Insertion domain within mammalian mitochondrial translation initiation factor 2 serves the role of eubacterial initiation factor 1

Aymen S. Yassin[a,1], Md. Emdadul Haque[b], Partha P. Datta[a,2], Kevin Elmore[b], Nilesh K. Banavali[a,c], Linda L. Spremulli[b], and Rajendra K. Agrawal[a,c,3]

[a]Wadsworth Center, New York State Department of Health, Empire State Plaza, Albany, NY 12201-0509; [b]Department of Chemistry, Campus Box 3290, University of North Carolina, Chapel Hill, NC 27599-3290; and [c]Department of Biomedical Sciences, School of Public Health, State University of New York, Albany, NY 12201

Edited by Peter B. Moore, Yale University, New Haven, CT 06520-8107, and approved January 21, 2011 (received for review November 23, 2010)

Mitochondria have their own translational machineries for the synthesis of thirteen polypeptide chains that are components of the complexes that participate in the process of oxidative phosphorylation (or ATP generation). Translation initiation in mammalian mitochondria requires two initiation factors, $IF2_{mt}$ and $IF3_{mt}$, instead of the three that are present in eubacteria. The mammalian $IF2_{mt}$ possesses a unique 37 amino acid insertion domain, which is known to be important for the formation of the translation initiation complex. We have obtained a three-dimensional cryoelectron microscopic map of the mammalian $IF2_{mt}$ in complex with initiator fMet-$tRNA_i^{Met}$ and the eubacterial ribosome. We find that the 37 amino acid insertion domain interacts with the same binding site on the ribosome that would be occupied by the eubacterial initiation factor IF1, which is absent in mitochondria. Our finding suggests that the insertion domain of $IF2_{mt}$ mimics the function of eubacterial IF1, by blocking the ribosomal aminoacyl-tRNA binding site (A site) at the initiation step.

protein synthesis | ribosome–$IF2_{mt}$ complex | cryo-EM structure | molecular modeling

Protein synthesis entails a complex series of events in which the ribosome interacts with a number of ligands and translation factors so as to conduct an efficient, accurate, and well-regulated process (1). The first event in translation is the initiation step, which requires the formation of an initiation complex, composed of the small ribosomal subunit (30S, in eubacteria), mRNA, initiator tRNA (fMet-$tRNA_i^{Met}$) and three initiation factors (IF1, IF2, and IF3) (2). IF3 helps in preparation of the 30S initiation complex by preventing premature docking of the large subunit (50S, in eubacteria) (3, 4). IF2 promotes the binding of the initiator tRNA to the peptidyl (P) site of the 30S subunit and facilitates the association of the 50S subunit to form the 70S initiation complex (5). Initiation factor IF1, together with IF2, stabilizes the 30S initiation complex (6). IF1 also plays a role in translation initiation fidelity in conjunction with IF3 (7–9). The binding site of IF1 overlaps with that of the aminoacyl-tRNA on the 30S subunit (the A site), thereby making this site unavailable for the initiator tRNA during translation initiation (10–13). However, no equivalent of eubacterial IF1 has been detected in mammalian mitochondria (14); mitochondrial protein synthesis requires only two initiation factors, $IF2_{mt}$ and $IF3_{mt}$.

There is a lack of consensus in the nomenclature used to describe the domains of IF2 by biochemists (15–18) and structural biologists (12, 19). A uniform nomenclature that addresses both structural and functional considerations was used in our previous study (20). According to this nomenclature the approximately 98 kDa *Eschericia coli* IF2 is composed of six domains (Fig. 1*A*). Domains I and II are not present in all of the eubacterial factors, are not highly conserved, and are also absent in the mammalian $IF2_{mt}$. Domain VI is divided into two structural subdomains, VI C1 and VI C2. $IF2_{mt}$ is composed of four domains (III–VI) that are homologous to their eubacterial counterparts (21). In addition, $IF2_{mt}$ possesses an insertion domain of 37 amino acid (aa) residues between domains V and VI that is not present in *E. coli* IF2 (Fig. 1*A* and Fig. S1). Mutations in this insertion domain severely affect the ability of $IF2_{mt}$ to bind to the mitochondrial small subunit and to promote formation of the initiation complex (18). Recent studies have indicated that $IF2_{mt}$ can replace both eubacterial IF2 and IF1 in an *E. coli* strain with IF2 and IF1 gene knockouts (20). However, deletion of the 37-aa insertion from $IF2_{mt}$ necessitates the presence of IF1 in *E. coli*. This observation suggests that the 37-aa insertion in $IF2_{mt}$ functionally substitutes for eubacterial IF1. As in the case of $IF2_{mt}$, $IF3_{mt}$ also has significant structural differences when compared to eubacterial IF3. $IF3_{mt}$ shares a central conserved region with eubacterial IF3; in addition, $IF3_{mt}$ possesses unique approximately 30-aa long extensions at the N- and C-termini (22). Biochemical studies have shown that these extensions are important for proper dissociation of $IF3_{mt}$ from the initiation complex (23).

The binding sites of eubacterial initiation factors to eubacterial ribosomes have been studied both by chemical methods (10, 24) and by cryoelectron microscopy [cryo-EM; IF3 (3), IF2 (12, 13, 25)]. The binding site of IF1, smallest of the three, has also been determined by X-ray crystallography (11). IF1 binds to the 30S ribosomal subunit; it sits in a pocket formed by the components of the 16S rRNA helices 18 and 44, and ribosomal protein S12 (11). IF2 makes extensive contacts with both ribosomal subunits. The G domain of IF2 interacts with the GTPase-associated center in the 50S ribosomal subunit, whereas the factor's domain VI interacts with proteins L14 and L16 on the 50S ribosomal subunit (12). IF2 also interacts with a large area of the 30S subunit through domain V. In addition, domain VI interacts with the CCA region of the initiator tRNA within the initiation complex. Here we show, using cryo-EM and 3D image reconstruction techniques that a 37 amino acid insertion present in the sequence

Author contributions: L.L.S. and R.K.A. designed research; A.S.Y., M.E.H., P.P.D., K.E., and N.K.B. performed research; A.S.Y. and R.K.A. analyzed data; and A.S.Y., N.K.B., L.L.S., and R.K.A. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The cryo-EM map of the 70S • mRNA • fMet-$tRNA_i^{Met}$ • $IF2_{mt}$ • GDPNP complex and the extracted density corresponding to the tRNA • $IF2_{mt}$ complex have been deposited in the EM database, http://emdep.rutgers.edu (accession codes EMD-1854 and EMD-1855). The modeled and fitted coordinates of $IF2_{mt}$ and relevant components of the ribosome structure have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID codes 3IZY and 3IZZ).

