BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Structural Transitions of Translation Initiation Factor IF2 upon GDPNP and GDP Binding in Solution

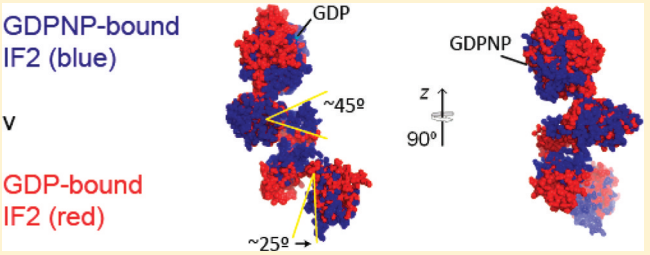
Louise Carøe Vohlander Rasmussen,[†] Cristiano Luis Pinto Oliveira,[‡,§] Jan Skov Pedersen,[‡] Hans Uffe Sperling-Petersen,[†] and Kim Kusk Mortensen*[,†]

[†]Department of Molecular Biology, Aarhus University, Gustav Wieds Vej 10, DK-8000 Aarhus C, Denmark

[‡]Department of Chemistry, Centre for mRNP Biogenesis and Metabolism, and iNANO Interdisciplinary Nanoscience Center, Aarhus University, Langelandsgade 140, DK-8000 Aarhus C, Denmark

[§]Institute of Physics, University of São Paulo, Rua do Matão Travessa R, São Paulo, Brazil

ABSTRACT: Three protein factors ensure rapid and accurate initiation of translation in bacteria. Translation initiation factor IF2 is a ribosome-dependent GTPase, which is important for correct positioning of initiator tRNA on the 30S subunit as well as ribosomal subunit joining. The solution structure of the free C-terminal part of IF2 (IF2C, comprising domains IV to VI-2) was previously determined by small-angle X-ray scattering (SAXS) [Rasmussen, L. C., et al. (2008) *Biochemistry 47*, 5590−5598]. In this study, adding GDP or nonhydrolyzable GTP analogue GDPNP to the protein in solution caused structural changes in the protein, in agreement with recent data determined via isothermal titration calorimetry [Hauryliuk, V., et al. (2009) *J. Mol. Biol. 394*, 621−626]. The $p(r)$ function indicated an elongated conformation supported by radii of gyration of 40.1 and 44.9 Å and maximum dimensions of ∼125 and ∼150 Å for IF2C with GDPNP and GDP, respectively. The SAXS data were used to model the structure of IF2C bound to either GDPNP or GDP. The structural transitions of IF2C upon GDPNP binding and following nucleotide hydrolysis support the concept of cofactor-dependent conformational switching rather than the classical model for GTPase activity.



Protein synthesis in bacteria is initiated on the ribosomes by the combined action of three translation initiation factors, one of them being the essential translation initiation factor IF2. In this process, the AUG start codon of an mRNA is positioned in the peptidyl site (P-site) of the 30S ribosomal subunit for in-frame translation. IF2 ensures correct binding of initiator tRNA and accelerates association of the 50S ribosomal subunit to the 30S initiation complex.[1] IF2 has a ribosome-dependent GTPase activity, and when the subunits are joined, this is contacted by the GTPase-activating center of the 50S subunit[2,3] and GTP is hydrolyzed to GDP by IF2. This GTP hydrolysis drives the release of IF2 from the 70S complex[2] and the transition to an elongation-competent conformation.[3]

Escherichia coli IF2 can be divided into six domains, where the C-terminal domain may be further divided into two subdomains. The domain breakdown is based on interspecies homology[4] (Figure 1). The C-terminal part (domains IV to VI-2) is highly conserved among species, whereas the N-terminal region (domains I−III) varies in both length and amino acid sequence but shows intraspecies homology in being completely conserved within the strains of *E. coli*.[5] Domain IV is the GTP-binding domain also known as the G-domain, and switch I and switch II regions in the nucleotide binding site allow discrimination between GTP and GDP.

Orthologues of IF2 are found in both eukaryotes [eukaryotic initiation factor 5B (eIF5B)] and archaeabacteria [archaeal initiation factor 5B (aIF5B)],[8−10] and aIF5B from *Methanobacterium thermoautotrophicum* (*M. therm.*) is the only full-length IF2 orthologue crystallized to date.[6] This archaeal orthologue and IF2 both facilitate ribosomal subunit joining and exhibit GTPase activity, but only IF2 is involved in initiator tRNA adjustment.[7] The aIF5B protein corresponds to domains IV to VI-2 of *E. coli* IF2, because it lacks all of the N-terminal domains (I−III). The structure was determined both for the protein without ligand and for the protein with GDP and GDPNP (a nonhydrolyzable GTP analogue) ligands. The protein displays a chalicelike structure, and a modest conformational change in domain IV (the G-domain) results in a minor ∼5 Å movement of the C-terminal domain upon nucleotide binding (domain IV is called domain I in ref 6 because of the use of a different domain nomenclature). However, these crystals are in the same space group and have the same crystal packing, so the extent of conformational difference might be larger in solution.

NMR studies of the isolated G-domain of *Bacillus stearothermophilus* (*B. stearo.*) IF2 (corresponding to domain IV of *E. coli* IF2), either free or in complex with GDP, also revealed changes in structural conformation upon nucleotide

Received: June 19, 2011
Revised: September 18, 2011
Published: October 11, 2011

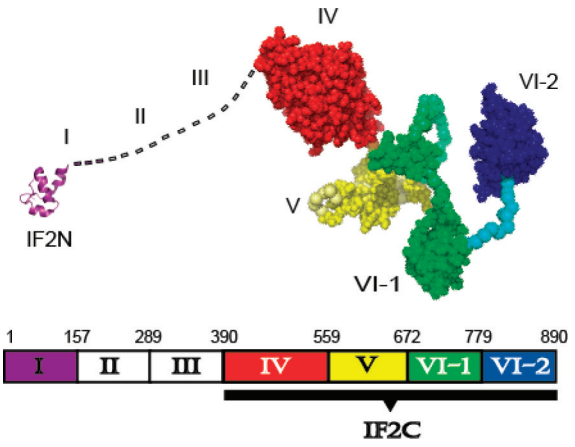


Figure 1. *E. coli* translation initiation factor IF2. The structure of *E. coli* IF2C shown next to the structure of the first 50 residues of the protein (IF2N). The structure of IF2C in solution was determined using SAXS.[16] The structure of IF2N was determined by NMR spectroscopy and is colored purple on the left (PDB entry 1ND9). The orientation of the two structures relative to each other is unknown and is depicted here only for illustrative purposes. The domains of IF2 are color-coded according to the chart, which displays the domain nomenclature. This color-coding of domains is used throughout this work.

binding.[8] The GTP binding site of this domain was also shown to bind the alarmone, ppGpp, with an affinity similar to that of GTP.[9] The level of this alarmone increases during conditions of nutritional stress, while GTP levels decrease. When ppGpp binds to IF2 domain IV, it inhibits translation, thus mediating a response to the shortage of nutrients. This has led to the proposal that domain IV is a metabolic sensor and a translational regulator.

