Title: Structure of the Bacterial Ribosome at 2 Å Resolution

Authors: Zoe L. Watson[1], Fred R. Ward[2], Raphaël Méheust[3,5], Omer Ad[4], Alanna Schepartz[1,2], Jillian F. Banfield[3,5,6], Jamie H.D. Cate[1,2,7]

Author Affiliations:

1. Department of Chemistry, University of California Berkeley, Berkeley, CA 94720, United States
2. Department of Molecular and Cell Biology, University of California Berkeley, Berkeley, CA 94720, United States
3. Innovative Genomics Institute, University of California Berkeley, Berkeley, CA, USA.
4. Department of Chemistry, Yale, New Haven, CT 06520, United States
5. Earth and Planetary Science, University of California Berkeley, Berkeley, CA, USA.
6. Environmental Science, Policy and Management, University of California Berkeley, Berkeley, CA, USA.
7. Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, United States

Corresponding Author: Jamie H.D. Cate, j-h-doudna-cate@berkeley.edu

Key Words: Cryo-EM, RNA, ribosome, antibiotics, aminoglycosides, post-translational modifications, post-transcriptional modifications

Abstract

Continuing advances in cryo-electron microscopy (cryo-EM) demonstrate the promise it holds for revealing biological structures at chemical resolution, in which noncovalent interactions, RNA and protein modifications, and solvation can be modeled accurately. At present, the best cryo-EM-derived models of the bacterial ribosome are of the large (50S) ribosomal subunit with effective global resolutions of 2.4-2.5 Å, based on map-to-model Fourier shell correlation (FSC). Here we present a model of the *E. coli* 70S ribosome with an effective global resolution of 2.0 Å, based on maps showcasing unambiguous positioning of residues, their detailed chemical interactions, and chemical modifications. These modifications include the first examples of isopeptide and thioamide backbone substitutions in ribosomal proteins, the former of which is likely conserved in all domains of life. The model also defines extensive solvation of the small (30S) ribosomal subunit for the first time, as well as interactions with A-site and P-site tRNAs, mRNA, and the antibiotic paromomycin. The high quality of the maps now allows a deeper phylogenetic analysis of ribosomal components, and identification of structural conservation to the level of solvation. The maps and models of the bacterial ribosome presented here should enable future structural analysis of the chemical basis for translation, and the development of robust tools for cryo-EM structure modeling and refinement.

Introduction

The ribosome performs the crucial task of translating the genetic code into proteins and varies in size from 2.3 MDa to over 4 MDa across the three domains of life (Melnikov et al., 2012). Polypeptide synthesis occurs in the peptidyl transferase center (PTC), where the ribosome acts primarily as an "entropic trap" for peptide bond formation (Rodnina, 2013). To carry out the highly coordinated process of translation, the ribosome orchestrates the binding and readout of messenger RNA (mRNA) and transfer RNAs (tRNAs), coupled with a multitude of interactions between the small and large ribosomal subunits and a host of translation factors. These molecular interactions are accompanied by a wide range of conformational dynamics that contribute to translation accuracy and speed. Because of the ribosome's essential role in supporting life, it is naturally the target of a plurality of antibiotics with diverse mechanisms of action (Arenz and Wilson, 2016). The ribosome also plays a unique role in our ability to study the vast array of RNA secondary and tertiary structural motifs found in nature, as well as RNA-protein interactions. Although X-ray crystallography has been central in revealing the molecular basis of many steps in translation, the resolution of available X-ray crystal structures of the ribosome in key functional states remains too low to provide accurate models of non-covalent bonding, i.e. hydrogen bonding, van der Waals contacts, and ionic interactions. Furthermore, development of small molecule drugs such as antibiotics is hampered at the typical resolution of available X-ray crystal structures of the ribosome (~3 Å) (Arenz and Wilson, 2016; Yusupova and Yusupov, 2017). Thus, understanding of the molecular interactions in the ribosome in chemical detail would provide a foundation for biochemical and biophysical approaches that probe ribosome function and aid antibiotic discovery.

The ribosome has been an ideal target for cryo-EM since the early days of single-particle reconstruction methods, as its large size, many functional states, and multiple binding partners

lead to conformational heterogeneity that make it challenging for X-ray crystallography (Frank, 2017). Previous high-resolution structures of the bacterial ribosome include X-ray crystal structures of the *E. coli* ribosome at 2.4 Å (2.1 Å by CC ½) (Noeske et al., 2015) and *T. thermophilus* ribosome at 2.3 Å, and cryo-EM structures reported at 2.1-2.3 Å (Halfon et al., 2019; Pichkur et al., 2020; Polikanov et al., 2015; Stojković et al., 2020). While resolution is well-defined for crystallography, which measures information in Fourier space, it is more difficult to assign a metric to global resolution of cryo-EM structures. Fourier Shell Correlation (FSC) thresholds are widely used for this purpose (Frank and Al-Ali, 1975; Harauz and van Heel, 1986). However, the "gold-standard" FSC most commonly reported, wherein random halves of the particles are refined independently and correlation of the two half-maps is calculated as a function of spatial frequency, is more precisely a measure of self-consistency of the data (Henderson et al., 2012; Rosenthal and Henderson, 2003; Subramaniam et al., 2016). Reporting of resolution is further complicated by variations in map refinement protocols and post-processing by the user, as well as a lack of strict standards for deposition of these data (Subramaniam et al., 2016). A separate measure, which may be more useful for higher resolution, is to compare the cryo-EM experimental map to the structural model derived from it, or so-called map-to-model FSC (DiMaio et al., 2013; Subramaniam et al., 2016). In this work we use both metrics to report on map quality but emphasize the use of the map-to-model FSC criterion as it more directly reports on the utility of the model for drawing conclusions about molecular interactions.