[1]Present address: Department of Microbiology and Immunology, Faculty of Pharmacy, Cairo University, Cairo 11562, Egypt.

[2]Present address: Department of Biology, Indian Institute of Science Education and Research-Kolkata, Mohanpur - 741252, Nadia, WB, India.

[3]To whom correspondence should be addressed. E-mail: agrawal@wadsworth.org.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1017425108/-/DCSupplemental.

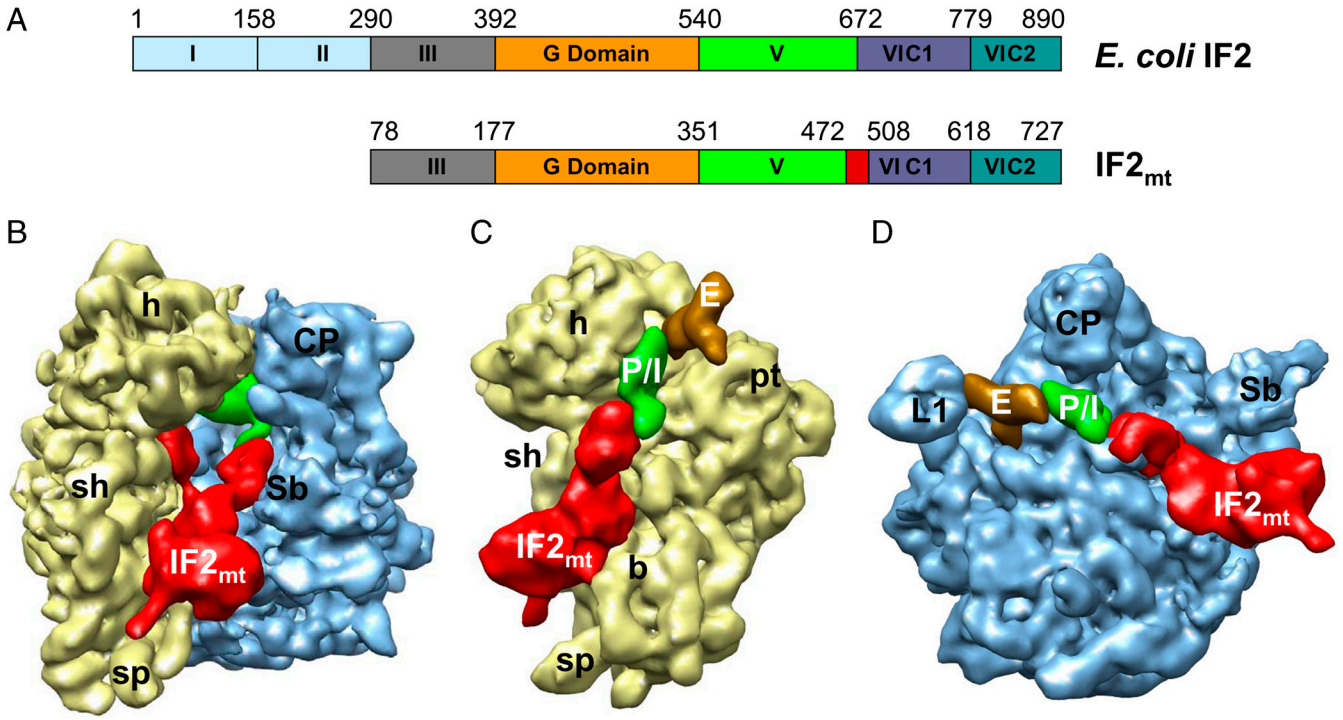

Fig. 1. Comparison of the domain organization of *E. coli* IF2 and mammalian $IF2_{mt}$, and cryo-EM structure of the 70S • mRNA • fMet – RNA_i^{Met} • $IF2_{mt}$ • GDPNP complex. (*A*) Domain alignment of *E. coli* IF2 and bovine $IF2_{mt}$. Note that the mature $IF2_{mt}$ (i.e., after deletion of the import sequence) starts from aa residue 78. The 37-aa insertion in $IF2_{mt}$ is highlighted in red. (*B*) The cryo-EM map of the *E. coli* 70S ribosome (30S subunit, yellow; 50S subunit, blue) in complex with $IF2_{mt}$ (red), initiator tRNA (green) at the P/I position, and an additional tRNA density (brown, visible in panels *C* and *D*) at the ribosomal E site. (*C* and *D*) The 30S and 50S subunit portions of the map are shown from their interface sides to reveal the overall binding positions of $IF2_{mt}$ and tRNAs. Landmarks of the 30S subunit: h, head; sh, shoulder; b, body; and pt, platform. Landmarks of the 50S subunit: L1, L1 protein protuberance; CP, central protuberance; and Sb, L7/L12 stalk base. For stereo viewing, see Fig. S2.

of $IF2_{mt}$ interacts with the same binding site on the small subunit of the ribosome that would be occupied by the eubacterial initiation factor IF1. Our finding suggests that the insertion domain of $IF2_{mt}$ mimics the function of eubacterial IF1 by structurally occupying the same binding site.