A cryo-EM reconstruction of the *E. coli* 70S initiation complex, including mRNA, initiator tRNA, and all three initiation factors (though IF3 was not visible in the final structure), was obtained using nonhydrolyzable GDPNP to stall progression of initiation.[10] Individually determined X-ray structures of the components were fitted into the corresponding density in the cryo-EM map. Because only one crystal structure was available for IF2 orthologues at present, the crystal structure of free *M. therm.* aIF5B in its GTP form was used, even though this protein is from a different organism. The crystal structure did not fit the density expected to be IF2 in the cryo-EM map, so the domains of the aIF5B crystal structure were reorganized to fit this density. IF2 in this complex was located in the intersubunit space contacting both 30S and 50S subunits as well as initiator tRNA.

Another cryo-EM study of *Thermus thermophilus* 70S initiation complexes, including mRNA, initiator tRNA, and only one initiation factor, IF2, was conducted with either GDP or nonhydrolyzable GTP analogue GMPPCP.[11] Here the crystal structure of *M. therm.* IF2 was docked without domain rearrangements into the IF2 electron density. The transition from the GMPPCP-bound to GDP-bound state was found to involve substantial conformational changes in both IF2 and the entire ribosome. This conformational change resulted in weakened interaction of IF2 with the ribosome, and because this complex is lacking IF1 and IF3, it might correspond to a late stage in initiation, post-GTP hydrolysis and just prior to IF2 release.

A concept of cofactor-dependent conformational switching of GTPases was developed[12] to account for the observations that some GTPases remain in an inactive (D) conformation even when GDP is replaced with GTP, while adopting an active (T) conformation in complex with the ribosome or ligands. This could explain the modest conformational changes observed between GDP- and GDPNP-bound aIF5B and the domain rearrangements of aIF5B observed in complex with the ribosome. The theory holds that when the equilibrium between the D and T conformations of the ligand-free GTPase is greatly shifted toward the D conformation, exchanging GDP with GTP may fail to significantly drive the GTPase to its T conformation. However, in the presence of a cofactor, such as tRNA, GDP to GTP exchange may readily switch the GTPase into its active T form.

The structure of *E. coli* IF2 domains IV to VI-2 (denoted IF2C, corresponds to domains G and II–IV in alternative domain nomenclature[13]) in solution is distinct from the chalicelike crystal structure of its orthologue, aIF5B.[14] IF2C is slightly more compact, rendering the link between subdomains VI-1 and VI-2 less accessible. This is in agreement with biochemical data showing that domain VI as a whole resists trypsinization. If this link were easily accessible as in aIF5B, it would be trypsinated as a result of the two lysines located here.

Like aIF5B, IF2 has also been shown to change conformation upon ligand binding.[15] Isothermal titration calorimetry showed a large decrease in heat capacity upon binding of either GDP or GTP, suggesting conformational changes reducing the solvent-accessible area of nonpolar amino acids. The structural differences between the protein without ligand and the GDP form were much larger than the differences between the GDP and GTP forms. This may be a result of the lack of a cofactor to drive the GTPase to its active conformation.

The structure of IF2C in solution and in the absence of any ligand has been determined previously using small-angle X-ray scattering.[14] The aim of this study has been to apply the same technique in analyzing the structure of IF2C bound to either GDP or GDPNP and to compare this to existing data on the structural transitions involved in ligand binding.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification. IF2C protein was expressed from the pET24d-IF2C vector[14] in BL21(DE3) cells and purified essentially as described in ref 14. Briefly, the cells were cultured in 2×TY medium [16 g/L peptone from casein (AppliChem), 10 g/L yeast extract, and 5 g/L NaCl] containing 50 μg/mL kanamycin, and protein expression was induced by addition of IPTG. IF2C was purified from the filtered supernatant of the opened cells by IMAC and IEC. Concentration and a change to SAXS buffer [20 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1 mM $MgCl_2$, 0.1 mM PMSF, and 1 mM DTT] were performed using a Vivaspin 20 filter unit (Sartorius).

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

GDPNP and GDP (Sigma-Aldrich) were added in molar excess to IF2C (50 μM GDPNP or GDP to 36.3 μM IF2C), resulting in >90% binding efficiency (based on the K_d for full-length IF2 and GDP[15]).

Small-Angle X-ray Scattering (SAXS). SAXS data were recorded for 2 mg/mL IF2C with either ligand at room temperature with the laboratory-based SAXS instrument of the Department of Chemistry of Aarhus University.[16] The samples were thermostated and kept in reusable home-built quartz capillaries. Home-written software [SUPERSAXS package (C.

L. P. Oliveira and J. S. Pedersen, unpublished work)] was used for all data reduction. This program is available by request. 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]. The q range was from 0.0085 to 0.35 Å^{-1}. Pure water was used as primary standard for absolute scale normalization.[16] Indirect Fourier transformation of the scattering data resulted in the characteristic real-space distance distribution function $p(r)$.[17−19] 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. The SAXS data were analyzed using data analysis software from the ATSAS program suite. Indirect Fourier transformation was performed using GNOM,[20] and GASBOR[21,22] was used for ab initio model calculations. Rigid body and dummy chain modeling was performed using BUNCH.[23] The IF2C sequence was modeled by five rigid body subdomains connected by flexible linkers. The subdomains were selected on the basis of sequence alignments among *E. coli* IF2C, *M. therm.* aIF5B, and *B. stearo.* IF2-C2, and no rigid body subdomains overlap the domain borders shown in Figure 1. The sequence input for BUNCH was a hybrid of *E. coli* IF2C and the sequences of the rigid body subdomains. Subdomains and homology to the corresponding *E. coli* IF2C sequence are shown in Table 1. Model alignment and averaging

