crystal structure and ribosome-associated structure are compared to the experimental data on IF2C.

which gives a strong indication that the structure of IF2C is different from these two structures.

Ab Initio and Hybrid Modeling. Because the sequence and the number of amino acid residues of IF2C were known, the dummy chain approach could be applied in retrieving a low-resolution model for IF2C in solution (*44*). The protein backbone was modeled as a sequence of connected spherical units, of which the three-dimensional arrangement was optimized to produce the best fit of the SAXS data. The intrinsic low-resolution nature of SAXS prevents the retrieval

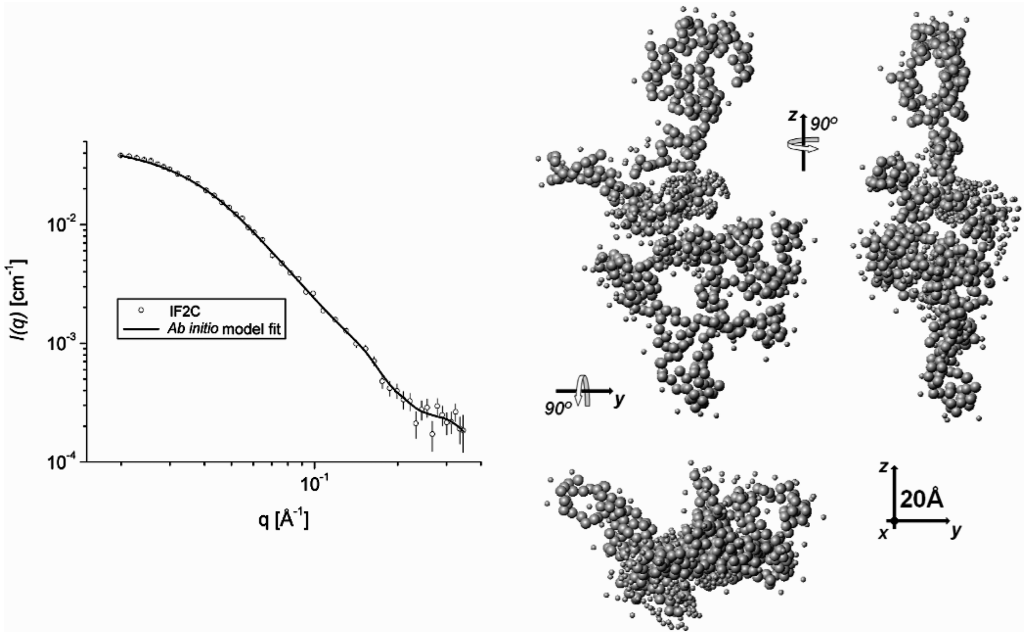

FIGURE 5: *Ab initio* modeling. To the left, the fit of the experimental data on IF2C given by the *ab initio* model. For the 10 calculated models, the fit quality is similar. To the right, the most representative structure of the 10 *ab initio* models is shown in three different perpendicular orientations. The small spheres are dummy water molecules added by the simulation program to mimic the hydration shell.

of one unique model structure, but general features of the protein shape can be recovered. A total of 10 independent models were determined, and they were averaged and compared (*45, 47*). From this averaging process, the best representative of the set was obtained, and it is shown in Figure 5 together with the fit of the experimental data. The elongated shape of the aIF5B structure suggested by the $p(r)$ function is also recognized in the *ab initio* model, which displays a relatively flat shape as seen from the rotation along the z axis. However, because of the low resolution of the model, it is difficult to indicate where the domains of IF2C should be located.

As already mentioned, the SAXS technique is intrinsically of low resolution, and the data contain only a modest amount of information. Therefore, it is necessary to use additional information (as known atomic resolution structures) on the structure to decrease the number of degrees of freedom of the problem, enabling the modeling programs to converge to more reasonable solutions.

