

OPEN

Structural analysis of 70S ribosomes by cross-linking/mass spectrometry reveals conformational plasticity

Christian Tüting^{1,7}, Claudio Iacobucci^{2,6,7}, Christian H. Ihling^{2,3}, Panagiotis L. Kastiris^{1,4,5} & Andrea Sinz^{2,3}✉

The ribosome is not only a highly complex molecular machine that translates the genetic information into proteins, but also an exceptional specimen for testing and optimizing cross-linking/mass spectrometry (XL-MS) workflows. Due to its high abundance, ribosomal proteins are frequently identified in proteome-wide XL-MS studies of cells or cell extracts. Here, we performed in-depth cross-linking of the *E. coli* ribosome using the amine-reactive cross-linker disuccinimidyl diacetic urea (DSAU). We analyzed 143 *E. coli* ribosomal structures, mapping a total of 10,771 intramolecular distances for 126 cross-link-pairs and 3,405 intermolecular distances for 97 protein pairs. Remarkably, 44% of intermolecular cross-links covered regions that have not been resolved in any high-resolution *E. coli* ribosome structure and point to a plasticity of cross-linked regions. We systematically characterized all cross-links and discovered flexible regions, conformational changes, and stoichiometric variations in bound ribosomal proteins, and ultimately remodeled 2,057 residues (15,794 atoms) in total. Our working model explains more than 95% of all cross-links, resulting in an optimized *E. coli* ribosome structure based on the cross-linking data obtained. Our study might serve as benchmark for conducting biochemical experiments on newly modeled protein regions, guided by XL-MS. Data are available via ProteomeXchange with identifier PXD018935.

Ribosomes, the molecular machines that are responsible for protein synthesis, have frequently attracted interest both from a biological¹ as well as from a methodological perspective^{2–4}. Due to their high abundance in cells, they are straightforward to isolate⁵, making them an ideal specimen for developing X-ray crystallography^{6–9} and electron microscopy (EM) methods². Ribosomes have been used in EM to develop positive and negative staining methods¹⁰, evaluate the efficiency of direct electron detectors¹¹, and they have served as a model system to develop image processing protocols². Due to their central role in protein translation, ribosomes possess an immense biological value, recognized by the award of the 2009 Chemistry Nobel Prize to Ramakrishnan, Steitz and Yonath “for studies of the structure and function of the ribosome”. Structural insights into ribosomes from various organisms, in a variety of conformations¹² and higher-order states^{13,14}, in complex with other biomolecules^{15,16}, inhibitors or antibiotics^{17,18} have been reported. These studies were influential to a detailed understanding of this highly intricate and large ribonucleoprotein machine^{1,19}.

¹Interdisciplinary Research Center HALOmem, Charles Tanford Protein Center, Martin Luther University Halle-Wittenberg, Kurt-Mothes-Str. 3a, 06120 Halle/Saale, Germany. ²Department of Pharmaceutical Chemistry and Bioanalytics, Institute of Pharmacy, Charles Tanford Protein Center, Martin Luther University Halle-Wittenberg, Kurt-Mothes-Str. 3a, 06120 Halle/Saale, Germany. ³Center for Structural Mass Spectrometry, Charles Tanford Protein Center, Martin Luther University Halle-Wittenberg, Kurt-Mothes-Str. 3a, 06120 Halle/Saale, Germany. ⁴Institute of Biochemistry and Biotechnology, Martin Luther University Halle-Wittenberg, Kurt-Mothes-Str. 3, 06120 Halle/Saale, Germany. ⁵Biozentrum, Martin Luther University Halle-Wittenberg, Weinbergweg 22, 06120 Halle/Saale, Germany. ⁶Present address: Corporate Preclinical R&D, Analytics and Early Formulations Department, CHIESI FARMACEUTICI S.P.A., Via Palermo 26/A, 43122 Parma, Italy. ⁷These authors contributed equally: Christian Tüting and Claudio Iacobucci. ✉email: panagiotis.kastiris@bct.uni-halle.de; andrea.sinz@pharmazie.uni-halle.de

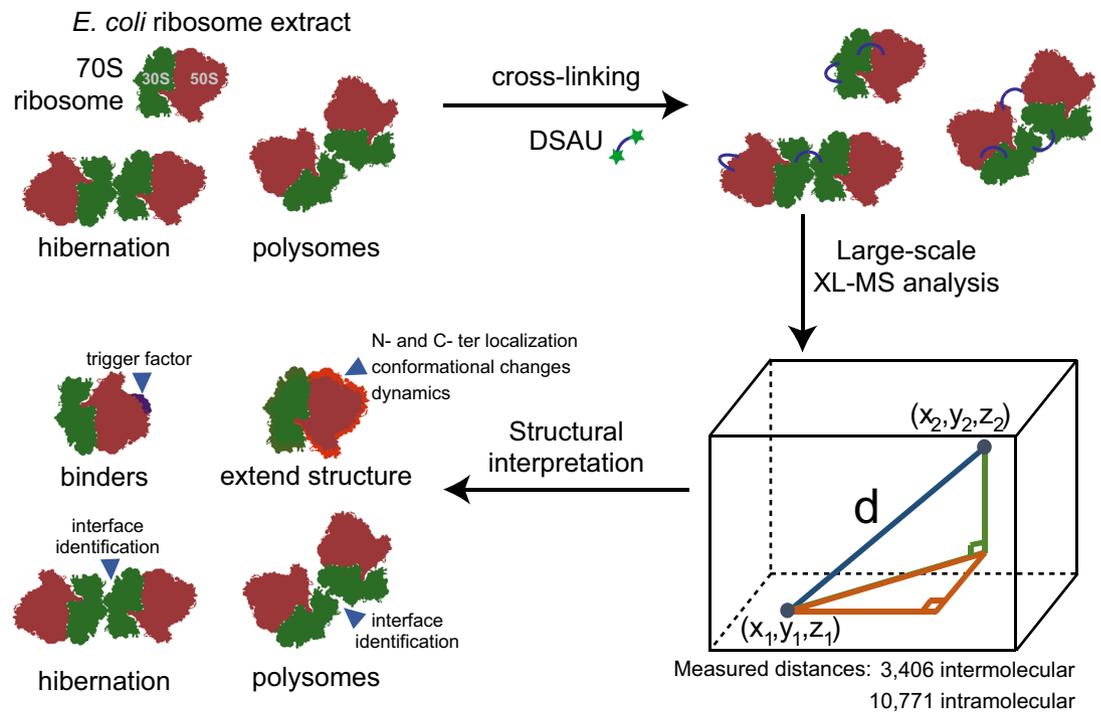


Figure 1. Workflow. We used the DSAU cross-linker with the 70S *E. coli* ribosome to identify binders, flexible regions, higher-order states and modeling of regions previously unresolved using derived data and structure-wide available data in the Protein Data Bank (<https://www.rcsb.org/>).

The *E. coli* ribosome has frequently been used in proteome-wide cross-linking/mass spectrometry methods (XL-MS) as a testbed for structural mapping of cross-linked peptides, and for determining the overall fit to the ribosome structure^{20–22}. Because ribosomes are abundant in the cell, most of the cross-linking reactions in cell extracts or in vivo are concentrated on the ribosome. Therefore, a large number of cross-linked peptides are available for evaluating cross-linking efficiency. As an example, 207 high-resolution atomic structures of *E. coli* ribosomes are currently deposited in the PDB database (date: February 1, 2020; resolution < 4.5; < 25 entities per biological unit). Surprisingly, only 87% of the ribosome's protein content has been structurally characterized in these complexes, meaning that more than 1,000 residues are highly flexible and impossible to recover in poor or absent electron densities.

XL-MS would allow providing structural information for all regions of the ribosome, including ordered as well as flexible regions, resulting in a detailed understanding of the overall ribosomal architecture. XL-MS has matured from a method to study isolated biomolecules into a proteome-wide method to understand cellular protein interactions^{23,24}. Towards this goal, novel cross-linkers have been developed to optimize the discovery of cross-linked peptides with sensitive LC/MS/MS protocols^{20,22,25,26} and attempts have been made to estimate false discovery rates (FDR) for cross-links²⁷. However, an important validation for proteome-wide XL-MS is the mapping of cross-links on structural models. There is still development in the field regarding the distance measure that is best applicable for mapping cross-links, i.e., Euclidian versus surface-exposed distance²⁸. Also, the usefulness of FDR calculations and their correlation to true positive protein–protein interactions and the corresponding structural models²⁹ as well as the choice of molecular models that are used for cross-link mapping are in some cases suboptimal. This is because (a) current studies have a bias for high-abundant proteins, but methods to address this issue are being implemented³⁰ and (b) few molecular models deposited in structure databases are being evaluated for cross-linking distances, and only one molecular model is selected for distance calculation per protein complex, despite the wealth of structural data. This obvious limitation subsequently confines recovered results, which could be of relevance for the structural biology and function of the protein of interest.

Here, we apply the *N*-hydroxysuccinimide (NHS) ester diacetyl dibutyric urea (DSAU) as an amine-reactive, urea-based MS-cleavable cross-linker³¹. It possesses a spacer length of ~ 10 Å and is shorter than the widely-used cross-linker DSBU. Therefore, DSAU allows a distinct subset of distance constraints during structural mapping and/or subsequent modeling to be measured. We cross-linked the *E. coli* ribosome with DSAU and recovered 126 intra- and 97 intermolecular cross-links at an FDR of 1%. We then comprehensively mapped *all* identified cross-links onto *all* suitable *E. coli* ribosomal structures. Based on the satisfaction of those data, we eventually (a) remodeled flexible protein regions and discovered underlying conformational plasticity, (b) localized the ribosome-associated chaperone, the trigger factor, (c) completed protein structures with additional residues and domains, and (d) unveiled higher-order ribosome states. We finally highlight the broad synergy of XL-MS with high-resolution structural methods, as our XL-MS experiments allowed remodeling of 2,057 residues in total, optimizing the current working model of the *E. coli* ribosome (Workflow, Fig. 1).