Table 1. Structures Used as Input for BUNCH Modeling

rigid body	part of IF2C modeled	source	residues	sequence identity (%)
1	1−169	*M. therm.* aIF5B	1−226	33
2	175−265	*M. therm.* aIF5B	232−338	26
3	295−323	*M. therm.* aIF5B	347−371	35
4	324−380	*M. therm.* aIF5B	372−433	20
5	404−500	*B. stearo.* IF2-C2	5−99	50

were performed using SUPCOMB[24] and DAMAVER.[25] Calculation of theoretical solution SAXS intensity curves from atomic coordinates of protein structure and its comparison with experimental data were performed using CRYSOL.[26] Model visualization and generation of figures were conducted with PyMOL.[27]

RESULTS AND DISCUSSION

SAXS Data. The C-terminal *E. coli* IF2 fragment (IF2C) containing domains IV to VI-2 was purified by column chromatography and concentrated in SAXS buffer to 2 mg/mL. GDPNP or GDP was added in molar excess prior to SAXS analysis. No aggregation was observed, and the monodispersity of each sample was confirmed both by the Guinier plot {ln[$I(q)$] vs q^2}, which gave a good linearity (Figure 2), and by the calculation of the protein molecular weights, which provided values very close to the expected ones for each case.

The radius of gyration (R_g) for a protein characterizes particle size and can be derived from SAXS data using two approaches. From the Guinier plot, we obtained values of 39.0 ± 1.5 Å for the GDPNP form and 41.5 ± 1.0 Å for the GDP form. From the indirect Fourier transformation performed in GNOM, values of 40.1 ± 0.4 Å for the GDPNP form and 44.9

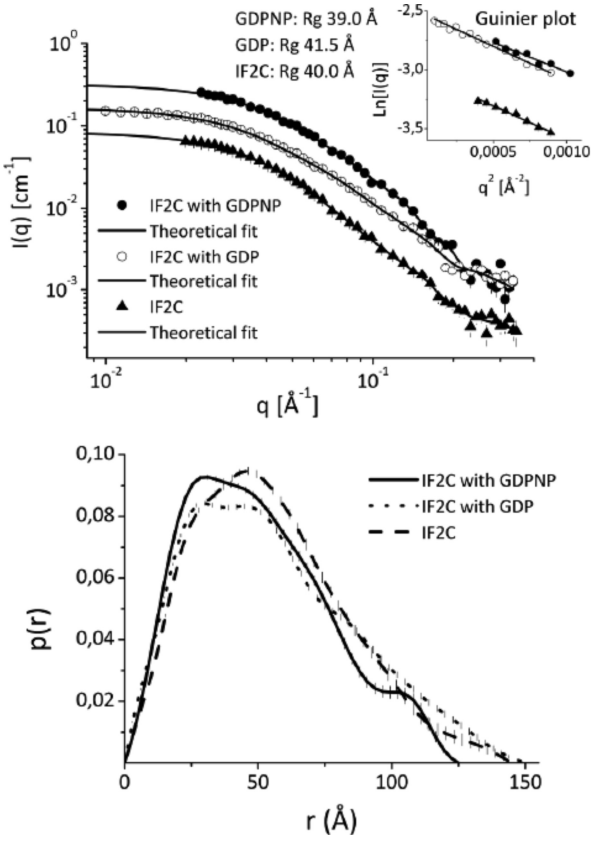


Figure 2. SAXS data for IF2C with GDP and GDPNP. The top panel shows experimental data for IF2C (2 mg/mL in all graphs) either alone or with GDP or GDPNP plotted as a function of scattering vector q and compared to the theoretical fit. The data for IF2C alone are from ref 14. Two of the curves are shifted by a constant factor for the sake of clarity. The inset shows Guinier plots for IF2C alone or IF2C with GDP or GDPNP. The bottom panel shows the distance distribution functions $p(r)$ that indicate the shape of the particle. A skewed distribution with a maximum at small distances is characteristic of elongated particles.[21] The data for IF2C alone are from ref 14.

± 0.4 Å for the GDP form were obtained, in good agreement with the Guinier analysis. In all cases, the general validity condition for the Guinier analysis ($R_g q_{max} \sim 1$) was fulfilled. The R_g for IF2C without ligand has previously been shown to be 40 ± 2 Å by Guinier analysis,[16] which is comparable to the R_g values obtained for the same protein with either GDPNP or GDP.

The maximum dimension D_{max} of the particle was estimated to be ∼125 and ∼150 Å for the GDPNP and GDP forms, respectively. Comparing R_g to the maximum particle dimension, we find the protein seems to have an elongated shape in both forms. The distance distribution function $p(r)$ (Figure 2) also resulting from indirect Fourier transformation of the SAXS data has a skewed appearance in both cases, which also indicates an elongated shape. Actually, the peak of the $p(r)$ functions for both GDPNP- and GDP-bound IF2C is more to the left than the peak of the ligand-free IF2C $p(r)$ function, indicating that binding of either nucleotide elongates the shape of IF2C further. Because the data are normalized to the absolute scale, the extrapolated forward intensity at $q = 0$, $I(0)$, is related to the molecular mass of the protein, which was determined to be 54.5 kDa (GDPNP form) or 56.4 kDa (GDP

dx.doi.org/10.1021/bi200938q | *Biochemistry* 2011, 50, 9779−9787

form), which is in very good agreement with the expected molecular mass of 55 kDa (Table 2).

Table 2. Values of SAXS Parameters

	IF2C	IF2C–GDPNP	IF2C–GDP
Guinier R_g (Å)	40 ± 2	39.0 ± 1.5	41.5 ± 1.0
IFT R_g (Å)	43 ± 1	40.1 ± 0.4	44.9 ± 0.4
D_{max} (Å)	~145	~125	~150
molecular mass (kDa)	45 ± 10	54.5	56.4

The SAXS intensity curve and distance distribution function $p(r)$ for IF2C without ligand are shown in Figure 2 for comparison (data from ref 14). It is obvious that the addition of either GDPNP or GDP results in distinct structural changes in IF2C. Interestingly, the $p(r)$ function for the GDPNP form has a small "shoulder" at large distances, which is not present on the other curves, indicating a correlation between distances within the protein structure. Another interesting point is the fact that the SAXS curve for IF2C with GDPNP is slightly different from the curves of IF2C alone and IF2C with GDP with large scattering angles. This is a consequence of an increasing flexibility for IF2C with GDPNP, which was confirmed by Kratky plot analysis (data not shown).