Considering the sequence homology between *B. stearothermophilus* IF2-C2 and *E. coli* IF2 domain VI-2, the structure of domain C2 was included in the modeling. The C2 structure was used to model domain VI-2, whereas domain VI-1 of the aIF5B crystal structure was divided into two regions of homology to the IF2 domain VI-1. These three regions were allowed to move as rigid bodies with respect to domain V of aIF5B, which was fixed in the simulation. Some loops (present in domain V of IF2C but not in aIF5B) were also simulated but with the known atomic structures of domain V fixed. Because the G-domain structure of IF2 domain IV is very well-conserved in all G proteins, this entire structure from the aIF5B crystal structure was included in the modeling and allowed to move in relation to domain V. These rigid bodies were connected by dummy residues, which were used to model the missing loops. A total of 10 calculations were made using the BUNCH program (*46*), and aligning and averaging these 10 independent fitting models resulted in the model depicted in transparent gray in Figure 6. The model constituting the best representative of the average is shown in colors according to the domain structure. The scattering curve of this model gives a good fit to the SAXS data for IF2C (Figure 6), with some difference in the beginning part of the curve but within the range of the experimental error bars of the measured data. The first part of the curve is very sensitive to the presence of small fractions of aggregates in the solution, and this effect can cause the observed small discrepancy. The DLS experiments excluded the formation of a significant fraction of aggregates, but because the scattering intensity scales with the volume square of the scattering particles, even a very small fraction of aggregates (not visible on the number distribution of sizes given on DLS measurements) can produce changes at low scattering vectors. However, for $q > 0.028$ Å^{-1} we have a very good fit of the experimental data, indicating that the positioning of the domains should correspond to the protein in solution. Indeed, this is also confirmed by the averaging of the models. As can be seen in Figure 6, the averaged model (semitransparent gray spheres) preserved the positions of the IF2 domains almost in the same place, indicating that in all simulated models (despite of small differences in orientation) the domains are arranged in the same way in all reconstructions. The small differences at high scattering vectors can be related to intrinsic flexibility of the protein (which cannot be described by the applied approach, because we are using a rigid-body modeling) and also with small differences of the shape of the used atomic structures from the homologous proteins when compared to the ones present in the IF2 structure.

The model is also very similar to the one obtained from the *ab initio* modeling, as seen when superpositioned (Figure 6C). The overall shapes of the hybrid and the *ab initio* model show many similarities, and the alignment of these two structures indicates that the positions of the IF2 domains can also be seen in the *ab initio* model. Apart from small

