Structures of ribosome-bound initiation factor 2 reveal the mechanism of subunit association

Thiemo Sprink,1 David J. F. Ramrath,1 Hiroshi Yamamoto,1 Kaori Yamamoto,1 Justus Loerke,1 Jochen Ismer,1* Peter W. Hildebrand,1* Patrick Scheerer,1 Jörg Bürger,1,2 Thorsten Mielke,1,2 Christian M. T. Spahn1,2

Throughout the four phases of protein biosynthesis—initiation, elongation, termination, and recycling—the ribosome is controlled and regulated by at least one specified translational guanosine triphosphatase (trGTPase). Although the structural basis for trGTPase interaction with the ribosome has been solved for the last three steps of translation, the high-resolution structure for the key initiation trGTPase, initiation factor 2 (IF2), complexed with the ribosome, remains elusive. We determine the structure of IF2 complexed with a nonhydrolyzable guanosine triphosphate analog and initiator fMet-tRNA\textsuperscript{Met} in the context of the Escherichia coli ribosome to 3.7-Å resolution using cryo-electron microscopy. The structural analysis reveals previously unseen intrinsic conformational modes of the 70S initiation complex, establishing the mutual interplay of IF2 and initiator transfer RNA (tRNA) with the ribosome and providing the structural foundation for a mechanistic understanding of the final steps of translation initiation.

INTRODUCTION

Initiation of translation is an intensively regulated multistep reaction (1). In bacteria, binding of initiator fMet-tRNA\textsuperscript{Met} to the AUG start codon at the P-site of the 30S subunit defines the correct reading frame of the mRNA. Formation of the 30S initiation complex (30S-IC) is aided by three universal initiation factors (IFs): IF1, IF2, and IF3. The translational guanosine triphosphatase (GTPase) IF2, a key player of the late steps of the initiation reaction (2), facilitates the association of the 30S-IC with the large 50S ribosomal subunit to form the 70S-IC. IF2 is composed of five domains, including a less conserved N-terminal domain followed by the C-terminal domain I to IV (3). Domain I confers GTPase activity, and, together with domains II and III, the ability to bind directly to the 30S subunit (4), whereas the C-terminal domain IV promotes interaction with the initiator fMet-tRNA\textsuperscript{Met} (5, 6). A recent x-ray study of IF2 revealed the structure of all the domains except domain IV (4, 7, 8), and low-resolution cryo-electron microscopy (cryo-EM) maps depict the overall location of IF2 between the ribosomal subunits (8–11). However, high-resolution structural information on ribosome-bound IF2 and its interactions with ribosome and initiator transfer RNA (tRNA) remain elusive, precluding a mechanistic understanding of the functional role of IF2.

RESULTS

To solve the structure of ribosome-bound IF2, the factor was stalled on programmed 70S ribosomes by the nonhydrolyzable GTP analog guanylylimidodiphosphate (GDPNP) in the presence of fMet-tRNA\textsuperscript{Met}. Although previous cryo-EM analysis depicted only one major state (4, 8), multiparticle refinement (12) of our cryo-EM data resulted in two cryo-EM maps for subpopulations of the 70S–fMet-tRNA\textsuperscript{Met}–IF2–GDPNP complex (Fig. 1 and fig. S1). These differ significantly in the rotational state of the ribosomal subunits, the position of the L1 stalk, and the binding mode of the Met-tRNA\textsuperscript{Met}. The smaller subpopulation (70S-IC I; 14,874 particle images, 4.6-Å resolution) is in a fully rotated ribosome state, with the 30S rotated by 8.4° and the 50S head mildly swiveled by 2.4° compared to the classical 70S ribosome conformation (13) (Fig. 1C). The larger subpopulation (70S-IC II: 54,585 particle images, 3.7-Å resolution) is in a semirotated state and undergoes only a moderate intersubunit rotation of 3.7° (Fig. 1F). The near-atomic resolution of major parts of the reconstruction of the 70S-ICII (Fig. 1, G to I, and fig. S2) facilitated atomic model building, aided by existing crystal structures. Although the overall structure is in general agreement with previous cryo-EM analysis (4, 8), there are several pronounced differences. For example, the C-terminal domain IV of IF2 is rotated by ~180°, compared to previous cryo-EM models (4, 8) in our present structure (fig. S3).