Ab Initio and Hybrid Modeling. Ab initio models can be generated without prior knowledge of the protein secondary structure. GASBOR implements a simulated annealing approach, in which an ensemble of dummy residues is molded into a model fitting the SAXS data and constrained by a sphere of diameter D_{max}, as determined by GNOM. The chain of dummy residues corresponds to the sequence and the number of amino acid residues of the analyzed protein, which is therefore required as input to model the protein backbone.[21] The intrinsic low-resolution nature of SAXS prevents the retrieval of one unique model structure, but general features of the protein shape can be recovered. Twenty independent models were generated, and the 10 best models (lowest χ^2) were averaged using DAMAVER.[24,25] In Figure 3, the best ab initio model for each form of IF2C is shown in colors and superpositioned onto the average of the 10 best models by SUPCOMB. This superposition shows that there is a good correlation between the generated models and that the best ab initio model is a good representation of IF2C in each form. Both ab initio models have an elongated shape as suggested by the $p(r)$ functions, but there are also differences. The ab initio model of IF2C with GDPNP is more compact around its longitudinal axis with two smaller protrusions giving it a crosslike structure in the y–z plane, whereas the ab initio model of IF2C with GDP is a bit more elongated and has an almost perpendicular protrusion from its longitudinal axis giving it a Y structure instead. Especially the GDPNP form has a relatively flat shape as seen from the rotation along the z-axis. The graphs in Figure 3 show a good fit of each ab initio model to the corresponding SAXS intensity data.

To go one step further in the modeling procedure, it is necessary to include additional information. Because known atomic-resolution structures for regions of the protein are available, they can be used to confine the translational degrees of freedom and allow the modeling programs to converge to more reasonable solutions. For modular proteins, high-resolution structures of domains can be assembled to give the

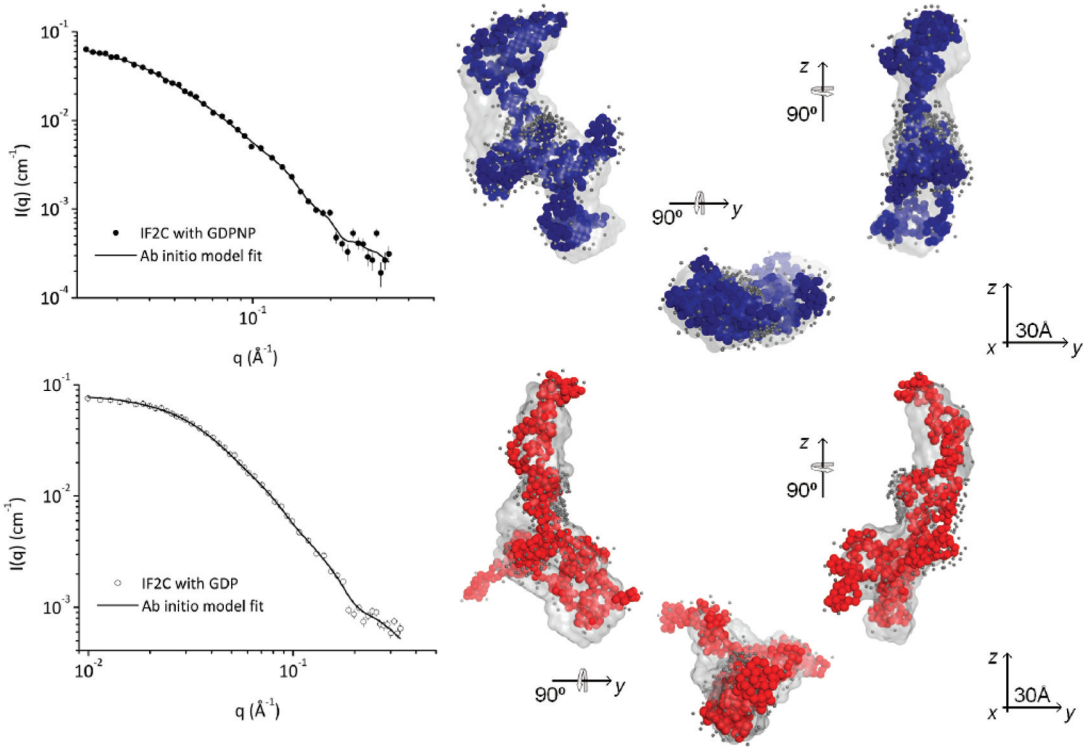


Figure 3. Ab initio modeling. Ab initio model fits to the experimental data were generated using GASBOR. The best representative of the 20 calculated models is shown in colors on top of an average of the 10 best models and in three different perpendicular orientations. The small gray spheres are dummy water molecules added by the simulation program to mimic the hydration shell. The graph shows the fit of the best ab initio model (generated by GASBOR) to the experimental data: (top) IF2C with GDPNP and (bottom) IF2C with GDP.