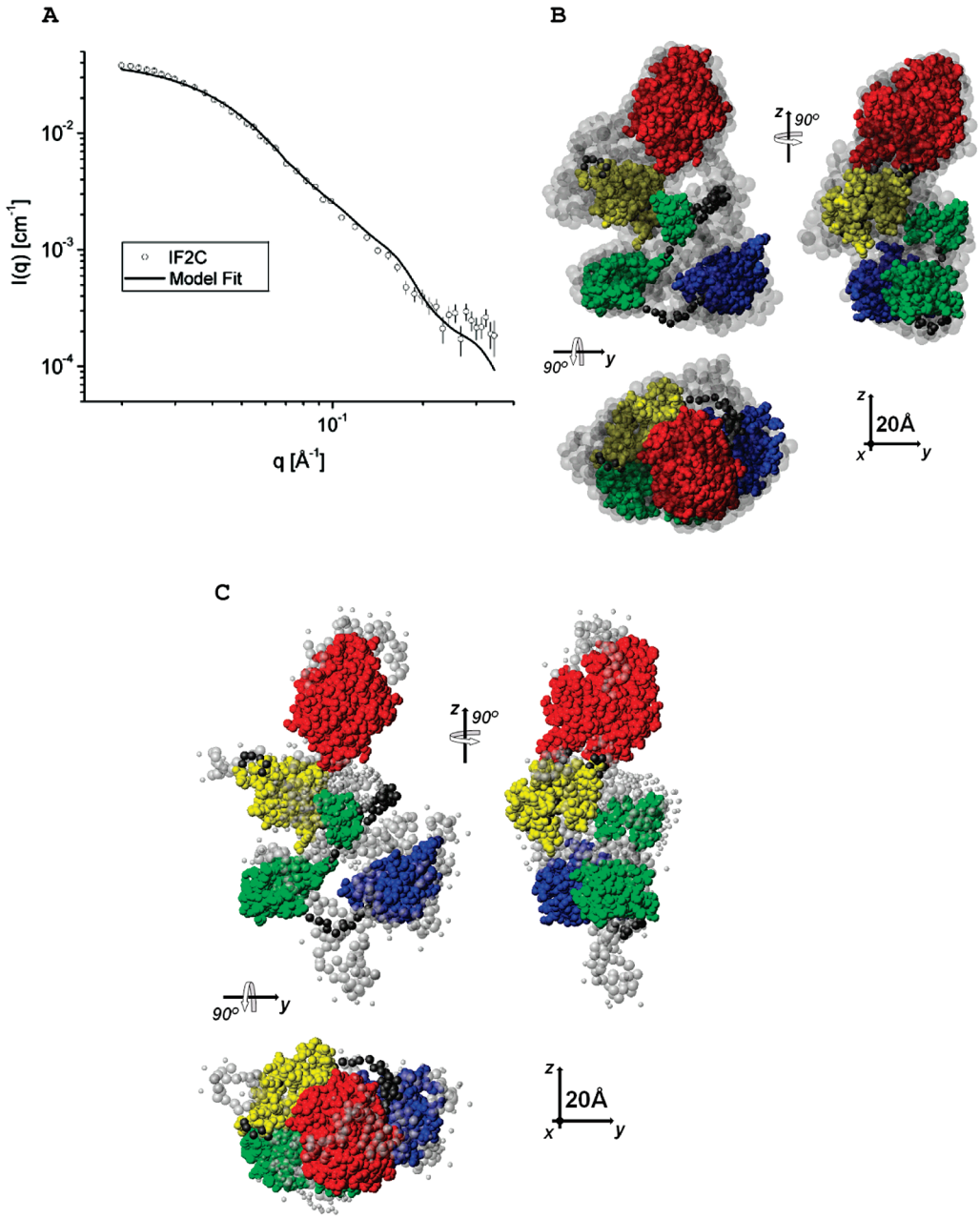

FIGURE 6: Hybrid modeling. (A) A total of 10 models were obtained by hybrid modeling using atomic resolution structures of homologous sequences. The scattering curve calculated from one of these models is shown together with the experimental data of IF2C. For the 10 calculated models, the fit quality is similar. (B) Transparent gray spheres represent an average of the 10 models. It is depicted here as a closed packing of spheres with a radius of 3 Å. The best representative of this average model is shown in colors, with the dummy residues of the flexible regions depicted in dark gray. (C) Superposition of the *ab initio* model with the hybrid model using known atomic structures. As can be seen, there is a very good agreement between the two model approaches, supporting the conclusions about the changes on the tertiary structure of the protein when compared to its homologue.

differences in the maximum size between the two model approaches (which can be induced by the presence of small fractions of aggregates on the system, which might force the *ab initio* model to larger dimensions), there is a very good agreement between the two obtained models. This agreement supports the rigid-body modeling strategy used to model the IF2C structure. Domains IV and V were assumed to be in the same conformation as in the crystal structure of aIF5B, whereas domain VI was enabled to change the disposition of its subdomains in the modeling, enabling the determination of the conformation for IF2C in solution.

To give further support to this model for IF2C, DLS measurements were performed on this protein in solution at a protein concentration of 1 and 0.1 mg/mL. DLS is a very low-resolution technique, which does not allow for the identification of structural differences at the same level as SAXS, but the method is very useful for measuring the overall protein shape. From these DLS data, a Stokes radius of 39 ± 3 Å was obtained for IF2C. The HYDROPRO (*50*) program was used to estimate the Stokes radius of the SAXS model to 41 ± 1 Å. This shows that the size of the model obtained from the SAXS data is in agreement with the one obtained by the light scattering measurements and, thus, that