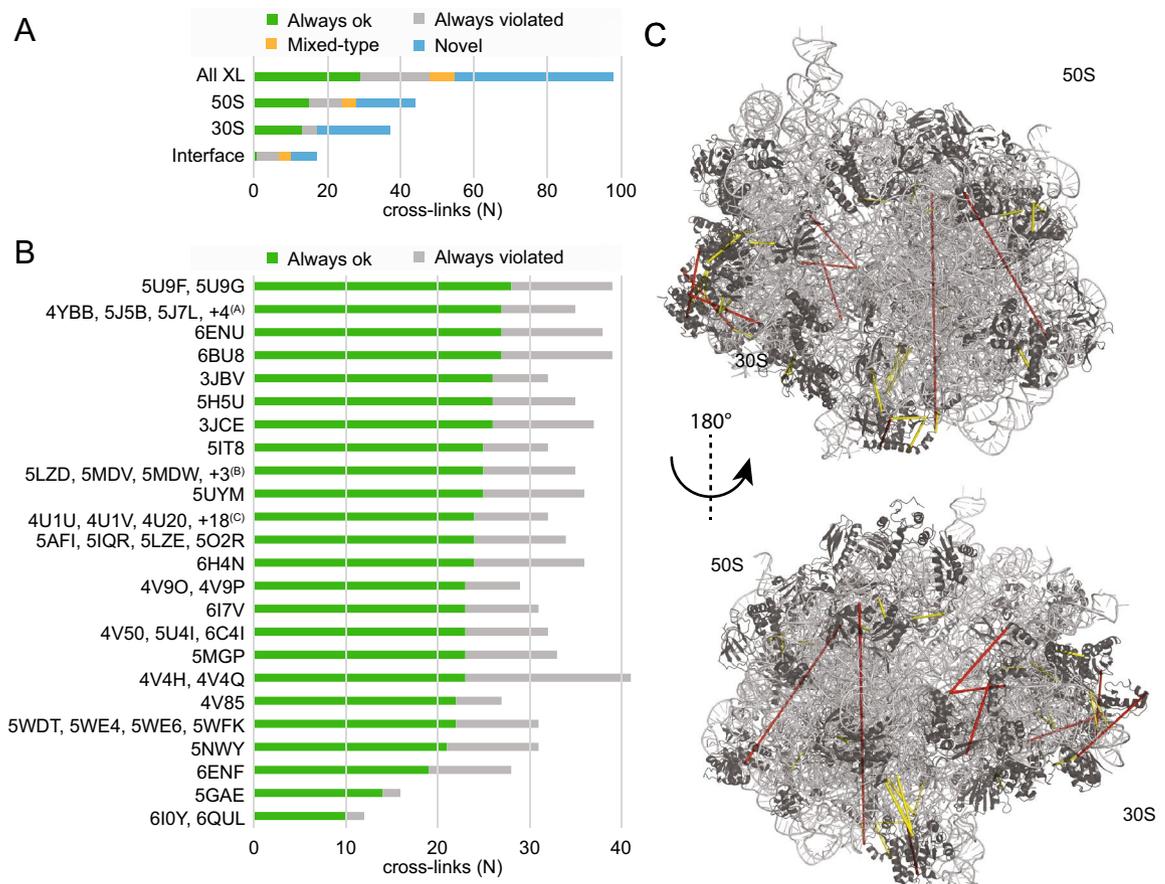


Figure 2. Statistics of database-wide mapping of ribosomal inter-molecular cross-links. **(A)** Classification of cross-links. Satisfied: cross-links always in the allowed range; Mixed-typed: cross-links both in the allowed and in the violated range, depending on the ribosome structure; Violated: cross-links always outside the allowed range; Novel: cross-links that could not be mapped on any high resolution ribosomal structure. **(B)** Classification of mappable cross-links on published structures. PDB IDs (<https://www.rcsb.org>) with the same count of satisfied and violated cross-links are grouped. Pruned pdb-codes: ^(A)5J88, 5J8A, 5J91, 5JC9; ^(B)5MDY, 5MDZ, 6HRM; ^(C)4U24, 4U25, 4U26, 4U27, 4V52, 4V54, 4V57, 4V64, 4V6C, 4V7S, 4V7T, 4V7U, 4V7V, 4V9C, 4V9D, 4WF1, 4WOI, 4WWW. **(C)** Cross-links on 5U9F, the published structure with the highest number of satisfied cross-links. Cross-links that satisfy the distance range are shown in yellow dashed lines; distances that are violated are shown in red. Ribosomal proteins (black) and rRNA (grey) are shown in cartoon presentation. The images of protein structures were generated with PyMOL (Version 2.3.2), Schrödinger, LCC. URL: <https://pymol.org>.

Results

Large-scale mapping of cross-links on *E. coli* ribosome structures. We have systematically recovered 126 intramolecular cross-links and 97 intermolecular cross-links at 1% FDR, mapping 47 proteins and 36 protein–protein interactions in the *E. coli* ribosome (Table S1). We first analyzed the intramolecular cross-links. 126 non-redundant intramolecular cross-links were mapped onto 142 ribosome structures, while only 11% (14/126) of these cross-links could not be mapped to any available structure (Table S2). The cross-links show a distance distribution (d_{intra}) that recapitulates the length of the DSAU cross-linker [$N = 10,769$; median(d_{intra}) = 13.4 Å]. As intramolecular ribosomal cross-links, except for L7/L12, do not involve protein–protein interactions, we set a threshold of 30 Å for evaluating those cross-links. By applying this threshold, 91.1% (102/112) of the mapped intramolecular cross-links were satisfied. As only intermolecular cross-links capture protein–protein-interactions in the ribosome, the intramolecular cross-links were not further analyzed, except for the interesting case of L7/L12, which is the only known multimeric protein in the *E. coli* ribosome.

For the 97 intermolecular cross-links, we identified 69 ribosomal structures with at least one intermolecular cross-linking pair (Table S3). The cross-links show a distance distribution (d_{inter}) higher than that of intramolecular cross-links, clearly indicating flexibility in the respective ribosomal protein–protein interactions [$N = 3,304$; median(d_{inter}) = 20.28 Å]. Out of the 97 cross-links, 54 could be mapped on ribosomal structures, while 44 were novel (Fig. 2A), signifying that the residues involved in the cross-linking reaction are not represented in any published ribosome structure (< 3.5 Å resolution). We systematically evaluated the threshold of the DSAU cross-linker for distance measurements of the intermolecular cross-links detected (Fig. S1). Interestingly, the number

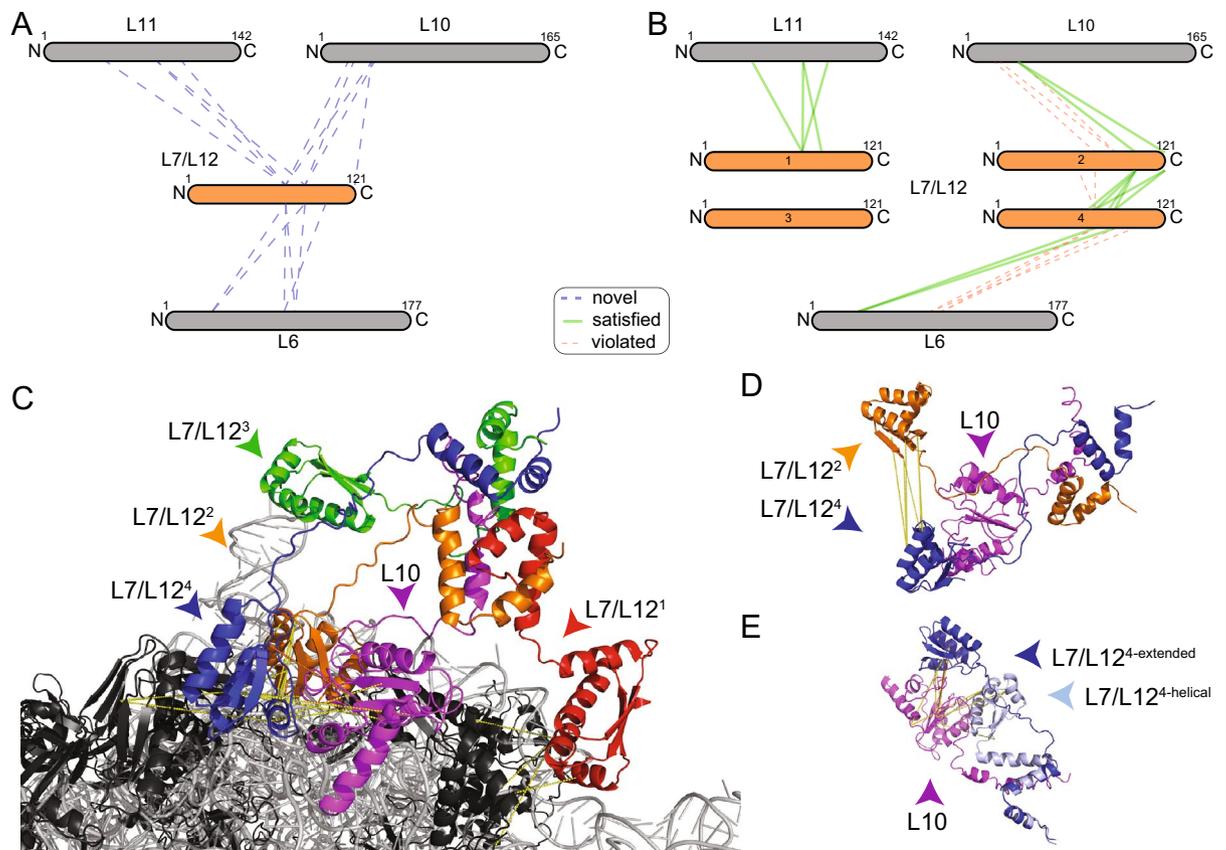


Figure 3. The dense interaction network of L7/L12 captured by XL-MS. Ribosomal proteins (black) and rRNA (grey) are shown in cartoon. Cross-linked proteins are color highlighted. Cross-links are shown as dotted lines, satisfaction (yellow) or minor violation (orange) are indicated. **(A)** Interaction network of L7/L12. Sizes of the boxes are scaled to the length of the protein sequences; each dotted line represents one identified novel cross-link pair. **(B)** Satisfied cross-links after remodeling; Only 4 cross-links of L6 to L7/L12 could not be unambiguously mapped (not shown); In addition, 5 intra-links show violation in the monomer, but satisfaction in the L7/L12 multimer. **(C)** Model of the interaction of an L7/L12 tetramer (blue, green, red, orange) bound to the helix 8 of L10 (magenta) based on the cross-linking information (see text). **(D)** Satisfaction of intra-molecular cross-links between CTD domains of L7/L12. **(E)** Inter-molecular cross-links satisfying both extended and helical conformation of the linker region of L7/L12; Note that all cross-links are satisfied in the helical linker conformation, and all but one cross-link are satisfied in the extended conformation. The images of protein structures were generated with PyMOL (Version 2.3.2), Schrödinger, LCC. <https://pymol.org>.

of violated cross-links remains constant up to a threshold of 40 Å (Fig. S1). Therefore, we chose a distance threshold of 37.5 Å as this (a) accounts for the extensive conformational flexibility of the ribosome (it is 150% higher compared to the atomic distance of 25 Å which is the maximum Ca–Ca distance of cross-linked Lys residues), (Fig. S2) and (b) does not influence the recovery of higher false positive numbers (Fig. S1).