dx.doi.org/10.1021/bi200938q | *Biochemistry* 2011, 50, 9779–9787

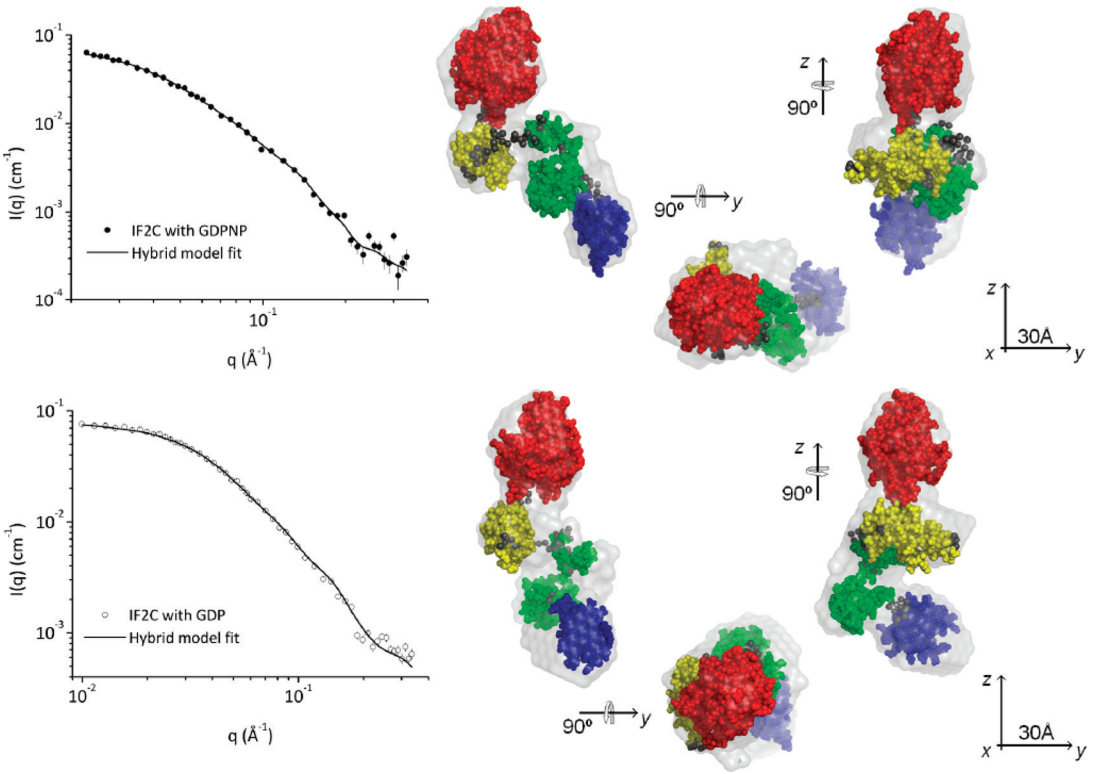


Figure 4. Hybrid modeling. Atomic-resolution structures of homologous sequences were connected by flexible dummy residues and modeled to fit the experimental data. Twenty hybrid models were generated using BUNCH. The best representative of the calculated models is shown in colors on top of an average of the 10 best models and in three different perpendicular orientations. The domains are highlighted according to the color code of Figure 1 with dummy residues colored gray. GDPNP and GDP (cyan) are marked where visible on the surface of the models. The graph shows the fit of the best hybrid model to the experimental data: (top) IF2C with GDPNP and (bottom) IF2C with GDP (PDB files of the best BUNCH models can be provided upon request presupposing proper citation).

tertiary structure of the protein. As the conserved residues of IF2 map to the core regions of each of the four globular domains of the aIF5B crystal structure,[6] some of these structures were included in the rigid body modeling of IF2. Because the G-domain structure of IF2 domain IV is very well conserved in all GTPases, this entire structure from the aIF5B crystal structure was applied. Domain V of IF2C contains loops not present in domain V of aIF5B, so these missing loops were simulated by dummy residues but with the known atomic structures of domain V fixed. Domain VI-1 of aIF5B was divided into two regions of homology to the corresponding IF2 domain VI-1 and also included.

Domain C2 from *B. stearo.* IF2 shares a high degree of sequence homology with *E. coli* IF2 domain VI-2, so the C2 NMR structure was used as input to model domain VI-2. All these separate structures of atomic resolution were connected by dummy residues and allowed to move as rigid bodies with respect to domain V of aIF5B, which was fixed in the simulation. Twenty independent calculations were conducted for both the GDPNP-bound and GDP-bound IF2C using BUNCH,[23] and aligning and averaging the 10 best hybrid models for each form of IF2C resulted in the structures depicted in transparent gray in Figure 4. The model constituting the best representative of the average is shown in colors according to domain structure and in three perpendicular orientations. The scattering curves of the GDPNP- and GDP-bound models give a very good fit to the experimental SAXS data (Figure 4), indicating that the positioning of the domains should correspond to the protein in solution. Small differences observed at high scattering vectors can be related to the intrinsic flexibility of the protein, which cannot be described by the rigid body modeling approach, as well as to small differences in shape between the homologous structures used for modeling and the actual IF2C structure. The averaging of the models shows a good correlation between individual models as seen from the very good agreement between the best model and the average. The positions of the individual domains are strongly preserved, indicating that the domains are arranged in the same way in all reconstructions despite small differences in orientation.

To further support the use of the rigid body modeling strategy, we superpositioned the hybrid models onto the GASBOR models for comparison (Figure 5). The overall shapes of the hybrid and the ab initio models show many similarities. For GDPNP-bound IF2C, the crosslike, flat structure of the GASBOR model can be recognized in the hybrid model, and for GDP-bound IF2C, domain VI-1 is placed to fit the protrusion giving the Y shape of the GASBOR model. Small differences in the maximum dimension of the two model approaches can be a consequence of modeling algorithms. For the ab initio (GASBOR) model, the maximum dimension obtained from the $p(r)$ curve is a model constraint, whereas for the rigid body modeling, it is not. In any case, there is a very good agreement between the two types of models obtained. The alignments show how the positions of the IF2C domains can be observed in the ab initio models, as well, showing the agreement of the two modeling approaches.

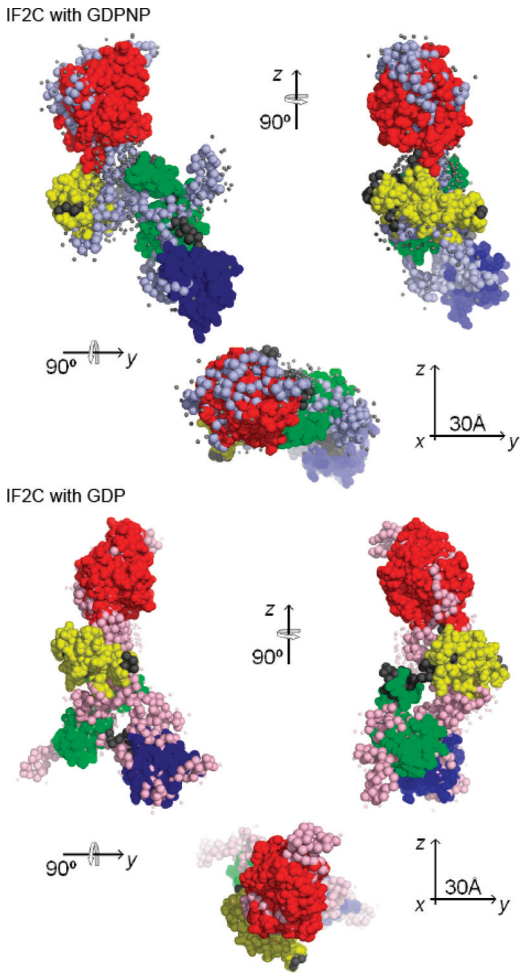


Figure 5. Model comparison. Superposition of each hybrid model onto the corresponding ab initio model shown in three different perpendicular orientations. As one can see, there is a very good agreement between the two model approaches, supporting the conclusions about the changes in tertiary structure of IF2C upon nucleotide binding: (top) IF2C with GDPNP (ab initio model colored light blue) and (bottom) IF2C with GDP (ab initio model colored light red).