Results and Discussion

Structure of the 70S • mRNA • fMet-$tRNA_i^{Met}$ • $IF2_{mt}$ • GDPNP Complex. To determine how the 37-aa insertion domain of $IF2_{mt}$ can mimic the function of eubacterial IF1 (20), we prepared an initiation complex of *E. coli* 70S ribosomes in which the eubacterial IF1 and IF2 were replaced by $IF2_{mt}$, and then obtained the 3D structure of the complex by single-particle cryo-EM technique (26) (see *Materials and Methods*). A comparison of the cryo-EM map of the 70S • mRNA • fMet-$tRNA_i^{Met}$ • $IF2_{mt}$ • GDPNP initiation complex (Fig. 1*B* and Fig. S2 for stereo viewing) with that of the control 70S ribosome revealed several distinct extra densities of mass in the intersubunit space that can be readily assigned to $IF2_{mt}$, the initiator fMet-$tRNA_i^{Met}$ (in the P/I position, see below) (12), and a deacylated tRNA at the ribosomal exit (E) site. The density corresponding to $IF2_{mt}$ makes extensive contacts with the ribosome, spanning major portions of the intersubunit surfaces of both ribosomal subunits (Fig. 1 *C* and *D*). On the interface side of the 30S subunit, the $IF2_{mt}$ density extends from the lower body region up to the decoding center, with a protruding mass occupying the position of the A-site tRNA. On the 50S subunit, the $IF2_{mt}$ density spans the GTPase-associated center and extends slightly beyond the ribosomal P site, where it maintains its interaction with the acceptor end of the initiator tRNA. The overall binding position of $IF2_{mt}$ on the 70S ribosome (Fig. 1*B*) closely matches to that of eubacterial IF2 (12), EF-G (27, 28), and EF-Tu (29, 30).

Structure of the $IF2_{mt}$ • fMet – $tRNA_i^{Met}$ • GDPNP Complex. To analyze the molecular interactions between different domains of $IF2_{mt}$ and the ribosome, we built a composite homology model of bovine $IF2_{mt}$ based on: the X-ray crystallographic structure of an archaeal IF2 (aIF5B) from *Methanobacterium thermoautotrophicum* [Protein Data Bank (PDB) ID code 1G7T] (19), the NMR structures of the C1 (PDB ID code 1Z9B) (31), and C2 (PDB ID code 1D1N) (32) portions of the IF2's domain VI from *Bacillus stearothermophilus*. The model took into account the corresponding sequence alignments between $IF2_{mt}$ and the homologous archaeal and eubacterial IF2 sequences and the cryo-EM density map (see *Materials and Methods* and *SI Text* for additional details). Our model of $IF2_{mt}$ shows a close structural resemblance to the corresponding regions in the atomic structures of the archaeal and eubacterial IF2 (Fig. S3), with the only major exception being the structure of the 37-aa insert that extends between domain V and the C1 portion of Domain VI (Fig. 2). This 37 aa insert was modeled separately, using the program I-TASSER (33). The compact structural domain formed by 3 α-helices of the insertion segment does not match the OB fold β-strand-rich X-ray crystallographic structure of the nonhomologous IF1 from *Thermus thermophilus* (11) (see *SI Text*). The N-terminal fragment (domain III, corresponding to aa 78–176) of the protein could not be modeled in the final structure. (Note that aa 1–77 constitute the signal sequence, which is cleaved off when $IF2_{mt}$ is imported into the mitochondrion; consequently, they are not present in the mature protein.) However, our cryo-EM map does show an extra mass (marked by an * in Fig. 2), which extends from the lower part of the $IF2_{mt}$ and its long axis is directed toward the small ribosomal subunit; that feature can snugly accommodate the 99-aa domain III, and is therefore tentatively assigned to domain III of $IF2_{mt}$. The final atomic model of $IF2_{mt}$ was obtained by flexible fitting

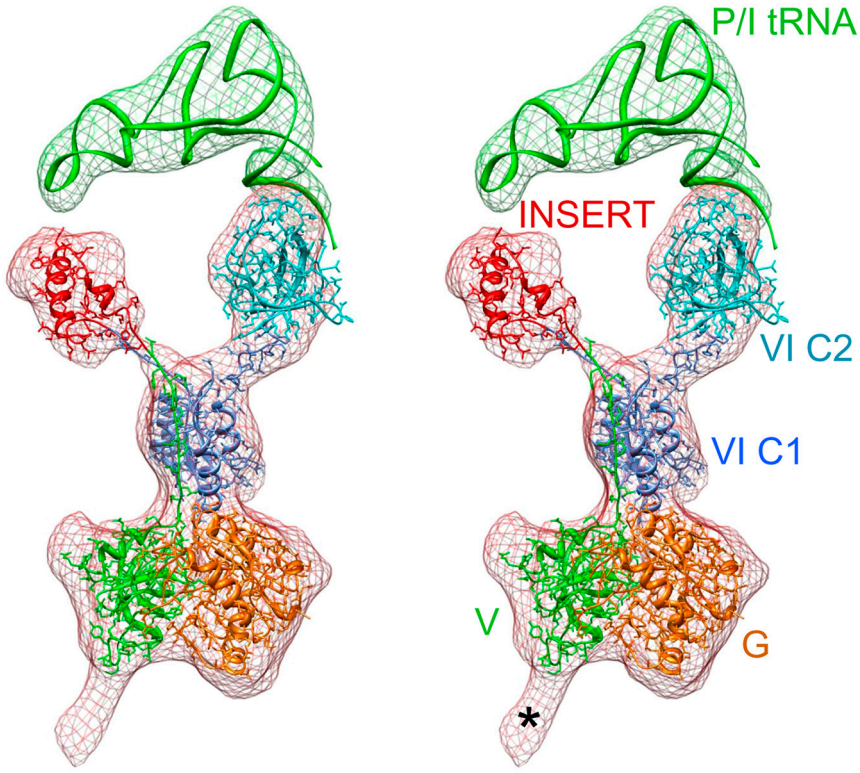

Fig. 2. Stereo-view representation of the fitting of atomic models of $IF2_{mt}$ and P/I tRNA into the corresponding cryo-EM densities. The cryo-EM densities are shown as meshwork. The color codes used for various domains of IF2mt are the same as in Fig. 1*A*. Asterisk (*) points to the region that would correspond to domain III of $IF2_{mt}$ and is not present in our composite homology model.

into the cryo-EM density corresponding to $IF2_{mt}$ (Fig. 2) using the Molecular Dynamics Flexible Fitting (MDFF) protocol (34). An atomic structure of the tRNA (PDB code 1ZO3) (12) was fit into the densities that correspond to the initiator and E-site tRNAs. The fitting of the atomic model of the initiator tRNA ($tRNA_i^{Met}$) into the corresponding cryo-EM density revealed a distortion in the ribosome-bound tRNA, the anticodon stem-loop occupied the P-site position on the 30S subunit, whereas the CCA end occupied a position between the canonical P and E sites, referred to as the initiator tRNA binding position [or P/I position (12)], on the large ribosomal subunit. The C2 portion of $IF2_{mt}$'s domain VI interacts with the CCA end of the initiator tRNA, and apparently helps dock the $tRNA_i$ in its proper orientation on the ribosome. Because the CCA end of the initiator tRNA is not positioned in the canonical P site, and the A site on the 30S subunit is still occupied by the insertion domain of $IF2_{mt}$, the present complex corresponds to a functional intermediate state between the 30S and 70S initiation complexes.