Therefore, 65% (35 out of 54) of the mapped cross-links are satisfied in at least one ribosomal structure (Fig. 2B). In particular, the ribosomal structure of the “ArfA–RF2 ribosome rescue complex” (5U9F)³² has a wide coverage of satisfied cross-links, amounting for 28 non-redundant residue pairs in total (Fig. 2B,C). Eventually, we are able to explain, clarify, and satisfy 89% (17 out of 19) of the violated cross-links, and therefore, confidently map >95% of cross-links to the available ribosome structures.

Plasticity of ribosomal active sites recapitulated by XL-MS. *Translation initiation and codon probing.* For efficient translation initiation, the 30S protein S21 is essential and consists of two α-helices, folding around the bound mRNA (PDB ID: 5U9F)³². Five cross-links were identified in S21, of which three map to the first α-helical region (residues 1–37) and two to the second (residues 38–71). Due to the localization of a violated cross-link (#86) in the second α-helix and the accessible space in the helical environment, only a planar movement of the second helical region can satisfy this cross-link. We performed a cross-link-driven reorientation of the second α-helix by molecular modeling, keeping the first α-helix rigid. By this, we could generate a model of S21, satisfying the violated cross-link in the second helix.

After initiation, aminoacyl-tRNA molecules must then be probed by trial and error, a procedure facilitated by the 50S ribosomal protein L7/L12, the only multimeric protein in the *E. coli* ribosome³³. Due to its flexible nature, it is only partially resolved in few published structures (PDB IDs: 4V85³⁴, 4V90³⁵, 6I0Y³⁶), however it

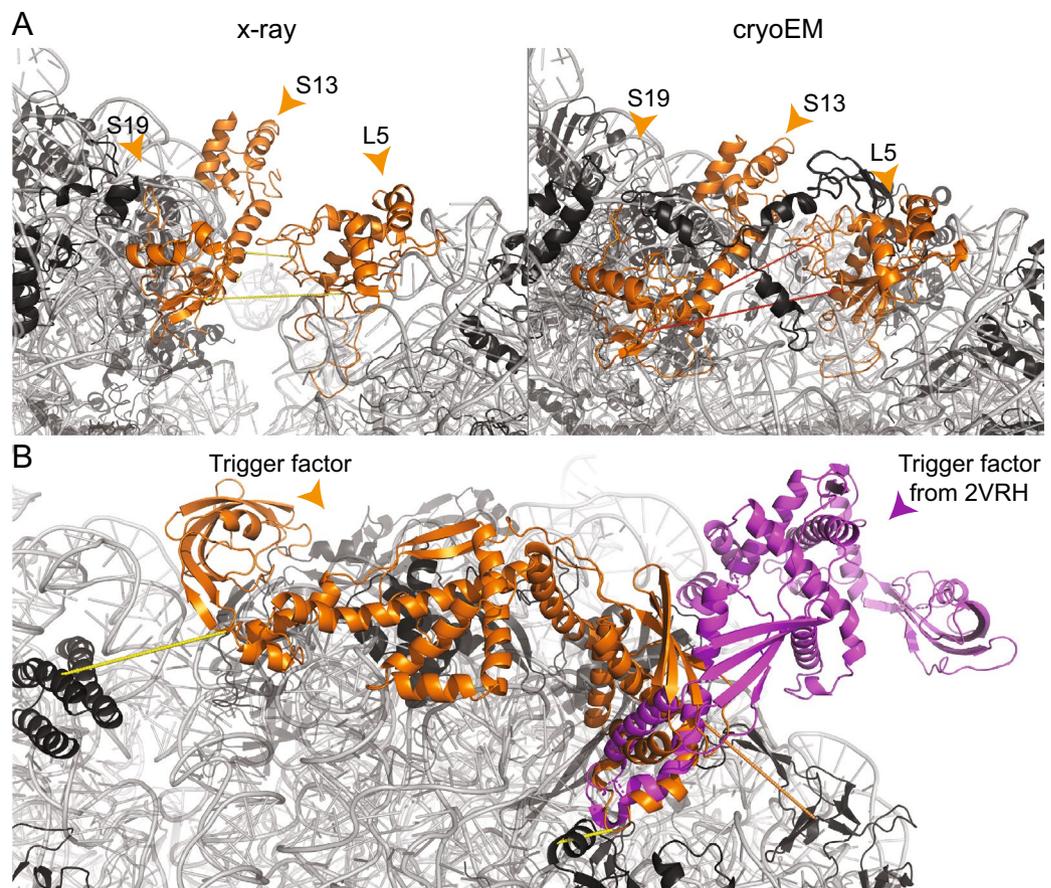


Figure 4. Visualization of differences in X-ray and cryo-EM structures, identified by cross-links falling in the mixed-type category, and localization of the trigger factor. Ribosomal proteins (black) and rRNA (grey) are shown in cartoon. **(A)** Capturing the flexibility between the 30S and 50S subunits. In the X-ray structure (4V4H), cross-links #62 and #73 are satisfied (yellow dashed lines); in the cryo-EM-structure (5U9F), a different conformation is visible, which violates the maximum distance (red dashed lines). **(B)** Structural model of the trigger factor (orange) in complex with the *E. coli* ribosome. Cross-link satisfaction is indicated as yellow lines, violation as orange lines. The known low-resolution structure (2VRH) is superimposed and shown in magenta. The images of protein structures were generated with PyMOL (Version 2.3.2), Schrödinger, LCC. <https://pymol.org>.

was found to be highly cross-linked (Fig. 3A). It is known that in *E. coli* up to four copies of L7/L12 can bind to a single ribosome³⁷. We identified 15 novel interactions involving L7/L12 that include three different proteins, L6, L10 and L11 (Fig. 3A), while one intermolecular and 21 intramolecular cross-links are located in L7/L12 itself. In addition to cross-links which clearly indicate that L7/L12 is multimeric, the intramolecular cross-links may also reveal a higher-order oligomerization state. For this, we mapped the intramolecular cross-links on the published monomeric structure of the C-terminal (C-ter) domain of L7/L12 (PDB ID: 1CTF)³⁸ and measured both Euclidian and solvent accessible distances using the xWalk software³⁹ (Table S4). Despite the Euclidian distances of all cross-links being in range ($< 30 \text{ \AA}$), the surface accessible distance for eight cross-links was above 30 \AA , pointing to a higher-order state.

Considering this information, we generated a tetrameric model of L7/L12 binding to the ribosome. We generated a tetramer because it is known that in *E. coli*, up to four copies of L7/L12 can bind to a single ribosome and was previously partially resolved in a tetrameric state³⁴. Our extended model satisfies 60% of the novel cross-links to other proteins (Fig. 3B,C) and 6 out of 8 intermolecular cross-links between copies of itself (Fig. 3D). In this model, different degrees of flexibility of the linker are visible: For one of the L7/L12 units, the cross-link pattern could be satisfied by either an α -helical or an extended fold (Fig. 3E).

Frameshifting and ratchet movement. Ribosome translocation is accompanied by the ratchet movement⁴⁰, while inhibitors can interfere with the ratchet movement of the ribosome. This plasticity is captured by two cross-links in the interface between 30 and 50S subunits: L5 was cross-linked to S13 and S19 in a mixed-type manner (Fig. 4A). Interestingly, cryo-EM structures resolve an open conformation, as indicated by a cross-link minor violation ($38\text{--}40 \text{ \AA}$), while X-ray structures capture various states of the ratchet movement. As an example, the crystal structure of the *E. coli* ribosome in complex with kasugamycin (PDB ID: 4V4H)⁴¹ was resolved with two ribosomes in the asymmetric unit with a variability in the 50S-30S interface: We mapped the cross-links and

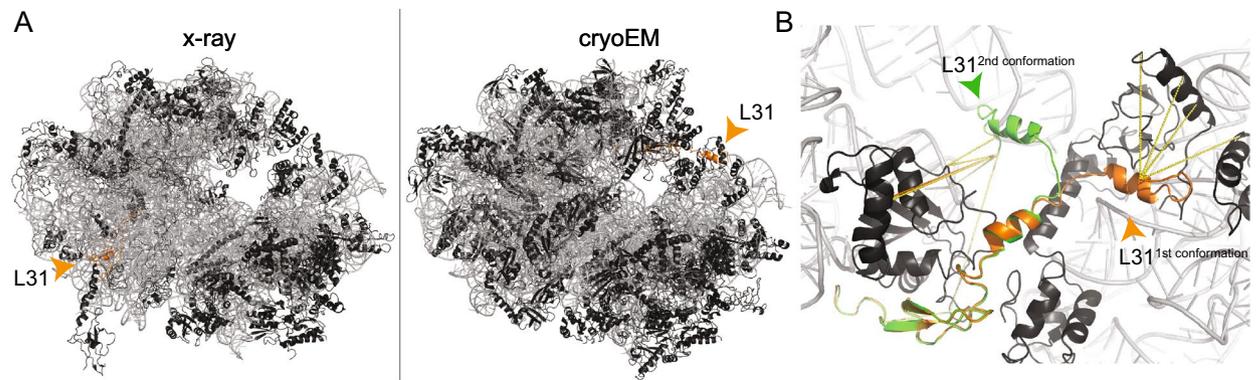


Figure 5. Localization and flexibility of the L31 ribosomal protein. Ribosomal proteins (black) and rRNA (grey) are shown in cartoon. L31, involved in cross-linking, is highlighted in orange. **(A)** Annotation discrepancy of L31 in the crystal structure (4V4H) in comparison to a recent cryo-EM structure (5U9F). The orange arrow indicates the annotated localization at the ribosomal structure. **(B)** Two distinct conformations of the CTD of L31. In orange, the first conformation is shown, where cross-links #66, #71–72 and #74 are satisfied. In green, the second conformation is shown, where cross-links #27–30 and #39 are satisfied and cross-link #38 (orange line) is only minorly violated (38.8 Å). Cross-link #26, located in the NTD, is satisfied in both conformations. The images of protein structures were generated with PyMOL (Version 2.3.2), Schrödinger, LCC. <https://pymol.org>.

discovered conflicting behavior of measured distances for each of those ribosomes (#62, 18.9 Å and 34.5 Å; #73, 24.2 Å and 41.2 Å). This indicates, that even in crystallized ribosomes, different conformational states could be existing, underlining, again, the high flexibility.