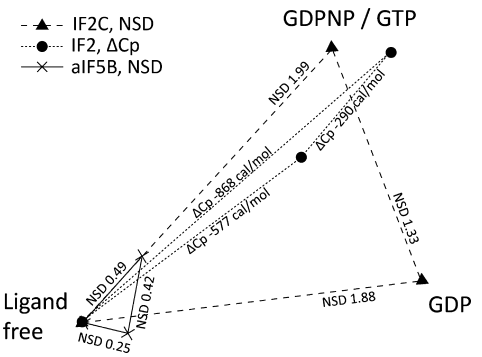


Figure 6. Graphic representation of NSD and ΔC_p values for structural comparisons. NSD values are represented as relative distances among the three forms of IF2C or aIF5B: ligand-free, GDPNP-bound, and GDP-bound. A short distance indicates structural proximity. The change in heat capacity, ΔC_p, is another measure of structural transition. Absolute ΔC_p values are represented in the same way using relative distances among the three forms of IF2 (GTP was used instead of GDPNP), and short distances (small changes in heat capacity) indicate structural proximity. The ΔC_p values are from ref 15. For both IF2 and IF2C, the GDPNP/GTP- and GDP-bound forms are closest to each other, whereas for aIF5B, the ligand-free and GDP-bound forms are closest to each other.

Structural Transitions upon Nucleotide Binding. From the analysis and modeling of the SAXS data, one can conclude that IF2C changes conformation when going from ligand-free to GDPNP-bound to GDP-bound form. The question is how to quantify these changes. The intrinsic low-resolution nature of SAXS prevents the direct comparison of structures at the atomic level, but an overall idea of the changes involved can be deduced. SUPCOMB is used to optimize the alignment of three-dimensional objects. The carbon backbones of the hybrid models are aligned, and the output from SUPCOMB includes a value for normalized spatial discrepancy (NSD),[24] which is used to quantify the agreement between individual structures. Values of <1.0 indicate close structural proximity. NSD for GDPNP-bound IF2C aligned with GDP-bound IF2C is 1.33. The values are much higher for alignments between ligand-free IF2C and GDPNP-bound or GDP-bound IF2C (1.99 and 1.88, respectively). This shows that the largest structural transition of IF2C occurs upon binding of either nucleotide. A smaller transition occurs upon GTP hydrolysis (between GDPNP- and GDP-bound forms), at least outside the context of the ribosome. This is illustrated in Figure 6. The points representing the GDPNP- and GDP-bound forms of IF2C are closer to each other than to that of the ligand-free form.

This is compared to the NSD values for the structures of aIF5B with and without ligand: aIF5B−GDPNP versus aIF5B−GDP, NSD = 0.42; aIF5B−GDPNP versus aIF5B, NSD = 0.49; aIF5B−GDP versus aIF5B, NSD = 0.25. The NSD values are markedly lower than those for the IF2C models, indicating structural proximity among the three aIF5B forms, which have identical crystal packing with different nucleotides (even in the absence of nucleotide). Here, the ligand-free and GDP-bound forms of aIF5B are very similar. The largest structural transition occurs upon GTP binding, but GTP hydrolysis induces an almost equally large structural transition. These transitions (as seen in the crystal structures) are conferred by modest changes in the switch 2 region of domain IV[6] correlating with the classical model for GTPase activity,[28] but the structural transitions observed here for IF2C do not show the same characteristics.

Another quantification of structural transitions can be deduced from the thermodynamic changes involved in ligand binding. These thermodynamic changes can be expressed as the heat capacity change, ΔC_p, which through an empirical expression can be linked to a change in solvent-accessible area.[29] A large decrease in heat capacity is characteristic of a large reduction in the solvent-exposed area of nonpolar amino acids, whereas globular protein unfolding usually has a positive ΔC_p. The heat capacity changes for full-length IF2 upon binding of either GTP or GDP have been recorded previously using isothermal titration calorimetry (ITC):[15] IF2 to IF2−GTP, ΔC_p = −868 cal/mol; IF2 to IF2−GDP, ΔC_p = −577 cal/mol; IF2−GDP to IF2−GTP, ΔC_p = −291 cal/mol (the difference between the two other ΔC_p values). These data are also visualized in Figure 6. The ΔC_p values were translated into changes in solvent-accessible surface area (ΔSA): IF2 to IF2−GTP, ΔSA = −2170−3210 Å^2; IF2 to IF2−GDP, ΔSA = −1440−2140 Å^2; IF2−GDP to IF2−GTP, ΔSA = −725−1074 Å^2. This study showed that the structural changes in IF2 upon