Ribosomal intersubunit rotation affects the binding states of initiator fMet-tRNA\textsuperscript{Met} (Fig. 2). Although in both substates the AUG start codon and fMet-tRNA\textsuperscript{Met} anticodon stem loop stay bound in the P-site of the 30S subunit (Fig. 2A), the tRNAs are differentially positioned relative to the 50S subunit (Fig. 2B). The codon-anticodon duplex within the 70S-IC I is similar in position to a canonical P/E tRNA (14, 15), and the elbow of the met-tRNA\textsuperscript{Met} also interacts with the L1 stalk (fig. S4). At first glance, this position of the fMet-tRNA\textsuperscript{Met} elbow is surprising because the 50S E-site is specific for deacylated tRNA (14, 15). However, the tRNA adopts a unique conformation and binding mode because the acceptor stem is bent toward the P-site and the 3′-CCA–fMet end is displaced by more than 20 Å to interact with IF2 domain IV (Fig. 2B). We term this novel chimeric tRNA state of the 70S-IC I complex the P/i state (30S P-site, 50S E-site, and IF2). In the 70S-IC II complex, the initiator tRNA is located closer to the P-site (Fig. 2B), and the tRNA elbow interacts with the central protuberance (CP) of the 50S subunit. Accordingly, we term this the P/pi state. Overall, the position of the P/pi-tRNA is similar but not identical to the P/I-tRNA previously delineated at lower resolution (8). Because the interaction with the L1 stalk, which has been previously reported for the P/I-tRNA, is only observed for the P/pi-tRNA but not for the P/i-tRNA, we may have disentangled a previously inseparable P/I-tRNA into P/pi and P/ei-states (fig. S4).
Relative to the 30S, the fMet-tRNA$_{\text{Met}}$ appears to follow the 30S head swivel, resulting in a small tilting toward the E-site for the P/psi-tRNA compared to the P/psi-tRNA (Fig. 2A). A similar observation has been made by Julián and colleagues (11) based on a lower-resolution cryo-EM map of a 30S-IC containing fMet-tRNA$_{\text{Met}}$ and all three initiation factors. The striking resemblance between the 30S-IC and the equivalent parts of the fully rotated 70S-IC I substate suggests that this state is close to the state immediately after subunit joining. The semirotated 70S-IC II substate in turn resembles an intermediate state on the way to the nonrotated classical elongation state. This assignment is in overall agreement with single-molecule fluorescence resonance energy transfer data reporting an early 70S-IC in the rotated state that is converted into an elongation-competent 70S complex by back rotation of the 30S subunit (16, 17). The fact that we have not detected a nonrotated ribosome state can be explained by the fact that the full back rotation requires GTP hydrolysis by IF2.

Going from the 70S-IC I to the 70S-IC II, both the body of the fMet-tRNA$_{\text{Met}}$ and the core domains of IF2 essentially follow the partial back rotation of the 30S subunit (Fig. 3A). In contrast, domain IV of IF2 is hardly moving, keeping its location in the neighborhood of the 50S peptidyl transferase center and interactions with ribosomal protein UL16 and 23S rRNA helices H89 and H71. Domain IV of IF2 in turn retains the 3′-A-fMet of the initiator tRNA, which is cradled by the two apical β1-β2 and β4-β5 loops (Fig. 3B). The IF2 domain IV/A76-fMet module is linked by flexible regions to the core of IF2 and the body of the tRNA, respectively, to uncouple the motion and to buffer the rotational movement of the 30S subunit. On the IF2 side, a flexible loop connects domain IV that extends from helix 12 of domain III. This is different to the homologous euukaryotic eIF5B, where a more rigid α helix links domain IV to the core of eIF5B, and where both elements together follow the 40S subunit rolling (18) (fig. S5). On the tRNA side, the 3′-CCA end is differently kinked in both states and appears to be dynamic. A defining feature of initiator tRNA is the absence of the first base pair in the acceptor stem. This may support the structural dynamics within the 70S-IC by increasing the length of the single-stranded 3′ end. Thus, the present finding suggests a role for the tRNA$_{\text{Met}}$ C1:A72 mismatch besides its functional role during formylation (19, 20).

Although there may be still some ambiguity in the exact details, the single-residue resolution of the 70S-IC II cryo-EM map provides insight into the contact between the IF2 domain IV and the CCA-fMet end of the tRNA (fig. S6). The terminal fMet-A76 is bound in a pocket formed by Ser$^{844}$, Arg$^{846}$, Arg$^{847}$, Phe$^{848}$, and Lys$^{849}$ from the β4-β5 loop; Phe$^{804}$, Phe$^{809}$, Ile$^{812}$, and Cys$^{815}$ from the β1-β2 loop; and Glu$^{860}$, Cys$^{861}$, and Gly$^{862}$ from β5 (Fig. 3B). The interaction of A76 with IF2 domain IV seems to resemble closely the interaction of the homologous elongation factor Tu (EF-Tu) domain II with elongator
tRNAs; the fMet is cradled by Arg847 and Phe848 and seems to be positioned by backbone interactions between the formyl group and the Na of Phe848. These binding site interactions are in agreement with biochemical and nuclear magnetic resonance spectroscopic data (21) and rationalize why the introduction of a partially negative charge from the formyl group leads to a 170-fold increased rate of initiation (22).