Flexibility and rotation of the L1 stalk and unfolding of the mRNA by S3. The L1 stalk is a flexible substructure consisting of the ribosomal protein L1 and helices 76–78 of the 23S rRNA. We observed five cross-links from L1 to the L9 protein and one cross-link from L1 to the L33 protein. The available structures satisfy five out of six cross-links; the remaining cross-link (#1; 63.3 Å) reflects the known extensive conformational variability of the L1 stalk, including movements of up to 20 Å⁴², which can be sampled only with extensive molecular dynamics simulations in combination with higher amount of data⁴³. In addition to the L1 stalk movement, mRNA needs to be single-stranded during translation. The 50S protein S3 at the mRNA entrance tunnel is directly involved in the unfolding of the mRNA⁴⁴. We identified a set of five cross-links, four to S10 and one to S2, including the structurally unresolved 26-aa, C-terminal (C-ter) region of S3. We identified a unique structural homolog (PDB ID: 2JPL⁴⁵; 38% similarity) using HHPRED⁴⁶ and modeled additional parts of the S3 protein. Without implementing any of the cross-linking constraints, the model satisfies all cross-links and, thereby, validates the structure with the additional C-ter region (Fig. S3).

Ribosomal chaperone. We identified three cross-links for the trigger factor, the only ribosome-associated chaperone in *E. coli*⁴⁷. The bound state of the trigger factor is resolved at low-resolution (19 Å, PDB ID: 2VRH)⁴⁸ and shows an open conformation (Fig. 4B). It is also partially resolved in its N-terminal (N-ter) region in a homologous high-resolution ribosomal structure from *Haloarcula marismortui* (PDB ID: 1W2B⁴⁹). In addition, the full-length structure in the unbound state has been solved by NMR spectroscopy (PDB ID: 1W26⁴⁹). Based on the published structures of the trigger factor homolog, a simple superposition of the unbound state on the bound N-ter region of the trigger factor resulted in a model that violated the cross-linking distances in the C-ter region, recapitulating the same binding mode observed in the low-resolution cryo-EM map previously published (PDB ID: 2VRH⁴⁸). We therefore applied a cross-link-driven structural modeling of the trigger factor. We confirmed that the N-ter region is bound near the exit tunnel of the nascent peptide chain, formed by L23 and L24⁴⁹. The distance violation (50.3 Å) between the trigger factor and L24 is justified by the inherent flexibility of the exit tunnel. Interestingly, we discovered that the C-ter region folds back to the 23S RNA, forming an extensive non-covalent interaction network with the 23S-RNA surface (Fig. 4B).

Extending the ribosomal model structure. We directly discovered that three cross-links involving the N-ter amine groups of S1, S8 and S18 could not be mapped on the ribosome because they were simply missing. Usually, the start Met is often cleaved and not considered for structure determination, but in our data it was involved in inter-molecular cross-links. By de novo completion, these novel cross-links could be mapped and were found to satisfy the distance threshold. However, the majority of missing residues are located in highly flexible regions and were subsequently identified and characterized, as described for L31 and S1 below.

Localization and C-terminal flexibility of L31. Based on the mixed-type distribution, meaning there is satisfaction in some, but violation in other structures, of a cross-link between L31 and L5, we were able to identify an annotation discrepancy for L31 in crystallographic structures of *E. coli* ribosome (PDB IDs: 4V4Q⁵⁰ and 4V4H⁴¹) (Fig. 5A). In the respective crystal structures, the electron density was mistakenly identified as L31

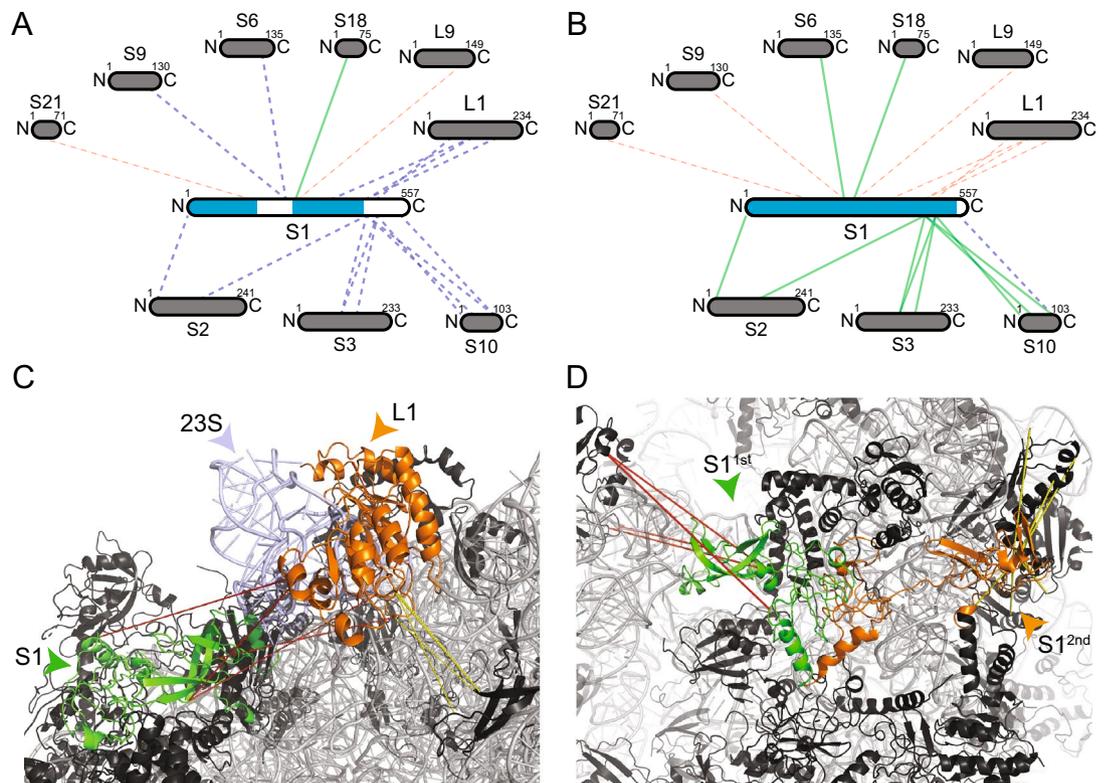


Figure 6. Identification and remodeling of the interaction network of S1. Interaction network of S1 before (A) and after (B) modeling. Sizes of the boxes are scaled to the length of the protein sequences and blue color in S1 indicates structurally characterized sequence parts. Each line represents one identified cross-link pair. Novel interactions are shown as dashed lines, satisfied (green) and violated (red) cross-link as colored lines. Ribosomal proteins (black) and rRNA (grey) are shown in cartoon. Remodeled proteins are highlighted in color (C, D). Cross-links are shown as dotted lines, and satisfaction (yellow) and violation (red) are also indicated. (C) Structural interaction network of the CTD of S1 (green) and the L1 stalk, formed by L1 (orange) and helices 76–78 of the 23S RNA (light blue). The cross-links between S1 and the L1 stalk show distance violation, but the L1 stalk itself to proteins in vicinity shows both, satisfaction (yellow) and violation (red) of cross-links, indicating different conformations as function of L1 stalk movement (see text for details). (D) Distinct conformations of the CTD of S1 (1st/2nd; green/orange). The 1st conformation (green) recapitulates the interaction of S1 with the L1 stalk in the 50S subunit. The 2nd conformation (orange) recapitulates an interaction with other 30S subunits and satisfies all cross-links. The images of protein structures were generated with PyMOL (Version 2.3.2), Schrödinger, LCC. <https://pymol.org>.

instead of L28, corroborating an inconclusive placement (Fig. 5A). In these structures, we also discovered 10 additional cross-links located in the *C-ter* region of L31 that are greatly violated (above 150 Å) and result from the incorrect placement of L31. In addition, although L31 is correctly placed in recent ribosomal structures, the *C-ter* region is still unresolved, and therefore, cross-linking distances involving the *C-ter* cannot be mapped. In addition, the *C-ter* residues 63–70 (FNIPGSK) of L31 are known to be proteolytically cleaved by protease VII during purification⁵¹. However, we reveal that, by identifying these cross-links, L31 also exists in a non-cleaved state. This was confirmed by the identification of the *C-ter* peptide of full-length L31 during peptide mass fingerprint analysis (Figs. S4, S5). By de novo modeling of the additional residues at the *C-ter* of L31, we could explain four out of ten (40%) cross-links, indicating a high flexibility of the L31 *C-ter*. By using the cross-links as distance constraints in data-driven homology modeling, we were finally able to satisfy nine out of ten (90%) crosslinks. Consequently, experimental data suggest two distinct conformations of L31's *C-ter* (Fig. 5B).