dx.doi.org/10.1021/bi200938q | Biochemistry 2011, 50, 9779–9787

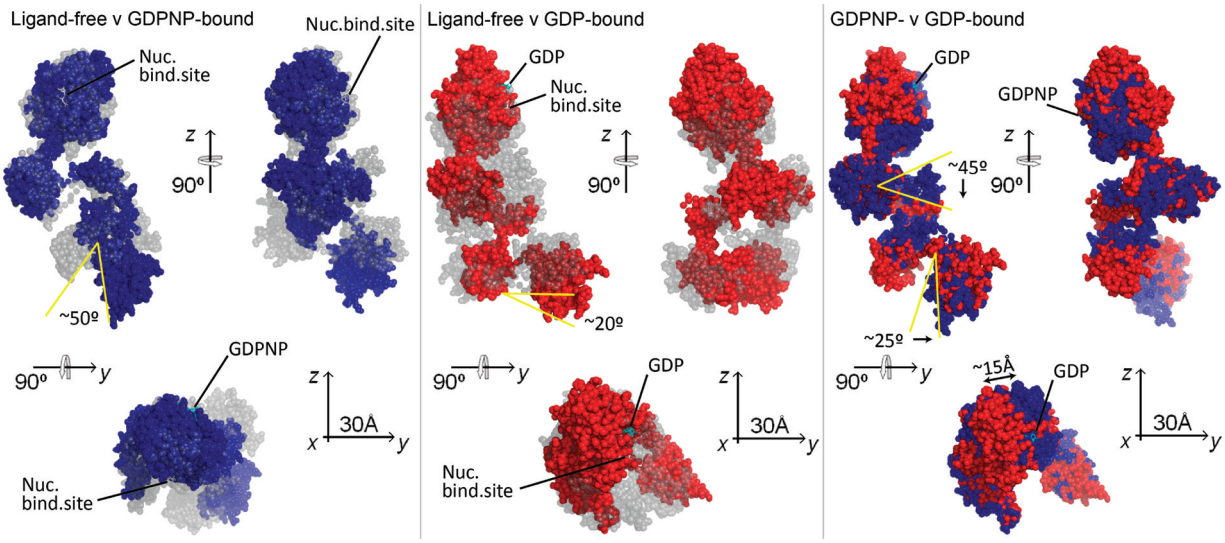


Figure 7. Structural transitions of IF2C. SUPCOMB was used to align the three different forms of IF2C: (left) GDPNP-bound IF2C (blue) superimposed on ligand-free IF2C (gray), (center) GDP-bound IF2C (red) superimposed on ligand-free IF2C, and (right) GDPNP-bound IF2C superimposed on GDP-bound IF2C. GDPNP and GDP (cyan) (nucleotide binding site colored white in the case of ligand-free IF2C) are marked where visible on the surface of the models. All superpositions are shown in three different perpendicular orientations. Indicated angles and distances were measured in PyMOL.

binding of either ligand are much larger than the difference between GTP- and GDP-bound IF2.

The hybrid models of ligand-binding IF2C presented here are in line with these findings as opposed to the classical GTPase behavior of aIF5B. CRYSOL output includes the surface area of the analyzed models: IF2C, SA = 9112 Å^2; IF2C−GDPNP, SA = 8362 Å^2; IF2−GDP, SA = 8071 Å^2. This allows calculation of the changes in surface area among the three hybrid models: IF2C to IF2C−GDPNP, ΔSA = −750 Å^2; IF2C to IF2C−GDP, ΔSA = −1041 Å^2; IF2C−GDP to IF2C−GDPNP, ΔSA = 291 Å^2. Compared to the ITC data, the hybrid models also show large decreases in surface area upon binding of either ligand. Because the hybrid models were generated for IF2C and the ITC data were recorded for full-length IF2, the surface area changes are expectedly smaller for the IF2C hybrid models, but the trend is the same. However, the change in surface area observed when changing from GDP- to GDPNP-bound IF2C has the opposite sign as measured by ITC, because the GDPNP-bound form has a larger surface area than the GDP-bound form. This discrepancy may relate to the use of GDPNP instead of GTP in the SAXS analysis, because some GTPases have a weaker affinity for GDPNP than for GTP.[12]

To visualize how changes in IF2C structure occur upon nucleotide binding, SUPCOMB was used to generate the best pairwise superposition of the models (Figure 7). It is clear that both GDPNP- and GDP-bound IF2C are more elongated than ligand-free IF2C. For the GDPNP-bound form, domain VI-2 seems to be flipped away from the rest of the protein at an angle of ~50°. Because this domain contacts initiator tRNA on the ribosome, it makes sense that this domain is stretched out to ensure this contact. For GDP-bound IF2C, domain VI-2 is pulled back a bit; however, the overall structure is still more elongated than ligand-free IF2C, and domain IV is also moved farther away. Upon comparison of the two nucleotide-binding forms of IF2C, they are obviously much more alike, even though differences can be observed. When GDPNP is exchanged for GDP, domain VI-1 moves ~15 Å downward at an angle of ~45° in relation to domain V. At the same time, domain VI-2 is moved ~25°, resulting in the observed less elongated shape. Domain IV is moved ~15 Å sideways, which is best seen from the top view (Figure 7, right panel).

IF2C on the Ribosome. On the basis of the classic model for GTPase activity, IF2C should show a more dramatic structural change between binding GDPNP and GDP than between the ligand-free and GDP-bound forms. However, this is not the case as shown by these SAXS data, and the same has been found for several other GTPases involved in ribosome function.[12] This led to the concept of cofactor-dependent conformational switching of GTPases, suggesting that the binding of GTP alone is not enough to shift the GTPase to its active form. Only on the ribosome in contact with cofactors as initiator tRNA does the GTPase adopt its active conformation and hydrolyze GTP. Our data seem to support this concept. There are differences between the GDPNP- and GDP-bound forms of IF2C, like the stretching of domain VI-2, but the overall structures are rather similar.

When IF2 in a GTP-bound form joins the 30S ribosomal subunit, it ensures the correct placement of initiator tRNA. The structure of *T. thermophilus* IF2 in the 30S initiation complex is very elongated, but here domains IV and V seem clustered and separated from domains VI-1 and VI-2 by a long thin stretch.[30] This structure of IF2 in the 30S initiation complex has not been deposited in the Protein Data Bank, so direct comparisons are not possible; however, this structure is clearly different from the models determined in this study. It is possible that IF2 has one conformation when GDPNP-bound and in solution, and this structure changes upon binding to 30S and initiator tRNA. When 50S joins, this induces further changes in IF2,[30] resulting in active GTP hydrolysis by domain IV (the G-domain) now positioned in the GTPase-activating center of the 50S subunit.