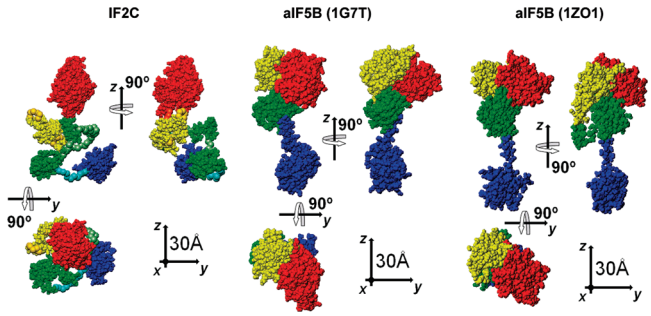

FIGURE 7: Visual comparison of the proposed structure for IF2C in solution to the crystal structure and ribosome-associated structure of aIF5B. For this visual comparison, domain V (yellow domain) was fixed in the same orientation for each structure. For IF2C, the dummy chains connecting the domains are depicted in orange, gold, light green, and cyan to reflect their connection to domains IV, V, VI-1, and VI-2, respectively.

there are no significant aggregation at the higher concentration used in the SAXS experiment.

Model of C-Terminal IF2 in Solution. The proposed structure of IF2C in solution is displayed next to the crystal structure of aIF5B and the suggested domain rearrangement of aIF5B on the ribosome (Figure 7). According to this visual comparison, domains IV and VI-1 are moved further away from domain V in the solution structure. Conversely, domain VI-2 is moved closer to the rest of the protein. When IF2-1 is trypsinated, domain VI as a whole resists trypsination and remains intact (*51*), and this does not correspond well with the highly exposed α-helical stretch observed between domains VI-1 and VI-2 in the crystal structure. The corresponding region in *E. coli* IF2-1 contains two lysines, which make this region prone to trypsination when present in solution; therefore, a closer packing of this region to the rest of domains VI-1 and VI-2 conforms to the biochemical data.

The difference between the suggested solution structure and ribosome-associated structure may be a result of induced fit. When IF2 binds to 30S through interactions of domain II and possibly domain V (to IF1), it may result in the rearrangement of domains to facilitate interactions with IF1 and initiator tRNA on the 30S subunit.

Biochemical data on IF2-1 in solution are scarce. The GTPase activity (*52*) as well as IF1 interaction (*53, 54*) are ribosome-dependent, just as a stable interaction with initiator tRNA is expected to be, because the observed binding of initiator tRNA to domain VI-2 in solution (*55*) is too weak to facilitate complex purification (*56*). However, IF2-1 in solution has been shown to possess protein chaperone activity and is capable of discriminating between unfolded and native proteins (*19*). It is possible that this activity depends upon a specific structure of IF2-1 in solution and that this flexibility in the IF2-1 structural conformation enables it to perform different functions in solution and associated with the ribosome.

Future work regarding the structure of IF2-1 involves attempts to crystallize IF2C. This will hopefully reveal whether the difference between the proposed model for IF2C and the crystal structure of aIF5B is a result of solution versus crystallization or *E. coli* versus *M. thermoautotrophicum*. Investigating the solution structure of aIF5B by SAXS analysis will add further results to this comparison. Structural knowledge of IF2-1 in complex with a permanently unfolded protein, such as reduced carboxymethyl α-lactalbumin (R-CMLA), would provide information on the particular chaperoning function of IF2-1 in solution.