Recapitulation of interaction networks for S1. A dense network of 18 cross-links (15 novel, two violated and one satisfied) between S1 and eight proximal proteins was identified in our XL-MS studies (Fig. 6A). S1 is the largest protein of the *E. coli* ribosome with a molecular weight of 61 kDa and is essential for docking and unfolding of structured mRNA⁵², but for mRNAs with a strong Shine-Dalgarno sequence and short 5'-UTR, the S1 is not needed⁵². Only the *N-ter* domain (NTD), bound to the ribosome, was recently resolved in complex with the ribosome (PDB IDs: 6H4N, 6BU8)^{53,54}, and most of the other sequence regions of S1 are structurally uncharacterized (central domain and *C-ter* domain (CTD)). Biophysical studies suggest that the structure of S1 could be very elongated (up to 230 Å)⁵⁵, proposing a model of a bound NTD and a flexible CTD, which probes mRNA present in the cytosol. Here, we generated two conformations of the S1's CTD that is represented by two bound states: For the structurally unknown central region (150–300), we identified a close structural homolog using

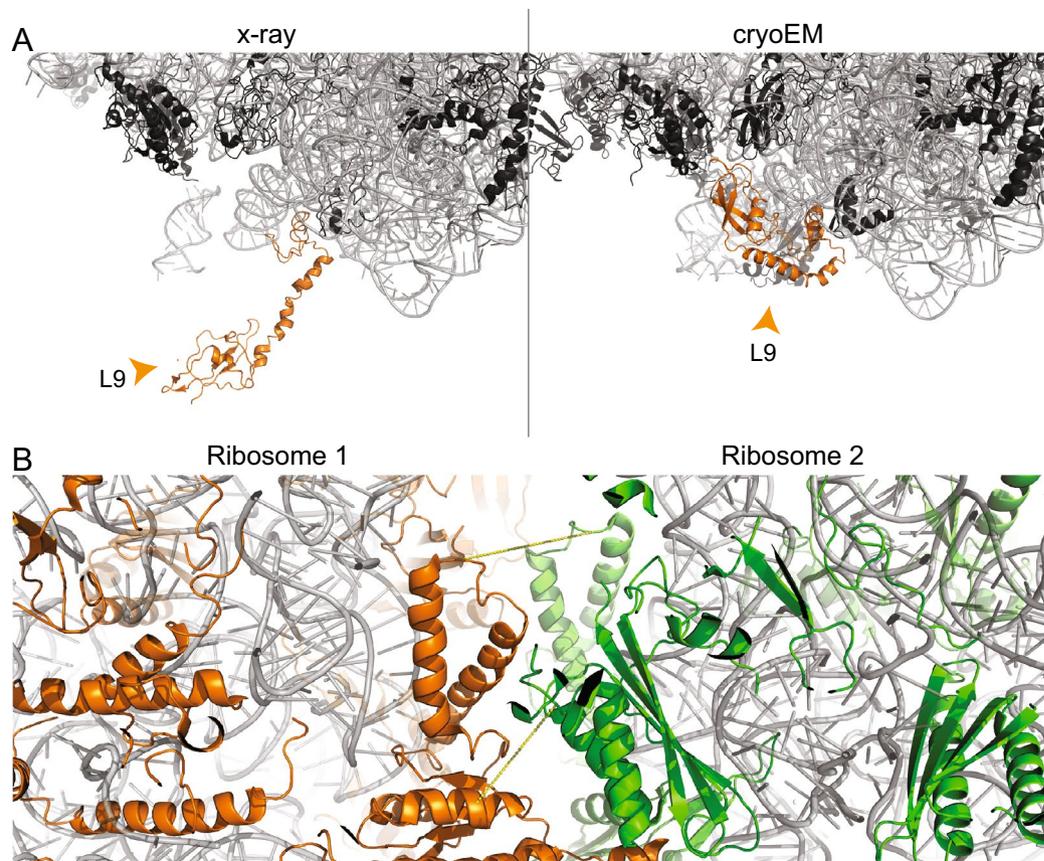


Figure 7. Multimeric states of the ribosome. Ribosomal proteins (black) and rRNA (grey) are shown in cartoon presentation. **(A)** Different conformations of L9. The protein is localized in the interface region of the asymmetric unit in the crystal structure (4V4H), and, thereby, is in an extended conformation. In the cryo-EM structure (5U9F), the protein is folding back onto the surface of the ribosome. **(B)** Cross-linking reveals presence of hibernated ribosomes in the sample (PDB ID 6H58). Ribosomal proteins of the two ribosomes are color-highlighted (green/orange). Cross-links #81 and #97 are shown. The images of protein structures were generated with PyMOL (Version 2.3.2), Schrödinger, LCC. <https://pymol.org>.

HHPRED (*E. coli* elongation factor Ts; PDB ID: 4Q7J)⁵⁶. In this region, five intermolecular cross-links were identified, of which two could be satisfied and one is only moderately violated (40.6 Å). The two additional cross-links are greatly violated (83.9 Å and 97.5 Å) and could thereby point to an extended conformation of S1⁵⁵.

Additionally, the majority of cross-links are located in the CTD of S1 that has so far been unresolved in existing ribosomal structures, but was resolved by solution NMR in its unbound state (free protein, PDB ID: 2KHJ)⁵⁷. Based on these data, we modeled the structure of S1 in its bound state, where only the 30 last amino acids at the *C-ter* are missing (Fig. 6B, structural coverage indicated in blue). We discovered two states of the CTD: A highly complex interaction with the L1 stalk (see “Flexibility and rotation of the L1 stalk”), and a second interaction within the 30S subunit, as corroborated by four and seven cross-links. Our models show that either (a) the CTD domain of S1 folds on the 23S rRNA of the L1 stalk (Fig. 6C), where extensive conformational changes are predicted to occur⁴² or (b) the CTD is bound to the 30S subunit that is validated by satisfying all seven cross-links (Fig. 6D).

Multimeric states of ribosomes. In crystallographic ribosomal structures, two cross-links between L2 and L9 are simultaneously violated and satisfied as apparently, L9 undergoes a conformational change during crystallization. The extended state of L9 is found in 33 crystallographically determined ribosomal structures from *E. coli* (Table S2) and is imposed by crystallographic contacts (Fig. 7A)¹². In ribosome hibernation, the dimeric interface is formed by the 30S subunit, while the 30S protein S2 of both ribosomes forms the core of the interface and 30S protein S1 is present in an inactive conformation⁵³. The structure of the hibernated 100S ribosome has been solved by cryo-EM (PDB ID: 6H58)⁵³, but was not considered for mapping the cross-links due to its moderate resolution of 7.9 Å. Nevertheless, that structure is highly valuable in explaining two violated cross-links, which are now in good agreement with the low-resolution cryo-EM structure of the 100S ribosome, thereby further corroborating the low-resolution cryo-EM model (calculated distances 18.3 Å and 25 Å, Fig. 7B). Finally, distance measurements in polysomes are not feasible due to the absence of high-resolution structures as they are heterogeneous and form various assemblies¹⁴. Usually, ribosomes are connected “top-to-

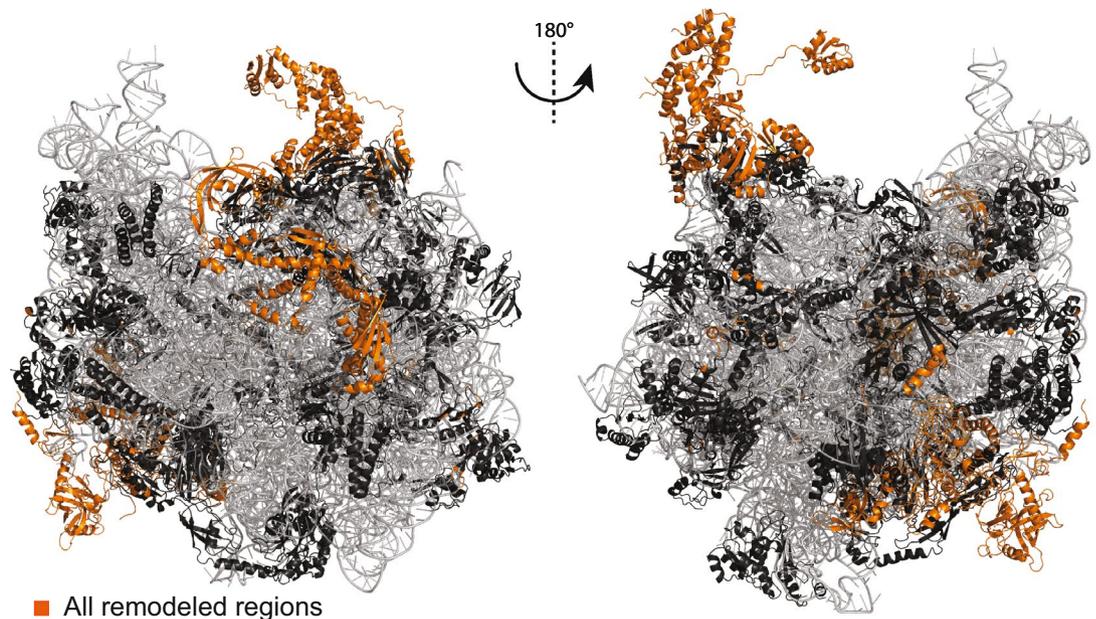


Figure 8. Final model of the *E. coli* 70S ribosome. Ribosomal proteins (black) and rRNA (grey) are shown in cartoon presentation. Added or altered ribosomal protein regions are highlighted in orange. The images of protein structures were generated with PyMOL (Version 2.3.2), Schrödinger, LCC. <https://pymol.org>.

top” in polysomes, with a 30S-30S interaction along the mRNA body and the 50S subunits facing outwards in a pseudo-helical manner¹⁴. Cross-links spanning through the ribosomes might be explained by two adjacent ribosomes bound in a top-to-top manner. This has been observed for the two cross-links in the 50S with violated distances of 120.9 Å and 166.6 Å, when mapped on the monomeric ribosomal structure. It might well be possible that these cross-links are satisfied in higher-order polysomal states.