Interaction Between $IF2_{mt}$ and the Ribosome. The crystal structure of the *E. coli* large ribosomal subunit (PDB code 2AW4) (35) and the crystal structure of the *T. thermophilus* small ribosomal subunit in complex with IF1 (11) (PDB code 1HR0) were fit into the corresponding portions of the cryo-EM map and used to identify (*i*) segments of the 16S and 23S ribosomal RNAs (rRNAs) and (*ii*) ribosomal proteins that interact with various structural domains of $IF2_{mt}$. Dockings of all molecular structures into the cryo-EM map showed that domains V and VI C1 of $IF2_{mt}$ interact with the 30S ribosomal subunit (Fig. 3*A*), mainly through its two 16S rRNA helices: h5 [nucleotides (nt) 55–56] and h15 (nt 368). Domain V is also in close proximity, within 5–10 Å, to 16S rRNA helix h14. The 37-aa insertion present in $IF2_{mt}$ fits tightly into the cryo-EM density feature that is close to the decoding region of the 30S subunit (Fig. 2). This region of the subunit is a pocket formed by the 530 loop (h18) and the tip of h44 of the 16S rRNA, and protein S12. We find that the above ribosomal components each undergo relative shifts to interact with the 37-aa insertion domain of $IF2_{mt}$ (Fig. S4). Therefore, atomic structures of each of the binding-pocket components (11) were individually fit into the cryo-EM density map, to allow us to delineate interactions. The insertion domain interacts with nts 517–519 and 530–531 of h18; nts 1397, 1411, and 1492–1493 of h44; and aa residue Leu52 of protein S12. The aa residues that form the linker between the insert and the rest of $IF2_{mt}$ interact with Arg41-Thr44, whereas domain VI CI contacts Arg59 and Glu79-His80 of S12. Because each of the above ribosomal components are conserved between *E. coli* and mammalian mitochondria (36), it is likely that the mode of interaction of the insertion domain is the same with small mitoribosomal subunit. $IF2_{mt}$ makes extensive contacts with the 50S ribosomal subunit as well (Fig. 3*B*). The G and VI C1 domains of $IF2_{mt}$ interact with the sarcin/ ricin loop (H95) of the 23S rRNA at nts 2660–2661 and 2662–2663, respectively. Domain VI C2 contacts nts 2482-2483 of H89. In addition, the insertion and VI C1 domains are in close proximities (within 4–6 Å) to the 23S rRNA helix H69 (nts 1913–1915) and protein L14 (Lys51), respectively.

Role of the 37 Amino Acid Insert in $IF2_{mt}$. As mentioned earlier, for interpretation of the 30S subunit portion of the cryo-EM map, we used the X-ray crystallographic structure of the 30S • IF1 complex (11), to enable assessment of the relative positions of IF1 and $IF2_{mt}$. This analysis revealed that the eubacterial IF1 and the 37-aa insertion domain of $IF2_{mt}$ occupy the same binding pocket on the 30S subunit (11) (Fig. 4 and Table S1). Despite the fact that the length of the $IF2_{mt}$ insertion is only about half of that of eubacterial IF1 (37 aa vs. 72 aa), their overlapping binding positions on the ribosome suggest that the insert plays a functional role similar to that of IF1 in eubacteria. Furthermore, there is no significant similarity between the sequences of the 37-aa