REFERENCES

1. Kyrpides, N. C., and Woese, C. R. (1998) Universally conserved translation initiation factors. *Proc. Natl. Acad. Sci. U.S.A. 95*, 224–228.
2. Sørensen, H. P., Hedegaard, J., Sperling-Petersen, H. U., and Mortensen, K. K. (2001) Remarkable conservation of translation initiation factors: IF1/eIF1A and IF2/eIF5B are universally distributed phylogenetic markers. *IUBMB Life 51*, 321–327.
3. Lee, J. H., Choi, S. K., Roll-Mecak, A., Burley, S. K., and Dever, T. E. (1999) Universal conservation in translation initiation revealed by human and archaeal homologs of bacterial translation initiation factor IF2. *Proc. Natl. Acad. Sci. U.S.A. 96*, 4342–4347.
4. Laursen, B. S., de A. Steffensen, S. A., Hedegaard, J., Moreno, J. M., Mortensen, K. K., and Sperling-Petersen, H. U. (2002) Structural requirements of the mRNA for intracistronic translation initiation of the enterobacterial infB gene. *Genes Cells 7*, 901–910.
5. Nyengaard, N. R., Mortensen, K. K., Lassen, S. F., Hershey, J. W., and Sperling-Petersen, H. U. (1991) Tandem translation of *E. coli* initiation factor IF2 β: Purification and characterization *in vitro* of two active forms. *Biochem. Biophys. Res. Commun. 181*, 1572–1579.
6. Hubert, M., Nyengaard, N. R., Shazand, K., Mortensen, K. K., Lassen, S. F., Grunberg-Manago, M., and Sperling-Petersen, H. U. (1992) Tandem translation of *Bacillus subtilis* initiation factor IF2 in *E. coli*. Over-expression of infBB.su in *E. coli* and purification of α- and β-forms of IF2B.su. *FEBS Lett. 312*, 132–138.
7. Giuliodori, A. M., Brandi, A., Gualerzi, C. O., and Pon, C. L. (2004) Preferential translation of cold-shock mRNAs during cold adaptation. *RNA 10*, 265–276.
8. Bae, W., Xia, B., Inouye, M., and Severinov, K. (2000) *Escherichia coli* CspA-family RNA chaperones are transcription antiterminators. *Proc. Natl. Acad. Sci. U.S.A. 97*, 7784–7789.
9. Antoun, A., Pavlov, M. Y., Andersson, K., Tenson, T., and Ehrenberg, M. (2003) The roles of initiation factor 2 and guanosine triphosphate in initiation of protein synthesis. *EMBO J. 22*, 5593–5601.
10. Lelong, J. C., Grunberg-Manago, M., Dondon, J., Gros, D., and Gros, F. (1970) Interaction between guanosine derivatives and factors involved in the initiation of protein synthesis. *Nature 226*, 505–510.
11. Lockwood, A. H., Sarkar, P., and Maitra, U. (1972) Release of polypeptide chain initiation factor IF-2 during initiation complex formation. *Proc. Natl. Acad. Sci. U.S.A. 69*, 3602–3605.
12. La Teana, A., Pon, C. L., and Gualerzi, C. O. (1996) Late events in translation initiation. Adjustment of fMet-tRNA in the ribosomal P-site. *J. Mol. Biol. 256*, 667–675.
13. Laalami, S., Timofeev, A. V., Putzer, H., Leautey, J., and Grunberg-Manago, M. (1994) *In vivo* study of engineered G-domain mutants of *Escherichia coli* translation initiation factor IF2. *Mol. Microbiol. 11*, 293–302.
14. Larigauderie, G., Laalami, S., Nyengaard, N. R., Grunberg-Manago, M., Cenatiempo, Y., Mortensen, K. K., and Sperling-Petersen, H. U. (2000) Mutation of Thr445 and Ile500 of initiation factor 2 G-domain affects *Escherichia coli* growth rate at low temperature. *Biochimie 82*, 1091–1098.
15. Laursen, B. S., Siwanowicz, I., Larigauderie, G., Hedegaard, J., Ito, K., Nakamura, Y., Kenney, J. M., Mortensen, K. K., and Sperling-Petersen, H. U. (2003) Characterization of mutations in the GTP-binding domain of IF2 resulting in cold-sensitive growth of *Escherichia coli*. *J. Mol. Biol. 326*, 543–551.
16. Luchin, S., Putzer, H., Hershey, J. W., Cenatiempo, Y., Grunberg-Manago, M., and Laalami, S. (1999) *In vitro* study of two dominant inhibitory GTPase mutants of *Escherichia coli* translation initiation factor IF2. Direct evidence that GTP hydrolysis is necessary for factor recycling. *J. Biol. Chem. 274*, 6074–6079.
17. La Teana, A., Gualerzi, C. O., and Dahlberg, A. E. (2001) Initiation factor IF 2 binds to the α-sarcin loop and helix 89 of *Escherichia coli* 23S ribosomal RNA. *RNA 7*, 1173–1179.
18. Tomsic, J., Vitali, L. A., Daviter, T., Savelsbergh, A., Spurio, R., Striebeck, P., Wintermeyer, W., Rodnina, M. V., and Gualerzi, C. O. (2000) Late events of translation initiation in bacteria: A kinetic analysis. *EMBO J. 19*, 2127–2136.