Discussion

The published ribosomal structures of *E. coli*, albeit of impressively high numbers, account for a large fraction of ribosomal proteins, binders, cofactors, and translation states. Interestingly, our cross-linking experiments identified ~50% of cross-links that are novel, and are not part of published high-resolution ribosomal structures. This means that current structural models only partially recapitulate the intrinsic flexibility of the 70S ribosome.

Our methodology, i.e. large-scale mapping of cross-linking data on *all* ribosomal structures and subsequent cross-link-based modeling, allowed us to capture novelties in technical, biochemical, and biological aspects. In particular, we have identified crystallographic contacts by visualizing violated cross-link distances, as in the case of L9. We have identified complex oligomerization states of L7/L12 and proposed a model for its structural role in the context of the ribosome; and we discovered discrete, but limited conformations for *N*- and *C*-*ter* regions of cross-linked ribosomal proteins S1 and L31. We have, ultimately, unraveled a novel interface for the *C*-*ter* region of the trigger factor, satisfying the cross-linking data. Advanced image processing methods using cryo-EM for sorting and classifying diverse ribosomal populations from a single sample can provide insights into ribosome heterogeneity. In particular, ribosomal classes including the trigger factor in complex with the *E. coli* ribosome are predicted to recapitulate our model where the trigger factor interacts with the 23S RNA in an extended conformation.

A testable hypothesis for ribosomologists based on our extended ribosomal model may include the L7/L12 system, for which efficient depletion has been developed in isolated ribosomes⁵⁸. The four copies of L7/L12 bind at the periphery of the ribosome with the *N*-*ter*, while the *C*-*ter* captures tRNA molecules and GTPases for efficient translation⁵⁸. A hinge region (residues 31–52) connects the terminal domains. Functional studies show that hinge truncation leads to inactivation while hinge extension seems to have no or only minor effects on translation⁵⁹. In addition, based on a single construct with a randomized sequence, it was concluded that the composition of hinge residues does not have an influence on translation activity⁵⁹. For the hinge region, we have observed that a coil-to-helix transition is encoded, satisfying the underlying cross-linking data (Fig. 3). We therefore hypothesize that local secondary structure transitions could underlie the function of L7/L12. A systematic approach to probe, promote or prevent formation of predicted secondary structure of the hinge region could decode specificity determinants in the recognition mechanism involving L7/L12.

In our final *E. coli* ribosome model, we added or altered 2,057 residues (Fig. 8). In total, we were able to satisfy seven novel and four violated intramolecular, and 28 novel and 10 violated intermolecular cross-links, in addition to the already satisfied cross-links. This final model involves 115 intra-molecular cross-links and 71 inter-molecular cross-links, now fulfilling the distance threshold (30 Å for intra-, 37.5 Å for inter-molecular cross-links) with 91% and 73%, respectively (Table S5). Our workflow of remodeling a single protein in a rigid environment can be further improved by considering the environment of the protein to be flexible and suitable

for remodeling, especially, if also rRNA is involved, as *e.g.* in the L1 stalk. Additionally, seven cross-links could not be satisfied for the L7/L12 multimer. Our proposed bound model of four L7/L12 proteins is a snapshot of the variable complexes that the multimer can adopt to probe aa-tRNAs, and therefore only a subset of cross-linking data is satisfied. The flexibility of the 30S–50S interface is further highlighted by five additional violated cross-links that are mapped on L5, S9 and S10, which are all known to be involved in tRNA binding⁶⁰, frameshifting⁶¹, and ratchet movement⁴⁰. An exercise to finally fit all derived cross-links in a single static model or in snapshots is impossible, and highlights the fact that the ribosome, even in its single, purified state, probes a significantly large and complex conformational landscape. In addition, this poses several questions regarding the model if we systematically consider conformational ensembles of ribosomal proteins. This is because it is difficult to disentangle which constraint originates from which conformation. Nevertheless, our generated model is an averaged model, which incorporates multiple conformations.

To conclude, high-resolution methods cannot capture completely the conformational variation and dynamics of large macromolecular complexes and the current work shows that XL-MS is an essential method to identify technical discrepancies, conformational variability, and additional interactors, and ultimately complete and extend the available structural data for the *E. coli* ribosome.

Biomolecules are inherently flexible and highly dynamic, and this underlying flexibility manifests substantially to larger biomolecular structures. As in the case of the ribosome, it is frequently observed in density reconstructions with very high local heterogeneity. These regions appear at low resolution and are difficult to model. In addition, modeling in low resolution densities often results in a single-state model, ignoring the fact that a lower-resolution density probably reflects structural variation. XL-MS can be highly beneficial for understanding local heterogeneity in unresolved or partially resolved regions and, therefore, deliver complementary information, which may be valuable to ultimately describe conformational variation in biomolecular complexes at high resolution.

Methods

Cross-linking. 33.3 mg/mL (13.3 μ M) *E. coli* ribosome was purchased from New England Biolabs, cat. no. P0763S20. 20 μ L of the commercial solution were diluted to 200 μ L with 20 mM HEPES (pH 7.6), 10 mM MgAc₂, 30 mM KCl, and 2 mM TCEP. Four 49- μ L aliquots were pipetted into four vials. DSAU was purchased from CF Plus Chemicals, cat. no. PCL006_0050. DSAU was dissolved in neat ACN to a final concentration of 5 mM by sonicating for 5 min at room temperature. 1 μ L of DSAU stock solution was added to each 49 μ L aliquots of 1.33 μ M ribosome. The mixtures were incubated for 60 min at room temperature. 1 μ L of 1 M ammonium bicarbonate was added and incubated for 15 min at room temperature for quenching the cross-linking reactions.

Digestion. SMART Digest Trypsin Kit (Thermo Fisher Scientific) was used for digesting 50 μ L of cross-linked sample. 150 μ L of SMART Digest buffer containing 5 μ g of beads were added and the resulting solutions were incubated at 70 °C for 3 h. After cooling down the samples were centrifuged and the supernatant was collected. Peptides were further incubated with DTT (4 mM) for 30 min at 56 °C and iodoacetamide (8 mM) for 20 min at room temperature in the dark. Alkylation was quenched by adding 4 mM of DTT. Finally, TFA was added to the samples.

Mass spectrometry. The digested samples were analyzed in two technical replicates by LC/MS/MS on an UltiMate 3000 RSLC nano-HPLC system (Thermo Fisher Scientific) coupled to an Orbitrap Q-Exactive Plus mass spectrometer (Thermo Fisher Scientific) using an established protocol¹⁹. A Nanospray Flex ion source (Thermo Fisher Scientific) was employed. Peptides were trapped on a C18 column (precolumn: Acclaim PepMap 100, 300 μ m \times 5 mm, 5 μ m, 100 Å (Thermo Fisher Scientific) and separated on a packed Picofrit nanospray C18 column, 75 μ m \times 250 mm, 1.8 μ m, 80 Å, tip ID 10 μ m (New Objective). After trapping, peptides were eluted by a linear 360 min water-acetonitrile gradient from 3 to 40% B (with solvent B: 85% (v/v) acetonitrile, 0.1% (v/v) formic acid) at a flow rate of 300 nl/min. The column was washed for 10 min at 300 nl/min using 42–99% B (5 min) and 99% B (5 min). The separation was performed at 45 °C using an external column heater (Phoenix S&T).

Data were acquired in data-dependent MS/MS mode using stepped higher-energy collision-induced dissociation (HCD) at normalized collision energies of 27%, 30%, and 33% according to¹⁹. High-resolution full scans (m/z 299 to 1799, $R=140,000$ at m/z 200) were followed by 10 high-resolution product ion scans ($R=35,000$) of the most intense signals (5 s). The isolation window was set to 2 Th. Target values for automated gain control were set to 3,000,000 (MS) and 250,000 (MS/MS) with maximum injection times of 100 ms (MS) and 250 ms (MS/MS). Precursor ions with charge states $>2+$ and $<8+$ were selected for fragmentation. Dynamic exclusion of 60 s (window 2 ppm) was enabled.

Data analysis. Proteome Discoverer 2.0 was used for converting mass spectrometric raw files to mzML. Cross-links search was performed with MeroX (version 2.0 beta) using a custom-made *E. coli* database (uniprot.org) (FASTA file is included in the Supporting Information) including only ribosomal proteins and the trigger factor. The following settings were applied: Proteolytic cleavage: C-*ter* at Lys and Arg (three missed cleavages were allowed), peptide length 4–30 aa, modification: alkylation of Cys by IAA (fixed), oxidation of Met (variable), cross-linker: DSAU with specificity towards Lys, Ser, Thr, Tyr, N-*ter* (site 1) and Lys and N-*ter* (site 2), search algorithm: RISEUP mode with up to two missing ions, precursor mass accuracy 4 ppm, product ion mass accuracy 8 ppm, signal-to-noise ratio min. 2, precursor mass correction enabled, 10% intensity as prescore cut-off, 1% FDR cut-off, and minimum score cut-off: 50. For a more stringent FDR estimation the cRAP database was included. Cross-links identified in two technical replicates of each sample were combined. MS data have

been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD018935.

Identification of *E. coli* ribosomal structures. For the identification of suitable structures, the structural information deposited in the UniProt-database was used. For intermolecular cross-links, the structural information of both interaction partners were cross correlated and the PDB IDs of all structures, in which both proteins are included, were selected. For intramolecular cross-links, no cross-correlation was needed. The structures were then filtered by their annotated resolution, only using structures with a resolution of 3.5 Å or better. NMR structures that do not have a deposited resolution were always included.