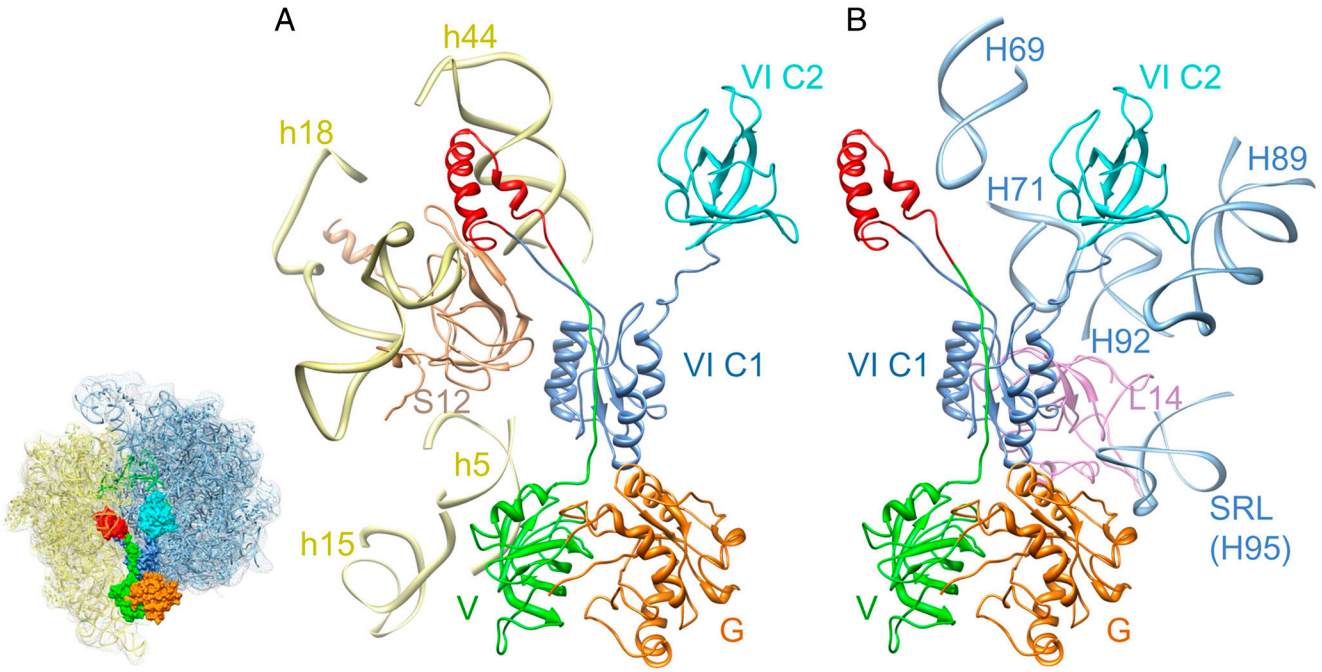

Fig. 3. Interactions of the individual domains of $IF2_{mt}$ with components of the (*A*) small and (*B*) large ribosomal subunits. Helices of the 16S and 23S rRNAs are identified with h and H, respectively, whereas ribosomal proteins of the small and large subunits are prefixed by S and L, respectively. $IF2_{mt}$ domains are color coded as in Fig. 2. Thumbnail to the left shows the overall orientation of the ribosome, with fitted atomic coordinates of *E. coli* 70S ribosome (35), $IF2_{mt}$ (space-filling model), and the P/I tRNA (green ribbons).

insertion and IF1 (Fig. S5). However, surfaces of both IF1 and the 37-aa insertion that interact with the 30S subunit show resemblance in their overall charge distribution; i.e., both surfaces have strong patches of electropositive charge (Fig. 5 *A* and *B*). Our study suggests that the functional similarity between the IF1 and the 37-aa insertion (20) is likely because they both sterically block the ribosomal A site (Fig. 4), thereby promoting binding of the initiator tRNA to the ribosomal P site during translation initiation.

Notably, the smallest of the eubacterial initiation factors has been physically integrated into $IF2_{mt}$. The mitochondrial matrix is densely crowded (approximately 800 mg/mL of protein) compared to the cytoplasm of an *E. coli* cell (approximately 300 mg/mL of proteins) (37). In addition, the structure of the mitochondrion is characterized by the convoluted cristae that form the inner membrane (38). Such topologically intricate structural organization, combined with the matrix density, could hinder smaller factors of the size of eubacterial IF1 from finding their target sites. Furthermore, because mitochondrial initiation factors are present in very low amounts [less than 20 copies per mitochondrion (39)], diffusion restrictions could pose problems for efficient initiation of translation. We hypothesize that, over the course of evolution, the mitochondrial version of IF1 was eliminated; instead, a smaller peptide bearing its function became an integral part of the larger $IF2_{mt}$ factor. Such transfer of structural and functional characteristics of IF1 to the $IF2_{mt}$ molecule would eliminate the necessity to import a very small polypeptide such as IF1, which is as small as some of the import (signal) sequences, from the cytoplasm into the mitochondrial inner matrix.

Materials and Methods

Preparation of the Ribosome • mRNA • fMet – $tRNA_i^{Met}$ • $IF2_{mt}$ • GDPNP Complex. Reassociated 70S ribosomes were prepared as described by Blaha and coworkers (41), using crude ribosomes prepared from *E. coli*. The 30S and 50S ribosomal subunits were purified by dissociation of crude ribosomes on 10–30% sucrose gradients in a buffer containing 1 mM $MgAc_2$. Ribosomal subunits were reassociated by incubation for 20 min at 25 °C in reassociation buffer (20 mM Hepes-KOH, pH 7.6, 10 mM $MgAc_2$, 30 mM KCl, and 4 mM β-mercaptoethanol) and subsequently purified on 10–30% sucrose gradients

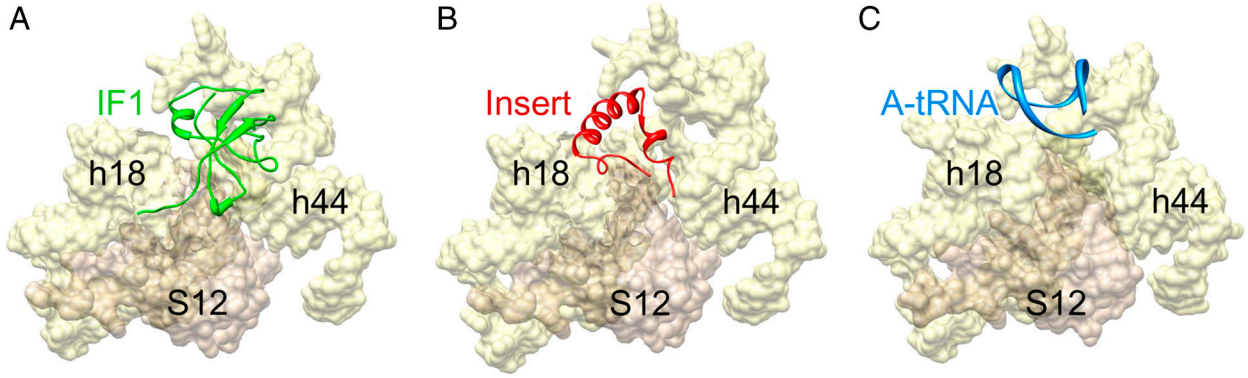

Fig. 4. Location of the insertion domain of $IF2_{mt}$ relative to the binding site of eubacterial IF1 and anticodon stem-loop of the A-site tRNA on the small subunit. (*A*) IF1 (11) (green); (*B*) 37-aa insertion domain (red); and (*C*) anticodon stem-loop of the A-site tRNA (40) (blue). Landmarks: h44 and h18, 16S rRNA helices; and S12, ribosomal protein S12.