19. Caldas, T., Laalami, S., and Richarme, G. (2000) Chaperone properties of bacterial elongation factor EF-G and initiation factor IF2. *J. Biol. Chem. 275*, 855–860.
20. Jones, P. G., VanBogelen, R. A., and Neidhardt, F. C. (1987) Induction of proteins in response to low temperature in *Escherichia coli*. *J. Bacteriol. 169*, 2092–2095.
21. Grill, S., Gualerzi, C. O., Londei, P., and Blasi, U. (2000) Selective stimulation of translation of leaderless mRNA by initiation factor 2: Evolutionary implications for translation. *EMBO J. 19*, 4101–4110.
22. Grill, S., Moll, I., Hasenohrl, D., Gualerzi, C. O., and Blasi, U. (2001) Modulation of ribosomal recruitment to 5′-terminal start codons by translation initiation factors IF2 and IF3. *FEBS Lett. 495*, 167–171.
23. Moll, I., Grill, S., Gualerzi, C. O., and Blasi, U. (2002) Leaderless mRNAs in bacteria: Surprises in ribosomal recruitment and translational control. *Mol. Microbiol. 43*, 239–246.
24. Mortensen, K. K., Kildsgaard, J., Moreno, J. M., Steffensen, S. A., Egebjerg, J., and Sperling-Petersen, H. U. (1998) A six-domain structural model for *Escherichia coli* translation initiation factor IF2. Characterisation of twelve surface epitopes. *Biochem. Mol. Biol. Int. 46*, 1027–1041.
25. Steffensen, S. A., Poulsen, A. B., Mortensen, K. K., and Sperling-Petersen, H. U. (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.
26. 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.
27. Laursen, B. S., Kjaergaard, 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.
28. Laursen, B. S., Mortensen, K. K., Sperling-Petersen, H. U., and Hoffman, D. W. (2003) A conserved structural motif at the N-terminus of bacterial translation initiation factor IF2. *J. Biol. Chem. 278*, 16320–16328.
29. Allen, G. S., Zavialov, A., Gursky, R., Ehrenberg, M., and Frank, J. (2005) The cryo-EM structure of a translation initiation complex from *Escherichia coli*. *Cell 121*, 703–712.
30. Myasnikov, A. G., Marzi, S., Simonetti, A., Giuliodori, A. M., Gualerzi, C. O., Yusupova, G., Yusupov, M., and Klaholz, B. P. (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.
31. Sorensen, H. P., Sperling-Petersen, H. U., and Mortensen, K. K. (2003) A favorable solubility partner for the recombinant expression of streptavidin. *Protein Expression Purif. 32*, 252–259.
32. Laursen, B. S., Sorensen, H. P., Mortensen, K. K., and Sperling-Petersen, H. U. (2005) Initiation of protein synthesis in bacteria. *Microbiol. Mol. Biol. Rev. 69*, 101–123.
33. Kjeldgaard, M., Nyborg, J., and Clark, B. F. (1996) The GTP binding motif: Variations on a theme. *FASEB J. 10*, 1347–1368.
34. Rodnina, M. V., Stark, H., Savelsbergh, A., Wieden, H. J., Mohr, D., Matassova, N. B., Peske, F., Daviter, T., Gualerzi, C. O., and Wintermeyer, W. (2000) GTPases mechanisms and functions of translation factors on the ribosome. *Biol. Chem. 381*, 377–387.
35. Vetter, I. R., and Wittinghofer, A. (2001) The guanine nucleotide-binding switch in three dimensions. *Science 294*, 1299–1304.
36. Wienk, H., Tomaselli, S., Bernard, C., Spurio, R., Picone, D., Gualerzi, C. O., and Boelens, R. (2005) Solution structure of the C1-subdomain of *Bacillus stearothermophilus* translation initiation factor IF2. *Protein Sci. 14*, 2461–2468.
37. Meunier, S., Spurio, R., Czisch, M., Wechselberger, R., Guenneugues, 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.
38. Allen, G. S., and Frank, J. (2007) Structural insights on the translation initiation complex: Ghosts of a universal initiation complex. *Mol. Microbiol. 63*, 941–950.