Mapping of cross-links and estimating distance violation. Each identified structure was downloaded from the RCSB protein data bank in the modern mmCIF-format and residue type, residue number, chain identifier, and coordinates of all Ca atoms were manually isolated. This simplified structure file was used for all further analysis. For each cross-link/structure pair, a local sequence alignment for both residues were performed, to guarantee the correct distance measurement and to compensate for annotation discrepancies between the structure and the UniProt annotated sequence (e.g. start Met excluded in structure and thereby residue numbering shifted by 1). To perform the alignment, a tripeptide with the cross-linked residue in the middle was fetched from UniProt and aligned to the sequence of the structure, assuming correct residue numbering. For terminal residues, only a dipeptide was used, but the terminal position was considered. If the first alignment failed, the tripeptide was shifted in a ± 5 residue window around the assumed position in the structural sequence. If the alignment failed (e.g. residue not resolved; incorrectly numbered structure), no unambiguous measurement of the distance was possible and the respective cross-link/structure pair were excluded. After successful sequence alignment, the distance d between the Ca atoms were calculated with $d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$, in which $P(x_1, y_1, z_1)$ and $P(x_2, y_2, z_2)$ are the coordinates of both Ca atoms. To differentiate distances in monomeric ribosomal structures from distances in asymmetric units, which would lead to false positive results, the chain identifier label was compared. According to common convention, chains in asymmetric units are distinguishable, and distances from asymmetric units were identified by following rules: (a) first chain uppercase but second chain lowercase or vis-à-vis (e.g. A to b instead of A to B); (b) if the chain identifier does have two characters, but the first character is differently (e.g. AA to BB, instead of AA to AB). All distances were classified by the given threshold of 37.5 Å and categorized in “always satisfied” when all distances of a given cross-link pair are below the threshold, “always violated” when all distances are above the threshold and “mixed-type”, when distances are both, below and above the threshold. Cross-links not mappable to any structure were classified as “novel”.

Modeling methods. For remodeling, homology modeling, and de novo addition of residues MODELLER (Version 9.22)⁶² was used as described in the manual (chapter 2.2.2), including VTFM and MD optimization of the generated models. The published ribosomal structure “3.2 Å cryo-EM ArfA-RF2 ribosome rescue complex (Structure II)” (PDB ID: 5U9F)³² was used as the starting point for modeling, but the bound ribosome-rescue factor A was removed. Superimposing of the trigger factor was done with PyMOL. To reduce calculation time, only proteins and rRNA within a distance of 10 Å and cross-linked interaction partners of the target protein were considered, and the remodeling was only applied to this single protein, while all others were kept static. Homologue structures for unknown regions in S1 and S3 were identified using HHPRED. The identified structures (PDB ID: 4Q7L for S1⁶³; PDB ID: 2JPL for S3⁴⁵) were used to model the unknown parts in their ribosomal environment. Missing residues were added in the alignment file and added by de novo modeling. Cross-links were included as Gaussian distributed distance restraints (see MODELLER manual chapter 2.2.11). The mean distance of the constraint was stepwise reduced from 45 to 25 Å (step size 5 Å) or until the cross-link distance was below the threshold. The standard deviation of the distance restraints was 0.1. The stepwise remodeling was applied to maintain structural integrity of flexible domains and regions. Remodeled proteins were used to replace their incomplete or incorrect counterparts in the working model by using the alignment function from PyMOL, sequentially building an extended ribosomal structure.

Data availability

Data are available via ProteomeXchange with identifier PXD018935. All models and modeling protocols are available from the corresponding authors upon request.

Received: 2 May 2020; Accepted: 10 July 2020

Published online: 28 July 2020

References

- Klinge, S. & Woolford, J. L. Jr. Ribosome assembly coming into focus. *Nat. Rev. Mol. Cell Biol.* **20**, 116–131. <https://doi.org/10.1038/s41580-018-0078-y> (2019).
- Wagenknecht, T., Carazo, J. M., Radermacher, M. & Frank, J. Three-dimensional reconstruction of the ribosome from *Escherichia coli*. *Biophys J* **55**, 455–464. [https://doi.org/10.1016/S0006-3495\(89\)82839-5](https://doi.org/10.1016/S0006-3495(89)82839-5) (1989).
- Yonath, A. *et al.* Characterization of single crystals of the large ribosomal particles from *Bacillus stearothermophilus*. *J Mol Biol* **187**, 633–636. [https://doi.org/10.1016/0022-2836\(86\)90342-6](https://doi.org/10.1016/0022-2836(86)90342-6) (1986).
- Noller, H. F. & Chaires, J. B. Functional modification of 16S ribosomal RNA by kethoxal. *Proc. Natl. Acad. Sci. USA* **69**, 3115–3118. <https://doi.org/10.1073/pnas.69.11.3115> (1972).
- Falvey, A. K. & Staehelin, T. Structure and function of mammalian ribosomes. I. Isolation and characterization of active liver ribosomal subunits. *J. Mol. Biol.* **53**, 1–19. [https://doi.org/10.1016/0022-2836\(70\)90042-2](https://doi.org/10.1016/0022-2836(70)90042-2) (1970).

6. Wimberly, B. T. *et al.* Structure of the 30S ribosomal subunit. *Nature* **407**, 327–339. <https://doi.org/10.1038/35030006> (2000).
7. Evers, U., Franceschi, F., Boddeker, N. & Yonath, A. Crystallography of halophilic ribosome: the isolation of an internal ribonucleoprotein complex. *Biophys. Chem.* **50**, 3–16. [https://doi.org/10.1016/0301-4622\(94\)85015-1](https://doi.org/10.1016/0301-4622(94)85015-1) (1994).
8. Clemons, W. M. Jr. *et al.* Structure of a bacterial 30S ribosomal subunit at 5.5 Å resolution. *Nature* **400**, 833–840. <https://doi.org/10.1038/23631> (1999).
9. Ban, N. *et al.* A 9 Å resolution X-ray crystallographic map of the large ribosomal subunit. *Cell* **93**, 1105–1115. [https://doi.org/10.1016/s0092-8674\(00\)81455-5](https://doi.org/10.1016/s0092-8674(00)81455-5) (1998).
10. Huxley, H. E. & Zubay, G. Electron microscope observations on the structure of microsomal particles from *Escherichia coli*. *J. Mol. Biol.* **2**, 10–18. [https://doi.org/10.1016/S0022-2836\(60\)80003-4](https://doi.org/10.1016/S0022-2836(60)80003-4) (1960).
11. Bai, X. C., Fernandez, I. S., McMullan, G. & Scheres, S. H. Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *Elife* **2**, e00461. <https://doi.org/10.7554/eLife.00461> (2013).
12. Brown, A. & Shao, S. Ribosomes and cryo-EM: A duet. *Curr. Opin. Struct. Biol.* **52**, 1–7. <https://doi.org/10.1016/j.sbi.2018.07.001> (2018).
13. Matzov, D. *et al.* The cryo-EM structure of hibernating 100S ribosome dimer from pathogenic *Staphylococcus aureus*. *Nat. Commun.* **8**, 723. <https://doi.org/10.1038/s41467-017-00753-8> (2017).
14. Brandt, F. *et al.* The native 3D organization of bacterial polysomes. *Cell* **136**, 261–271. <https://doi.org/10.1016/j.cell.2008.11.016> (2009).
15. Kohler, R., Mooney, R. A., Mills, D. J., Landick, R. & Cramer, P. Architecture of a transcribing-translating expressome. *Science* **356**, 194–197. <https://doi.org/10.1126/science.aal3059> (2017).
16. Demo, G. *et al.* Structure of RNA polymerase bound to ribosomal 30S subunit. *Elife*. <https://doi.org/10.7554/eLife.28560> (2017).
17. Garreau de Loubresse, N. *et al.* Structural basis for the inhibition of the eukaryotic ribosome. *Nature* **513**, 517–522. <https://doi.org/10.1038/nature13737> (2014).
18. Arenz, S. *et al.* Structures of the orthosomycin antibiotics avilamycin and evernimicin in complex with the bacterial 70S ribosome. *Proc. Natl. Acad. Sci. USA* **113**, 7527–7532. <https://doi.org/10.1073/pnas.1604790113> (2016).
19. Genuth, N. R. & Barna, M. The discovery of ribosome heterogeneity and its implications for gene regulation and organismal life. *Mol. Cell* **71**, 364–374. <https://doi.org/10.1016/j.molcel.2018.07.018> (2018).
20. Liu, F., Rijkers, D. T., Post, H. & Heck, A. J. Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. *Nat. Methods* **12**, 1179–1184. <https://doi.org/10.1038/nmeth.3603> (2015).
21. Kastritis, P. L. *et al.* Capturing protein communities by structural proteomics in a thermophilic eukaryote. *Mol. Syst. Biol.* **13**, 936. <https://doi.org/10.15252/msb.20167412> (2017).
22. Gotze, M., Iacobucci, C., Ihling, C. H. & Sinz, A. A simple cross-linking/mass spectrometry workflow for studying system-wide protein interactions. *Anal. Chem.* **91**, 10236–10244. <https://doi.org/10.1021/acs.analchem.9b02372> (2019).
23. Sinz, A. Divide and conquer: Cleavable cross-linkers to study protein conformation and protein-protein interactions. *Anal. Bioanal. Chem.* **409**, 33–44. <https://doi.org/10.1007/s00216-016-9941-x> (2017).
24. Iacobucci, C., Gotze, M. & Sinz, A. Cross-linking/mass spectrometry to get a closer view on protein interaction networks. *Curr. Opin. Biotechnol.* **63**, 48–53. <https://doi.org/10.1016/j.copbio.2019.12.009> (2019).
25. Sinz, A. Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes. *J. Mass. Spectrom.* **38**, 1225–1237. <https://doi.org/10.1002/jms.559> (2003).
26. Kaake, R. M. *et al.* A new *in vivo* cross-linking mass spectrometry platform to define protein-protein interactions in living cells. *Mol. Cell. Proteom.* **13**, 3533–3543. <https://doi.org/10.1074/mcp.M114.042630> (2014).
27. Fischer, L. & Rappsilber, J. Quirks of error estimation in cross-linking/mass spectrometry. *Anal. Chem.* **89**, 3829–3833. <https://doi.org/10.1021/acs.analchem.6b03745> (2017).
28. Bullock, J. M. A., Schwab, J., Thalassinou, K. & Topf, M. The importance of non-accessible crosslinks and solvent accessible surface distance in modeling proteins with restraints from crosslinking mass spectrometry. *Mol. Cell. Proteom.* **15**, 2491–2500. <https://doi.org/10.1074/mcp.M116.058560> (2016).
29. Yugandhar, K., Wang, T.-Y. & Yu, H. Structure-based validation can drastically under-estimate error rate in proteome-wide cross-linking mass spectrometry studies. *bioRxiv* <https://doi.org/10.1101/617654> (2019).
30. Fursch, J., Kammer, K. M., Krefl, S. G., Beck, M. & Stengel, F. Proteome-wide structural probing of low-abundant protein interactions by cross-linking mass spectrometry. *Anal. Chem.* **92**, 4016–4022. <https://doi.org/10.1021/acs.analchem.9b05559> (2020).
31. Muller, M. Q., Dreiocker, F., Ihling, C. H., Schafer, M. & Sinz, A. Cleavable cross-linker for protein structure analysis: Reliable identification of cross-linking products by tandem MS. *Anal. Chem.* **82**, 6958–6968. <https://doi.org/10.1021/ac101241t> (2010).
32. Demo, G. *et al.* Mechanism of ribosome rescue by ArfA and RF2. *Elife*. <https://doi.org/10.7554/eLife.23687> (2017).
33. Diaconu, M. *et al.* Structural basis for the function of the ribosomal L7/L12 stalk in factor binding and GTPase activation. *Cell* **121**, 991–1004. <https://doi.org/10.1016/j.cell.2005.04.015> (2005).
34. Zhou, J., Lancaster, L., Trakhanov, S. & Noller, H. F. Crystal structure of release factor RF3 trapped in the GTP state on a rotated conformation of the ribosome. *RNA* **18**, 230–240. <https://doi.org/10.1261/rna.031187.111> (2012).
35. Pulk, A. & Cate, J. H. Control of ribosomal subunit rotation by elongation factor G. *Science* **340**, 1235970. <https://doi.org/10.1126/science.1235970> (2013).
36. Tian, P. *et al.* Folding pathway of an Ig domain is conserved on and off the ribosome. *Proc. Natl. Acad. Sci. USA* **115**, E11284–E11293. <https://doi.org/10.1073/pnas.1810523115> (2018).
37. Mandava, C. S. *et al.* Bacterial ribosome requires multiple L12 dimers for efficient initiation and elongation of protein synthesis involving IF2 and EF-G. *Nucleic Acids Res.* **40**, 2054–2064. <https://doi.org/10.1093/nar/gkr1031> (2012).
38. Leijonmarck, M. & Liljas, A. Structure of the C-terminal domain of the ribosomal protein L7/L12 from *Escherichia coli* at 1.7 Å. *J. Mol. Biol.* **195**, 555–579. [https://doi.org/10.1016/0022-2836\(87\)90183-5](https://doi.org/10.1016/0022-2836(87)90183-5) (1987).
39. Kahraman, A., Malmstrom, L. & Aebersold, R. Xwalk: Computing and visualizing distances in cross-linking experiments. *Bioinformatics* **27**, 2163–2164. <https://doi.org/10.1093/bioinformatics/btr348> (2011).
40. Gao, H. *et al.* Study of the structural dynamics of the *E. coli* 70S ribosome using real-space refinement. *Cell* **113**, 789–801. [https://doi.org/10.1016/s0092-8674\(03\)00427-6](https://doi.org/10.1016/s0092-8674(03)00427-6) (2003).
41. Schuwirth, B. S. *et al.* Structural analysis of kasugamycin inhibition of translation. *Nat. Struct. Mol. Biol.* **13**, 879–886. <https://doi.org/10.1038/nsmb1150> (2006).
42. Valle, M. *et al.* Locking and unlocking of ribosomal motions. *Cell* **114**, 123–134. [https://doi.org/10.1016/s0092-8674\(03\)00476-8](https://doi.org/10.1016/s0092-8674(03)00476-8) (2003).
43. Trabuco, L. G. *et al.* The role of L1 stalk-tRNA interaction in the ribosome elongation cycle. *J. Mol. Biol.* **402**, 741–760. <https://doi.org/10.1016/j.jmb.2010.07.056> (2010).
44. Amiri, H. & Noller, H. F. Structural evidence for product stabilization by the ribosomal mRNA helicase. *RNA* **25**, 364–375. <https://doi.org/10.1261/rna.068965.118> (2019).
45. Rogne, P., Fimland, G., Nissen-Meyer, J. & Kristiansen, P. E. Three-dimensional structure of the two peptides that constitute the two-peptide bacteriocin lactococcin G. *Biochim. Biophys. Acta* **1784**, 543–554. <https://doi.org/10.1016/j.bbapap.2007.12.002> (2008).
46. Zimmermann, L. *et al.* A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.* **430**, 2237–2243. <https://doi.org/10.1016/j.jmb.2017.12.007> (2018).