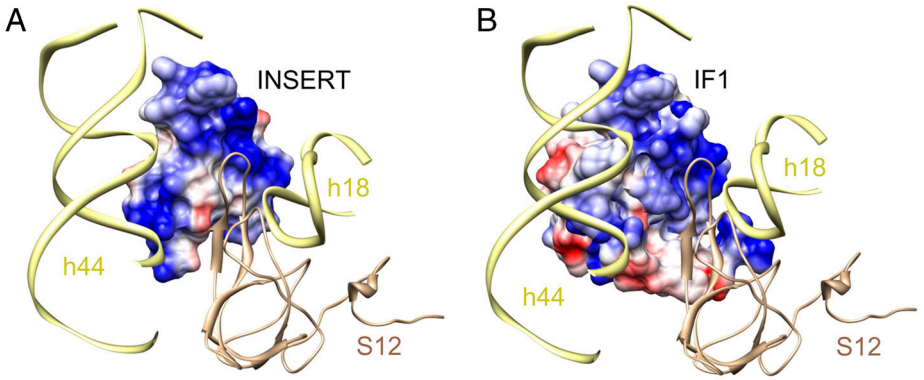

Fig. 5. Comparison of surface charge distribution on IF1 and the 37-aa insertion domain of $IF2_{mt}$. (*A*) Insert of $IF2_{mt}$, and (*B*) IF1. Both the insert and IF1 are shown in matching orientations with respect to their common binding pocket on the 30S subunit. Blue and red patches reflect positive and negative surface charge potentials, respectively. Landmarks are same as in Fig. 4.

containing 10 mM $MgAc_2$. Reassociated 70S ribosomes were pelleted at 40,000 × *g* for 27 h and resuspended in reassociation buffer. Mammalian mitochondrial translation initiation factors were expressed and purified as described previously (42). Ribosome activity was quantified by in vitro binding assays (18), using 0.2 μM 70S ribosomes and 1 μM fMet-$tRNA_i^{Met}$ in the presence of saturating levels of initiation factors (1 μM). Formation of the 70S • $IF2_{mt}$ • fMet-$tRNA_i^{Met}$ complex was optimized by testing binding with various buffer conditions. Assays showed that $IF3_{mt}$ is necessary to stimulate binding of $IF2_{mt}$. The optimal magnesium ion concentration for activity with the nonhydrolyzable GTP-analog (GDPNP) was 5 mM. Microcon (100 kDa cut-off) centrifugation (23) using 0.2 μM 70S ribosomes with 0, 0.32, and 0.64 μM $IF2_{mt}$, followed by SDS-PAGE analysis confirmed that 80% of the ribosomes contained $IF2_{mt}$. A complex in 25 μL was prepared containing 70S reassociated ribosomes (1 μM) with a 10-fold excess of the mitochondrial initiation factors (10 μM) in the presence of 5 μM fMet-$tRNA_i$ (43), 10 μM mRNA including a Shine–Dalgarno sequence GGCAA*GGAGG*AAUAAAAAUGUUCGCACGU (obtained from Dharmacon), 0.5 mM GDPNP, 50 mM Tris-HCl, pH 7.6, 5 mM $MgCl_2$, 80 mM KCl, and 1 mM dithiothreitol, by incubation at 25 °C for 20 min. The 70S ribosomes used for complex formation were 60% active in initiation complex formation. The complex was fast-frozen in a dry-ice isopropanol bath and stored at −80 °C.

Cryo-Electron Microscopy, Image Processing, and Three-Dimensional Reconstruction. Cryo-EM grids were prepared according to standard procedures (44). Data were collected on a Philips FEI Tecnai F20 field emission gun electron microscope with a magnification of 50,760. 392 micrographs were scanned on a Zeiss flatboard scanner with a step size of 14 μm, corresponding to 2.76 Å on the object scale and were sorted into 43 defocus groups. A total of 426,091 images were manually selected. SPIDER (45) was used for all image processing, including 2D image classification, reconstruction, and refinement. A previously determined cryo-EM structure of the 70S ribosome was used as the reference to align images for an initial reconstruction from fewer images, using projection matching procedure. The 3D volume so obtained was low pass filtered and used subsequently for alignment of a larger dataset. The original reconstruction and subsequent refinement yielded a weak mass of density corresponding to $IF2_{mt}$, as expected from the biochemical data showing partial occupancy of the factor on the ribosome. To enhance the density of the factor in the 70S ribosome complex, the method of supervised classification was applied (46). In supervised classification, the experimental images are realigned using 2D projections generated from two different reference maps, which in this case were the maps of a vacant 70S ribosome and a complex of the 70S ribosome in complex with IF1, IF2, IF3, fMet-$tRNA_i^{Met}$, and mRNA (12). 313,315 particle images (approximately 70% of the original total population, after removing images with poorer cross-correlation coefficient values with respect to reference projection images) were used in the supervised classification. 256,191 particle images classified with the projection images generated from the map of the IF2-bound 70S ribosome. The 3D map determined from these images showed an enhanced density for $IF2_{mt}$. However, our $IF2_{mt}$ cryo-EM density did not show the mass assigned by Allen and coworkers (12) to the N-terminus of *E. coli* IF2, thereby excluding the possibility of any reference bias due to use of the map of the *E. coli* 70S initiation complex (12) as a reference. To further enhance the structure of the 70S • mRNA • fMet – $tRNA_i^{Met}$ • $IF2_{mt}$ • GDPNP complex, another round of supervised classification was applied using the particles that had aligned to the projection images of the map of the IF2-bound 70S ribosome. In the next and final round, the newly reconstructed structure from the previous round was used as one of the two references. The other reference was a map of a factor-free, vacant ribosome. 121,742 particles classified with the 70S • mRNA • fMet-$tRNA_i^{Met}$ • $IF2_{mt}$ • GDPNP complex, and 135,323 particles classified with the vacant ribosome. After this round of classification of the dataset, the resolution of the cryo-EM map of the 70S • mRNA • fMet-$tRNA_i^{Met}$ • $IF2_{mt}$ • GDPNP complex improved slightly to 10.8 Å (at 0.5 cutoff of FSC), or 9.0 Å (using 0.143 cutoff of FSC) (47); however, the map showed much stronger density for $IF2_{mt}$, explaining all of the structural domains of the factor and the bound tRNAs. Modeling and fitting were done using O (48) and visualization was done using Chimera software (49).