The electron density of IF2C on the 70S ribosome was recorded,[10] and the domains of the crystal structure were rearranged to fit this electron density. This new structure was deposited in the Protein Data Bank as entry 1ZO1. Figure 8 shows GDPNP-bound IF2C superpositioned on this ribosome-associated structure, and it is clear that GDPNP-bound IF2C is

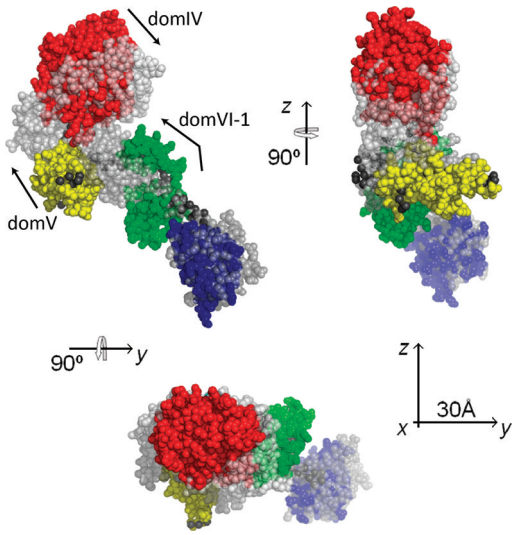


Figure 8. IF2C on the ribosome. GDPNP-bound IF2C (color-coded according to Figure 1) was superimposed on the structure of ribosome-associated IF2 (gray, PDB entry 1ZO1) using SUPCOMB. The superposition is shown in three different perpendicular orientations. Arrows indicate the required movement of GDPNP-bound IF2C domains to fit the ribosome-associated structure.

not yet in the right conformation found in the 70S ribosome. The stretching of domain VI-2 seems to fit very well with the ribosome-associated structure, but in addition, the three other domains need to be clustered at a point even farther from domain VI-2. This suggests that the binding of GDPNP results in only an intermediate structure in the transition to an active GTPase on the ribosome, and that the binding to the ribosome itself induces further conformational switching. It is possible that using GTP instead of GDPNP would induce larger conformational changes, but because the ITC data showed only modest structural transitions between GTP- and GDP-bound IF2, this hypothesis seems unlikely.

AUTHOR INFORMATION

Corresponding Author

*Department of Molecular Biology, Aarhus University, Gustav Wieds Vej 10, DK-8000 Aarhus C, Denmark. Fax: +45 86182812. Phone: +45 89424891. E-mail: kkm@mb.au.dk.

Funding

This research was supported by Danish Natural Science Research Council Grants 21-03-0592 and 21-03-0465 and by 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.).

ABBREVIATIONS

aIF5B, archaeal translation initiation factor 5B; cryo-EM, cryo-electron microscopy; GDP, guanosine 5′-diphosphate; GDPNP, guanylylimino diphosphate; GMPPCP, guanylyl 5′-(β,γ-methylenediphosphonate); IF, translation initiation factor; IF2C, domains IV to VI-2 of *E. coli* IF2; ITC, isothermal titration calorimetry; NSD, normalized spatial discrepancy; PDB, Protein Data Bank; P-site, peptidyl site; SA, surface area; SAXS, small-angle X-ray scattering.

REFERENCES

(1) 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.

(2) 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.

(3) Marshall, R. A., Aitken, C. E., and Puglisi, J. D. (2009) GTP hydrolysis by IF2 guides progression of the ribosome into elongation. *Mol. Cell 35*, 37−47.

(4) 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.

(5) 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.

(6) 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.

(7) Bell, S. D., and Jackson, S. P. (1998) Transcription and translation in Archaea: A mosaic of eukaryal and bacterial features. *Trends Microbiol. 6*, 222−228.

(8) Tishchenko, E. V. (2005) NMR studies of the GTP/GDP binding domain of translation initiation factor IF2. Ph.D. Thesis, Utrecht University, Utrecht, The Netherlands.

(9) Milon, P., Tischenko, E., Tomsic, J., Caserta, E., Folkers, G., La Teana, A., Rodnina, M. V., Pon, C. L., Boelens, R., and Gualerzi, C. O. (2006) The nucleotide-binding site of bacterial translation initiation factor 2 (IF2) as a metabolic sensor. *Proc. Natl. Acad. Sci. U.S.A. 103*, 13962−13967.

(10) 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.

(11) 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.

(12) Hauryliuk, V., Hansson, S., and Ehrenberg, M. (2008) Cofactor dependent conformational switching of GTPases. *Biophys. J. 95*, 1704−1715.

(13) 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.

(14) Rasmussen, L. C., Oliveira, C. L., Jensen, J. M., Pedersen, J. S., Sperling-Petersen, H. U., and Mortensen, K. K. (2008) Solution structure of C-terminal *Escherichia coli* translation initiation factor IF2 by small-angle X-ray scattering. *Biochemistry 47*, 5590−5598.

(15) Hauryliuk, V., Mitkevich, V. A., Draycheva, A., Tankov, S., Shyp, V., Ermakov, A., Kulikova, A. A., Makarov, A. A., and Ehrenberg, M. (2009) Thermodynamics of GTP and GDP binding to bacterial initiation factor 2 suggests two types of structural transitions. *J. Mol. Biol. 394*, 621−626.

(16) 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.

(17) Glatter, O. (1977) A new method for the evaluation of small-angle scattering data. *J. Appl. Crystallogr. 10*, 415−421.

(18) 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.

(19) Svergun, D. I., Semenyuk, A. V., and Feigin, L. A. (1988) Small-Angle-Scattering-Data Treatment by the Regularization Method. *Acta Crystallogr. A44*, 244−250.

(20) Svergun, D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr. 25*, 495−503.
(21) Svergun, D. I., and Koch, M. H. J. (2003) Small-angle scattering studies of biological macromolecules in solution. *Rep. Prog. Phys. 66*, 1735−1782.
(22) Svergun, D. I., Petoukhov, M. V., and Koch, M. H. (2001) Determination of domain structure of proteins from X-ray solution scattering. *Biophys. J. 80*, 2946−2953.
(23) 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.
(24) Kozin, M. B., and Svergun, D. I. (2001) Automated matching of high- and low-resolution structural models. *J. Appl. Crystallogr. 34*, 33−41.
(25) Volkov, V. V., and Svergun, D. I. (2003) Uniqueness of *ab-initio* shape determination in small-angle scattering. *J. Appl. Crystallogr. 36*, 860−864.
(26) 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.
(27) *The PyMOL Molecular Graphics System*, version 0.99 (2006) Schrödinger LLC, New York.
(28) Sprang, S. R. (1997) G protein mechanisms: Insights from structural analysis. *Annu. Rev. Biochem. 66*, 639−678.
(29) Prabhu, N. V., and Sharp, K. A. (2005) Heat capacity in proteins. *Annu. Rev. Phys. Chem. 56*, 521−548.
(30) Simonetti, A., Marzi, S., Myasnikov, A. G., Fabbretti, A., Yusupov, M., Gualerzi, C. O., and Klaholz, B. P. (2008) Structure of the 30S translation initiation complex. *Nature 455*, 416−420.