47. Hesterkamp, T., Hauser, S., Lutcke, H. & Bukau, B. *Escherichia coli* trigger factor is a prolyl isomerase that associates with nascent polypeptide chains. *Proc. Natl. Acad. Sci. USA* **93**, 4437–4441. <https://doi.org/10.1073/pnas.93.9.4437> (1996).
48. Merz, F. *et al.* Molecular mechanism and structure of Trigger Factor bound to the translating ribosome. *EMBO J* **27**, 1622–1632. <https://doi.org/10.1038/emboj.2008.89> (2008).
49. Ferbitz, L. *et al.* Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature* **431**, 590–596. <https://doi.org/10.1038/nature02899> (2004).
50. Schuwirth, B. S. *et al.* Structures of the bacterial ribosome at 3.5 Å resolution. *Science* **310**, 827–834. <https://doi.org/10.1126/science.1117230> (2005).
51. Eistetter, A. J., Butler, P. D., Traut, R. R. & Fanning, T. G. Characterization of *Escherichia coli* 50S ribosomal protein L31. *FEMS Microbiol. Lett.* **180**, 345–349. <https://doi.org/10.1111/j.1574-6968.1999.tb08816.x> (1999).
52. Duval, M. *et al.* *Escherichia coli* ribosomal protein S1 unfolds structured mRNAs onto the ribosome for active translation initiation. *PLoS Biol.* **11**, e1001731. <https://doi.org/10.1371/journal.pbio.1001731> (2013).
53. Beckert, B. *et al.* Structure of a hibernating 100S ribosome reveals an inactive conformation of the ribosomal protein S1. *Nat. Microbiol.* **3**, 1115–1121. <https://doi.org/10.1038/s41564-018-0237-0> (2018).
54. Loveland, A. B. & Korostelev, A. A. Structural dynamics of protein S1 on the 70S ribosome visualized by ensemble cryo-EM. *Methods* **137**, 55–66. <https://doi.org/10.1016/j.ymeth.2017.12.004> (2018).
55. Subramanian, A. R. Structure and functions of ribosomal protein S1. *Prog. Nucleic Acid Res. Mol. Biol.* **28**, 101–142. [https://doi.org/10.1016/s0079-6603\(08\)60085-9](https://doi.org/10.1016/s0079-6603(08)60085-9) (1983).
56. Takeshita, D., Yamashita, S. & Tomita, K. Molecular insights into replication initiation by Qbeta replicase using ribosomal protein S1. *Nucleic Acids Res.* **42**, 10809–10822. <https://doi.org/10.1093/nar/gku745> (2014).
57. Salah, P. *et al.* Probing the relationship between Gram-negative and Gram-positive S1 proteins by sequence analysis. *Nucleic Acids Res* **37**, 5578–5588. <https://doi.org/10.1093/nar/gkp547> (2009).
58. Carlson, M. A. *et al.* Ribosomal protein L7/L12 is required for GTPase translation factors EF-G, RF3, and IF2 to bind in their GTP state to 70S ribosomes. *FEBS J* **284**, 1631–1643. <https://doi.org/10.1111/febs.14067> (2017).
59. Babunenko, M. G., Chuiikov, S. V. & Gudkov, A. T. The length of the interdomain region of the L7/L12 protein is important for its function. *FEBS Lett.* **313**, 232–234. [https://doi.org/10.1016/0014-5793\(92\)81198-u](https://doi.org/10.1016/0014-5793(92)81198-u) (1992).
60. Hoang, L., Fredrick, K. & Noller, H. F. Creating ribosomes with an all-RNA 30S subunit P site. *Proc. Natl. Acad. Sci. USA* **101**, 12439–12443. <https://doi.org/10.1073/pnas.0405227101> (2004).
61. Hong, S. *et al.* Mechanism of tRNA-mediated +1 ribosomal frameshifting. *Proc. Natl. Acad. Sci. USA* **115**, 11226–11231. <https://doi.org/10.1073/pnas.1809319115> (2018).
62. Sali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815. <https://doi.org/10.1006/jmbi.1993.1626> (1993).
63. Bukowska, M. A. *et al.* A transporter motor taken apart: Flexibility in the nucleotide binding domains of a heterodimeric ABC exporter. *Biochemistry* **54**, 3086–3099. <https://doi.org/10.1021/acs.biochem.5b00188> (2015).

Acknowledgements

The authors would like to thank Mr. Daniele Ubbiali and Ms. Xiaohan Wang for their contribution in XL-MS data analysis. A.S. and P.L.K. acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), RTG 2467, project number 391498659 “Intrinsically Disordered Proteins—Molecular Principles, Cellular Functions, and Diseases”. A.S. acknowledges financial support by the DFG (project Si 867/15-2) and the region of Saxony-Anhalt. P.L.K. and C.T. are supported by the Federal Ministry for Education and Research (BMBF, ZIK program, grant number 03Z22HN23 to P.L.K.), the European Regional Development Funds for Saxony-Anhalt (Grant number EFRE: ZS/2016/04/78115 to P.L.K.) and the Martin Luther University Halle-Wittenberg. C.I. acknowledges funding by the Alexander von Humboldt Foundation and H2020 MSCA IF 2018 EU project 843585.

Author contributions

C.T. performed analysis and modeling, C.I. performed cross-linking experiments and data analysis, C.H.I. performed mass spectrometry, A.S. and P.L.K. supervised the study. All authors designed the study, conceived analysis, wrote, and reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-69313-3>.

Correspondence and requests for materials should be addressed to P.L.K. or A.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020