Construction of the Atomic Model of $IF2_{mt}$. The $IF2_{mt}$ sequence shows an additional 37 aa between domains V and VI when compared to the *E. coli* IF1 sequence, but shows an additional 49 aa when compared to the archaeal *M. thermautotropicum* IF2 (aIF5B) sequences (Fig. S1). This 49-aa insert includes the 37-aa insert obtained by comparison to the *E. coli* IF2 sequence, with additional 3 N-terminal and 9 C-terminal flanking aa residues. Therefore, the initial homology model for $IF2_{mt}$ was constructed using MODELLER (50) and the *M. thermautotrophicum* IF2 X-ray crystallographic structure (PDB ID 1G7T) (19) as the structural template, by excluding this 49 aa insert. The sequence alignment between *M. thermautotrophicum* and $IF2_{mt}$ sequences is weaker for the C1 and C2 portions of domain VI. Better sequence alignment for these two regions is obtained with the *B. stearothermophilus* IF2 domains VI C1 and C2, for which solution NMR structures [PDB IDs 1Z9B (31) and 1D1N (32)] have been solved (Table S2). Homology models were constructed separately using MODELLER for these domain VI regions based on alignments for the $IF2_{mt}$ sequence with defined secondary structure regions of the NMR structures (Figs. S1 and S3). These homology models for the VI C1 and VI C2 domains were structurally aligned with the model constructed using the *M. thermautotrophicum* IF2 using the program RAPIDO (51) and then incorporated to create a continuous composite model.

The 49 aa residues, present in the $IF2_{mt}$ sequence but absent in the archael IF2 sequence and crystal structure, were modeled separately using the I-TASSER server (33), which predicted a compact structure with 3 α-helices. This helical nature agrees with the secondary structure prediction for this sequence using JPred (52). This 49-residue model was then covalently linked to the rest of the modeled IF2 using the program loopy (53) to predict a structure for linker regions with a minimal number of residues to obtain a complete continuous structure. The binding site densities of this 49 residue region and the remainder of the IF2 are spatially separated in the cryo-EM map and required linker loops of 11 residues on the N-terminal side and 6 residues on the C-terminal side of the insert sequence, resulting in melting some of the helical structure of the 49 residue I-TASSER model. It is certainly possible that some of the secondary structure elements seen in the homologous *M. thermautotrophicum* or *B. stearothermophilus* IF2 structures refold into a different topology in $IF2_{mt}$ to accommodate the additional 49 residues, but we chose to preserve the known folds as much as possible by minimizing the number of residues assigned to the linker regions.

The resulting composite homology model (Fig. S3, *Center*) of the entire $IF2_{mt}$ was minimized using the Steepest Descent minimization algorithm in the program CHARMM (54) until an energy gradient convergence criterion of 0.01 was met. This minimized model was initially fit as a rigid body into the

cryo-EM density corresponding to $IF2_{mt}$ and then flexibly fit using the SCX module of the Yup program (55), with a step factor of 1.0 and a temperature of 2 °K to an average goodness of fit of 0.62, and then minimized again using the above criteria. The fit of the model was further refined by flexible fitting using MDFF (34). This final model fits extremely well into the excised $IF2_{mt}$ cryo-EM density, as judged by high cross-correlation coefficient (CCC) value of 0.85 calculated between the cryo-EM density corresponding to $IF2_{mt}$ and the fitted $IF2_{mt}$ model (Table S2). The CCC values between the fitted X-ray coordinates and the corresponding cryo-EM density maps were determined after conversion of the fitted coordinates into the density map, through computation of averaged densities within volume elements scale-matched to those of the cryo-EM map (i.e., with a pixel size of 2.76 Å) and after filtration of the atomic model to the resolution of the cryo-EM density map (27). The achievable accuracy of fittings of atomic models into cryo-EM maps has been estimated to be in the range of 1/6 to 1/10 of the resolution of the cryo-EM map (see ref. 56); this would yield approximately 1.8 Å in our case.

ACKNOWLEDGMENTS. We thank Timothy Booth for some of the cryo-EM data collection, and Tanvir Shaikh for help with image processing. We also thank Manjuli Sharma and Takeshi Yokoyama for helpful discussions. N.K.B. gratefully acknowledges Wadsworth Center new investigator funds for equipment support. This work was supported by National Institutes of Health Grant R01 GM61576 (R.K.A.).