39. Unbehaun, A., Marintchev, A., Lomakin, I. B., Didenko, T., Wagner, G., Hellen, C. U., and Pestova, T. V. (2007) Position of eukaryotic initiation factor eIF5B on the 80S ribosome mapped by directed hydroxyl radical probing. *EMBO J. 26*, 3109–3123.
40. Pedersen, J. S. (2004) A flux- and background-optimized version of the NanoSTAR small-angle X-ray scattering camera for solution scattering. *J. Appl. Crystallogr. 37*, 369–380.
41. Glatter, O. (1977) A new method for the evaluation of small-angle scattering data. *J. Appl. Crystallogr. 10*, 415–421.
42. Pedersen, J. S., Hansen, S., and Bauer, R. (1994) The aggregation behavior of zinc-free insulin studied by small-angle neutron scattering. *Eur. Biophys. J. 22*, 379–389.
43. Svergun, D. I., Semenyuk, A. V., and Feigin, L. A. 1988) Small-angle-scattering-data treatment by the regularization method. *Acta Crystallogr., Sect. A: Found. Crystallogr. 44*, 244–250.
44. Svergun, D. I., and Koch, M. H. J. (2003) Small-angle scattering studies of biological macromolecules in solution. *Rep. Prog. Phys. 66*, 1735–1782.
45. Volkov, V. V., and Svergun, D. I. (2003) Uniqueness of *ab initio* shape determination in small-angle scattering. *J. Appl. Crystallogr. 36*, 860–864.
46. Petoukhov, M. V., and Svergun, D. I. (2005) Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophys. J. 89*, 1237–1250.
47. Kozin, M. B., and Svergun, D. I. (2001) Automated matching of high- and low-resolution structural models. *J. Appl. Crystallogr. 34*, 33–41.
48. Svergun, D., Barberato, C., and Koch, M. H. J. (1995) CRYSOL—A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr. 128*, 768–773.
49. Koradi, R., Billeter, M., and Wuthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graphics 14*, 51–55, 29−32.
50. Garcia de la Torre, J., Huertas, M. L., and Carrasco, B. (2000) Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys. J. 78*, 719–730.
51. Moreno, J. M., Kildsgaard, J., Siwanowicz, I., Mortensen, K. K., and Sperling-Petersen, H. U. (1998) Binding of *Escherichia coli* initiation factor IF2 to 30S ribosomal subunits: A functional role for the N-terminus of the factor. *Biochem. Biophys. Res. Commun. 252*, 465–471.
52. Kolakofsky, D., Dewey, K. F., Hershey, J. W., and Thach, R. E. (1968) Guanosine 5′-triphosphatase activity of initiation factor f2. *Proc. Natl. Acad. Sci. U.S.A. 61*, 1066–1070.
53. Boileau, G., Butler, P., Hershey, J. W., and Traut, R. R. (1983) Direct cross-links between initiation factors 1, 2, and 3 and ribosomal proteins promoted by 2-iminothiolane. *Biochemistry 22*, 3162–3170.
54. Marintchev, A., Kolupaeva, V. G., Pestova, T. V., and Wagner, G. (2003) Mapping the binding interface between human eukaryotic initiation factors 1A and 5B: A new interaction between old partners. *Proc. Natl. Acad. Sci. U.S.A. 100*, 1535–1540.
55. Petersen, H. U., Roll, T., Grunberg-Manago, M., and Clark, B. F. (1979) Specific interaction of initiation factor IF2 of *E. coli* with formylmethionyl-tRNA $_{f}^{Met}$. *Biochem. Biophys. Res. Commun. 91*, 1068–1074.
56. Wu, X. Q., and RajBhandary, U. L. (1997) Effect of the amino acid attached to *Escherichia coli* initiator tRNA on its affinity for the initiation factor IF2 and on the IF2 dependence of its binding to the ribosome. *J. Biol. Chem. 272*, 1891–1895.
57. Caserta, E., Tomsic, J., Spurio, R., La Teana, A., Pon, C. L., and Gualerzi, C. O. (2006) Translation initiation factor IF2 interacts with the 30 S ribosomal subunit via two separate binding sites. *J. Mol. Biol. 362*, 787–799.

BI8000598