Our model also explains the role of Cys861 and Cys815 that presumably stabilize the backbone geometry of the binding pocket through a disulfide bond (23), and further rationalizes that mutations of Gly862 to larger amino acids substantially reduce the binding affinity of IF2 to fMet-tRNAMet (21) because this would lead to a clash with the ligand (fig. S6).

In both 70S-IC complexes, the IF2 core domains I to III are bound in the classical factor binding site (Fig. 1) in an overall similar configuration to EF-G and EF-Tu (24). Switch I (sw I) and sw II are well defined and found in a different configuration compared to recent x-ray structures of the isolated IF2 (7) (Fig. 3C). The sarcin-ricin loop [SRL; helix 95 of 23S rRNA] of the 50S subunit is bound at the interface between domains I and III. The universally conserved Tyr772 within IF2 helix 12 of domain III is in close proximity to the SRL, similar to the eukaryotic system (18, 25). Remarkably, the partial back rotation of the 30S subunit brings the G-domain of IF2, in particular His448 of the sw II loop, closer to the SRL when transitioning from the 70S-IC I to the 70S-IC II. A contact between His448 (His84 in EF-Tu) of the sw II loop and the SRL has been deemed important for GTP hydrolysis of EF-G (26–28) (fig. S7). Therefore, GTP hydrolysis may be regulated by the back rotation of the 30S subunit. Our results indicate how IF2, by modulating the state of fMet-tRNAfMet and by exploiting spontaneous intersubunit rotation of the ribosome, facilitates joining of the large subunit and the subsequent transition into an elongation-competent 70S complex (Fig. 4). As proposed (10), fMet-tRNAfMet and IF2 aid subunit joining by creating a large scaffold at the subunit interface and providing contact regions with the 50S subunit. However, similar to the eukaryotic system (18), a major problem during initiation is not subunit joining per se but subunit joining with the 30S-IC, where initiator tRNA is already bound to the P-site of the small ribosomal subunit. The space between helix 69 of the 23S rRNA and uL5 of the CP of the 50S subunit is too narrow for the tRNA to pass through in its canonical conformation (fig. S8). Our results suggest how this problem is solved during bacterial initiation. Aided by IF2, the initiator tRNA is first bound to the 50S E-site region in the fully rotated 70S-IC I. Back rotation of the 30S...
subunit then facilitates a partial reverse translocation of initiator tRNA relative to the 50S subunit (Fig. 4). A comparison to our mammalian initiation complex (18) shows that the late steps of initiation use large-scale conformational changes of the ribosome in both domains of life. Nevertheless, there are striking differences in the molecular strategy to load initiator tRNA. Whereas the eukaryotic initiation uses eukaryotic-specific subunit rolling (18, 29), the bacterial IF2-dependent initiation uses subunit rotation. The unique P/ei and P/pi states of the tRNA require the tRNA to adopt a specific conformation to simultaneously undergo codon-anticodon interaction and interactions with domain IV of IF2. Specific features of the initiator tRNA appear to facilitate these particular intermediate states. Thus, the P/ei and P/pi states presumably resemble checkpoints to ensure binding of the correct tRNA. In this sense, the complex pathway of IF2-promoted subunit joining may not only ensure efficient loading of the fMet-tRNA\textsubscript{Met} into the 50S P-site but may also be used to provide discriminative power for accurate selection of initiator over elongator tRNAs.

**MATERIALS AND METHODS**

**Methods**

**Formation of the 70S–fMet-tRNA\textsubscript{Met}–IF2–GDPNP complex.** His-tagged IF2 from *E. coli* was produced using the T5 expression system. Recombinant IF2 was isolated by affinity chromatography and ion-exchange chromatography. The mRNA with the sequence 5′-GGGCAAAACAAAAAGGAGCCUAUAUAUGUUCUAGCAAAACA- AAACAAAAGAUU-3′ was transcribed using the T7-MEGAshortscript (Ambion). The mRNA contains a strong Shine-Dalgarno sequence (underlined) and the initiator AUG start codon (bold). The fMet-tRNA\textsubscript{Met} was made from tRNA\textsubscript{Met} (Sigma). The tRNA was aminocylated and N-formylated as described previously (30) and purified via Nucleosil C\textsubscript{4} column (Macherey-Nagel). Tightly coupled 70S ribosomes from the *E. coli* CAN20 strain were purified as described (31). The fMet-tRNA\textsubscript{Met}–IF2–GDPNP complex was formed by incubating 24-pmol ribosomes with 400-pmol mRNA, 100-pmol tRNA, and 115-pmol IF2 in the presence of 50 nmol of the nonhydrolyzable GTP analog GDPNP at 37°C for 15 min in Buffer B15 [20 mM Heps-KOH (pH 7.5), 15 mM magnesium acetate, 150 mM potassium acetate, 4 mM β-mercaptoethanol, 2 mM spermidine, and 0.05 mM spermine]. The occupancy of IF2 in the complex was about 40 to 50%, as determined by Western blotting against the 6xHis tag of IF2 and the uL1 protein following a sucrose density centrifugation.