1. Ramakrishnan V (2002) Ribosome structure and the mechanism of translation. *Cell* 108:557–572.
2. Gualerzi CO, Pon CL (1990) Initiation of mRNA translation in prokaryotes. *Biochemistry* 29:5881–5889.
3. McCutcheon JP, et al. (1999) Location of translational initiation factor IF3 on the small ribosomal subunit. *Proc Natl Acad Sci USA* 96:4301–4306.
4. Grigoriadou C, Marzi S, Pan D, Gualerzi CO, Cooperman BS (2007) The translational fidelity function of IF3 during transition from the 30 S initiation complex to the 70 S initiation complex. *J Mol Biol* 373:551–561.
5. La Teana A, Gualerzi CO, Dahlberg AE (2001) Initiation factor IF 2 binds to the alpha-sarcin loop and helix 89 of *Escherichia coli* 23S ribosomal RNA. *RNA* 7:1173–1179.
6. Wintermeyer W, Gualerzi C (1983) Effect of *Escherichia coli* initiation factors on the kinetics of N-Acphe-tRNAPhe binding to 30S ribosomal subunits. A fluorescence stopped-flow study. *Biochemistry* 22:690–694.
7. Pon CL, Gualerzi CO (1984) Mechanism of protein biosynthesis in prokaryotic cells. Effect of initiation factor IF1 on the initial rate of 30 S initiation complex formation. *FEBS Lett* 175:203–207.
8. Milon P, Konevega AL, Gualerzi CO, Rodnina MV (2008) Kinetic checkpoint at a late step in translation initiation. *Mol Cell* 30:712–720.
9. Qin D, Fredrick K (2009) Control of translation initiation involves a factor-induced rearrangement of helix 44 of 16S ribosomal RNA. *Mol Microbiol* 71:1239–1249.
10. Moazed D, Samaha RR, Gualerzi C, Noller HF (1995) Specific protection of 16 S rRNA by translational initiation factors. *J Mol Biol* 248:207–210.
11. Carter AP, et al. (2001) Crystal structure of an initiation factor bound to the 30 S ribosomal subunit. *Science* 291:498–501.
12. 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.
13. Simonetti A, et al. (2008) Structure of the 30S translation initiation complex. *Nature* 455:416–420.
14. Koc EC, Spremulli LL (2002) Identification of mammalian mitochondrial translational initiation factor 3 and examination of its role in initiation complex formation with natural mRNAs. *J Biol Chem* 277:35541–35549.
15. 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.
16. Caserta E, et al. (2006) Translation initiation factor IF2 interacts with the 30 S ribosomal subunit via two separate binding sites. *J Mol Biol* 362:787–799.
17. Caserta E, et al. (2010) Ribosomal interaction of *Bacillus stearothermophilus* translation initiation factor IF2: characterization of the active sites. *J Mol Biol* 396:118–129.
18. Spencer AC, Spremulli LL (2005) The interaction of mitochondrial translational initiation factor 2 with the small ribosomal subunit. *Biochim Biophys Acta* 1750:69–81.
19. 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.
20. Gaur R, et al. (2008) A single mammalian mitochondrial translation initiation factor functionally replaces two bacterial factors. *Mol Cell* 29:180–190.
21. Lee JH, Choi SK, Roll-Mecak A, Burley SK, Dever TE (1999) Universal conservation in translation initiation revealed by human and archaeal homologs of bacterial translation initiation factor IF2. *Proc Natl Acad Sci USA* 96:4342–4347.
22. Haque ME, Spremulli LL (2008) Roles of the N- and C-terminal domains of mammalian mitochondrial initiation factor 3 in protein biosynthesis. *J Mol Biol* 384:929–940.
23. Haque ME, Grasso D, Spremulli LL (2008) The interaction of mammalian mitochondrial translational initiation factor 3 with ribosomes: Evolution of terminal extensions in IF3mt. *Nucleic Acids Res* 36:589–597.
24. Dallas A, Noller HF (2001) Interaction of translation initiation factor 3 with the 30S ribosomal subunit. *Mol Cell* 8:855–864.
25. Myasnikov AG, et al. (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.
26. Frank J, Penczek P, Agrawal RK, Grassucci RA, Heagle AB (2000) Three-dimensional cryoelectron microscopy of ribosomes. *Methods Enzymol* 317:276–291.
27. Agrawal RK, Penczek P, Grassucci RA, Frank J (1998) Visualization of elongation factor G on the *Escherichia coli* 70S ribosome: The mechanism of translocation. *Proc Natl Acad Sci USA* 95:6134–6138.
28. Gao YG, et al. (2009) The structure of the ribosome with elongation factor G trapped in the posttranslocational state. *Science* 326:694–699.
29. Stark H, et al. (1997) Visualization of elongation factor Tu on the *Escherichia coli* ribosome. *Nature* 389:403–406.
30. Schmeing TM, et al. (2009) The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science* 326:688–694.
31. Wienk H, et al. (2005) Solution structure of the C1-subdomain of *Bacillus stearothermophilus* translation initiation factor IF2. *Protein Sci* 14:2461–2468.
32. Meunier S, et al. (2000) Structure of the fMet-tRNA(fMet)-binding domain of *B. stearothermophilus* initiation factor IF2. *EMBO J* 19:1918–1926.
33. Zhang Y (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9:40.
34. Trabuco LG, Villa E, Schreiner E, Harrison CB, Schulten K (2009) Molecular dynamics flexible fitting: a practical guide to combine cryo-electron microscopy and X-ray crystallography. *Methods* 49:174–180.
35. Schuwirth BS, et al. (2005) Structures of the bacterial ribosome at 3.5 A resolution. *Science* 310:827–834.
36. Sharma MR, et al. (2003) Structure of the mammalian mitochondrial ribosome reveals an expanded functional role for its component proteins. *Cell* 115:97–108.
37. Srere PA (1980) The infrastructure of the mitochondrial matrix. *Trends Biochem Sci* 5:120–121.
38. Mannella CA, Marko M, Buttle K (1997) Reconsidering mitochondrial structure: new views of an old organelle. *Trends Biochem Sci* 22:37–38.
39. Liao HX, Spremulli LL (1990) Identification and initial characterization of translational initiation factor 2 from bovine mitochondria. *J Biol Chem* 265:13618–13622.
40. Selmer M, et al. (2006) Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* 313:1935–1942.
41. Blaha G, et al. (2000) Preparation of functional ribosomal complexes and effect of buffer conditions on tRNA positions observed by cryoelectron microscopy. *Methods Enzymol* 317:292–309.
42. Grasso DG, Christian BE, Spencer A, Spremulli LL (2007) Overexpression and purification of mammalian mitochondrial translational initiation factor 2 and initiation factor 3. *Methods Enzymol* 430:59–78.
43. Graves MC, Spremulli LL (1983) Activity of Euglena gracilis chloroplast ribosomes with procaryotic and eucaryotic initiation factors. *Arch Biochem Biophys* 222:192–199.
44. Grassucci RA, Taylor DJ, Frank J (2007) Preparation of macromolecular complexes for cryo-electron microscopy. *Nat Protoc* 2:3239–3246.
45. Shaikh TR, et al. (2008) SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. *Nat Protoc* 3:1941–1974.
46. Valle M, et al. (2002) Cryo-EM reveals an active role for aminoacyl-tRNA in the accommodation process. *EMBO J* 21:3557–3567.
47. Rosenthal PB, Henderson R (2003) Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J Mol Biol* 333:721–745.
48. Jones TA, Zou JY, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* 47:110–119.
49. Pettersen EF, et al. (2004) UCSF Chimera-a visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612.
50. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234:779–815.
51. Mosca R, Brannetti B, Schneider TR (2008) Alignment of protein structures in the presence of domain motions. *BMC Bioinformatics* 9:352.
52. Cuff JA, Barton GJ (1999) Evaluation and improvement of multiple sequence methods for protein secondary structure prediction. *Proteins* 34:508–519.
53. Xiang Z, Soto CS, Honig B (2002) Evaluating conformational free energies: the colony energy and its application to the problem of loop prediction. *Proc Natl Acad Sci USA* 99:7432–7437.
54. Brooks BR, et al. (2009) CHARMM: The biomolecular simulation program. *J Comput Chem* 30:1545–1614.
55. Tan RKZ, Petrov AS, Harvey SC (2006) YUP: A molecular simulation program for coarse-grained and multiscaled models. *J Chem Theor Comput* 2:529–540.
56. Rossmann MG (2000) Fitting atomic models into electron-microscopy maps. *Acta Crystallogr D* 56:1341–1349.