**Cryo-EM and image processing.** Ribosomal complexes were diluted to a concentration of 30 nM and flash-frozen in liquid ethane on glow-discharged carbon-coated grids (Quantifoil Micro Tools GmbH) using a Vitrobot device (FEI). In total, two sets of images were collected under low-dose conditions (20 e\textsuperscript{−}Å\textsuperscript{2}) on a Tecnai G2 Polara Microscope (FEI) operating at 300 kV and equipped with a K2 Summit detector (Gatan) at ×31,000 nominal magnification in SuperResolution mode resulting in a pixel size of 0.615 Å per pixel. Data were collected automatically using Leginon (32) at a defocus range of 0.5 to 5 μm.

The initial contrast transfer function defocus values for the micrographs were estimated from the contrast transfer function using CTFFind4 (33). Ribosomal projection images were (semi)automatically identified with SIGNATURE (34) and e2boxer from the EMAN2 software package (35). In total, 232,659 particle images were selected from 3715 micrographs. The data set was individually aligned by using a low-resolution 70S from *E. coli* as reference using the multiparticle refinement (12) in SPIDER (36). After initial refinement, bad particle images were removed. During unsupervised classification, the data set split into structures showing IF2 bound to ribosomes in two different conformations. These two conformations were isolated from each other and refined separately. Particle images that did not contain the two large ligands IF2 and fMet-tRNA\textsubscript{Met} were selected out and not considered for further analysis. After interpolation to a pixel size of 1.23 Å per pixel, a total number of 14,874 particle images corresponding to the fully rotated conformation were refined using SPIDER (36) and SPARX (37). A total number of 54,585 particle images corresponding to an intermediate rotated conformation were interpolated to a pixel size of 1.025 Å per pixel and refined using SPIDER (36) and SPARX (37).

**Structural models.** Existing atomic models of the *E. coli* ribosome were adjusted to the derived densities. Initial rigid body docking of x-ray structures of the ribosomal 50S subunit [PDB ID 4V9O for subpopulation II and 3R8S for subpopulation I (13)], 30S head, and 30S body [4GD1 for subpopulation I, 4GD2 for subpopulation II (13)] was performed using UCSF Chimera (38). Flexible elements of the docked models were subsequently adjusted to the derived densities in COOT (39). The near-atomic resolution of substate II enabled us to build a model for IF2 in COOT (39) based on a homology model composed of the core region of the solution structure of *T. thermophilus* IF2 (7) and the C-terminal domain IV from *Bacillus stearothermophilus*.
(40). Subsequently, the model was globally idealized using REFMAC (41) and validated with MolProbity (42). For the AUG start codon and fMet-tRNA\textsuperscript{Met} (PDB 5AFI), two rigid units comprising either the ASL and AUG start codon or the acceptor arm, together with the tRNA elbow (TP- and D-loop), were combined by allowing flexibility around ASL-tRNA elbow junction, a flexible hinge region that was already observed when tRNAs adopt hybrid A/T or P/E tRNA binding states (13, 43–45). The single-stranded CCA end was built manually into the derived density in COOT (39), and the initial fitting was globally idealized using ERRASER (46). The structure of poorly resolved, flexible loops in IF2 was predicted by FragFit, a fragment-based tool for modeling of missing protein segments into cryo-EM density maps (see section below). These initial models were then refined by real-space refinement in Phenix 1.10. To prevent overfitting of the weight of the map versus model, geometry was estimated using cryo-EM maps from half-sets as introduced for reciprocal space refinement (47, 48) (fig S2B).

The derived models of IF2 and the mRNA tRNA module from subpopulation II were subsequently docked as rigid bodies into the map of subpopulation I.

Validation with FragFit. The FragFit database (http://proteininformatics.charite.de/fragfit) is constructed from about 90,000 protein structures, deposited in PDB (as of June 2013). Extraction of all overlapping peptide fragments with lengths between 3 and 35 amino acids yielded more than 700 million fragments. For each fragment, the sequence, PDB identifier, and a geometrical fingerprint of the stem atoms are deposited in PDB (as of June 2013). Extraction of all overlapping pep-

tides as introduced for reciprocal space refinement (47, 48) (fig. S2B).

The derived models of IF2 and the mRNA tRNA module from subpopulation II were subsequently docked as rigid bodies into the map of subpopulation I.

Validation with FragFit. The FragFit database (http://proteininformatics.charite.de/fragfit) is constructed from about 90,000 protein structures, deposited in PDB (as of June 2013). Extraction of all overlapping peptide fragments with lengths between 3 and 35 amino acids yielded more than 700 million fragments. For each fragment, the sequence, PDB identifier, and a geometrical fingerprint of the stem atoms are deposited in PDB (as of June 2013). Extraction of all overlapping pep-

tides as introduced for reciprocal space refinement (47, 48) (fig. S2B).

The derived models of IF2 and the mRNA tRNA module from subpopulation II were subsequently docked as rigid bodies into the map of subpopulation I.

Validation with FragFit. The FragFit database (http://proteininformatics.charite.de/fragfit) is constructed from about 90,000 protein structures, deposited in PDB (as of June 2013). Extraction of all overlapping peptide fragments with lengths between 3 and 35 amino acids yielded more than 700 million fragments. For each fragment, the sequence, PDB identifier, and a geometrical fingerprint of the stem atoms are deposited in PDB (as of June 2013). Extraction of all overlapping pep-

tides as introduced for reciprocal space refinement (47, 48) (fig. S2B).

The derived models of IF2 and the mRNA tRNA module from subpopulation II were subsequently docked as rigid bodies into the map of subpopulation I.

Validation with FragFit. The FragFit database (http://proteininformatics.charite.de/fragfit) is constructed from about 90,000 protein structures, deposited in PDB (as of June 2013). Extraction of all overlapping peptide fragments with lengths between 3 and 35 amino acids yielded more than 700 million fragments. For each fragment, the sequence, PDB identifier, and a geometrical fingerprint of the stem atoms are deposited in PDB (as of June 2013). Extraction of all overlapping pep-

tides as introduced for reciprocal space refinement (47, 48) (fig. S2B).

The derived models of IF2 and the mRNA tRNA module from subpopulation II were subsequently docked as rigid bodies into the map of subpopulation I.

Acknowledgments: We thank T. Hilal, E. Behrmann, A. Schmidt, J. Kalms, and J.T. Christiansen for the helpful discussions. The plasmid for IF2 was provided by K. Nierhaus. We also thank R. Albrecht for the preparation of the reassociated 70S ribosomes. We further acknowledge the North-German Supercomputing Alliance (HLRN) for providing computing resources that have contributed to the reported research results. Funding: This work was supported by grants from the German Research Foundation (DFG) Forschergruppe 1805 to C.M.T.S.; SBF740 to C.M.T.S., T.M., P.S., and P.W.H.; SBF1078 to P.S.; B893/8 to P.W.H.; and the DFG Cluster of Excellence Unifying Concepts in Catalysis Research Field D3/E3-1 to P.S. Author contributions: T.S., J.D.F.R., and V. Ramakrishnan designed the experiments. K.Y., prepared Met-tRNA\textsuperscript{Met}. T.S. prepared the 70S–Met-tRNA\textsuperscript{Met}–IF2–GDPNP complex. J.B and T.M collected cryo-EM data. T.S., J.L., and C.M.T.S. did the image processing. T.S., H.Y., D.J.F.R., and C.M.T.S. discussed the results and wrote the paper. Competing interests: The authors declare that they have no competing interests. Data and materials availability: The electron density maps and models of the 70S-IC complexes have been deposited in the 3D-EM and PDB database with accession numbers EMD-6559 (70S-IC II) and EMD-3285 (70S-IC I) and PDB IDs 3JCN (70S-IC I) and 236 (70S-IC I). All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 23 October 2015
Accepted 12 January 2016
Published 4 March 2016
10.1126/sciadv.1501502

Structures of ribosome-bound initiation factor 2 reveal the mechanism of subunit association
Thiemo Sprink, David J. F. Ramrath, Hiroshi Yamamoto, Kaori Yamamoto, Justus Loerke, Jochen Ismer, Peter W. Hildebrand, Patrick Scheerer, Jörg Bürger, Thorsten Mielke and Christian M. T. Spahn

Sci Adv 2 (3), e1501502.
DOI: 10.1126/sciadv.1501502