Here we have determined the structure of the *E. coli* 70S ribosome to a global resolution of 2.02 Å according to the "gold-standard" Fourier Shell Correlation (FSC) 0.143 cutoff, with focused refinement leading to local resolutions as high as 1.8 Å. The map-to-model FSC for the 50S subunit is 1.95 Å, 2.17 Å for the 30S subunit, and 2.06 Å for the 70S complex as a whole. We highlight well-resolved features of the map with particular relevance to ribosomal function,

including contacts to mRNA and tRNA substrates, a detailed description of the aminoglycoside antibiotic paromomycin bound in the mRNA decoding center, and interactions between the ribosomal subunits. We also describe solvation and ion positions, as well as features of post-transcriptional modifications and post-translational modifications seen here for the first time. Discovery of these chemical modifications, as well as a new RNA-interacting motif found in protein bS21, provide the basis for addressing phylogenetic conservation of ribosomal protein structure and clues towards the role of a protein of unknown function. These results open new avenues for studies of the chemistry of translation and should aid future development of tools for refining structural models into cryo-EM maps.

Results

Overall map quality

We determined the structure of the *Escherichia coli* 70S ribosome in the classical (non-rotated) state with mRNA and tRNAs bound in the Aminoacyl-tRNA and Peptidyl-tRNA sites (A site and P site, respectively). Partial density for Exit-site (E-site) tRNA is also visible in the maps, particularly for the 3'-terminal C75 and A76 nucleotides. Our final maps were generated from two 70S ribosome complexes that were formed separately using P-site $tRNA^{fMet}$ that differed only by being charged with two different non-amino acid monomers (see Methods). In both complexes, we used the same mRNA and A-site Val-$tRNA^{Val}$. Both complexes yielded structures in the same functional state, with similar occupancy of the tRNAs. As neither A-site nor P-site tRNA 3'-CCA-ends were resolved in the individual cryo-EM maps, we merged the two datasets for the final reconstructions (**Figure 1–figure supplement 1**). Global resolution of the entire complex reached 2.02 Å resolution, using conventional half-map FSC measures, with local resolution reaching 1.8 Å (**Figure 1A**, **Figure 1–figure supplement 2, Table 1**). The global resolution reached 2.06 Å using the map-to-model FSC criterion (**Figure 1–figure supplement 2**). We also used focused refinement of the large (50S) and small (30S) ribosomal subunits, and further focused refinements of smaller regions that are known to be conformationally flexible, to enhance their resolution (**Figure 1–figure supplement 3**). In particular, focused refinement of the 30S subunit improved its map quality substantially, along with its immediate contacts to the 50S subunit and the mRNA and tRNA anticodon stem-loops. Additional focused refinement of the Central Protuberance (CP) in the 50S subunit, the 30S head domain, and 30S platform aided in model building and refinement. In the following descriptions, maps of specific ribosomal subunits or domains refer to the focused-refined maps. Details of the conventional and map-to-model FSC resolutions are given in **Figure 1–figure supplements 2-3** and **Table 2**.

The high resolution indicated by the FSC curves is supported by several visual features observable in the maps, including holes in many aromatic rings and riboses, as well as ring puckers, directionality of non-bridging phosphate oxygens in ribosomal RNA, and numerous well-resolved ions, waters, and small molecules (**Figure 1B-D, Figure 1–figure supplement 4**). The maps also reveal post-transcriptional and post-translational modifications of ribosomal RNA and proteins in detail (**Figure 1–figure supplement 5, Figure 1–figure supplement 6**). We additionally observed evidence for possible ribose damage in some nucleotides due to the electron dose (**Figure 1–figure supplement 7**). As seen in other ribosome structures, elements of the ribosome at the periphery are less ordered, including the uL1 arm, the GTPase activating center, bL12 proteins and central portion of the A-site finger in the 50S subunit (23S rRNA helix H38), as well as the periphery of the 30S subunit head domain and spur (16S rRNA helix h6). The resolution of the maps for the elbow and acceptor ends of the P-site and A-site tRNAs is also relatively low, likely as a result of poor accommodation of the unnatural substrates used.

High-resolution structural features of the 30S ribosomal subunit

The 30S ribosomal subunit is highly dynamic in carrying out its role in the translation cycle (Munro et al., 2009). Previously published structures demonstrate that even in defined conformational states of the ribosome, the 30S subunit exhibits more flexibility than the core of the 50S subunit. Furthermore, the mass of the 50S subunit dominates alignments in cryo-EM reconstructions of the 70S ribosome. Solvation of the 30S subunit has not been extensively modeled, again owing to the fact that it is generally more flexible and less well-resolved than the 50S subunit, even in available high-resolution structures (Noeske et al., 2015; Polikanov et al., 2015). Using focused-refined maps of the 30S subunit, we achieved the resolution necessary for in-depth chemical analysis of key contacts to mRNA and tRNAs, the 50S subunit, and to generate more complete

models of 30S subunit components, Mg^{2+} ion positions and solvation (**Figure 2 - figure supplement 1**).

mRNA and tRNA interactions with the 30S ribosomal subunit

The 30S ribosomal subunit controls interactions of tRNAs with mRNA and helps maintain the mRNA in the proper reading frame. The present structure reveals the interactions of the 30S subunit with mRNA and tRNA in the A and P sites including solvation. The tight binding of tRNA to the P site (Lill et al., 1986) is reflected in extensive direct contacts between the tRNA anticodon stem loop (ASL) and both the 30S subunit head and platform domains (**Figure 2A-D**). The C-terminal residues Lys129 and Arg130 of protein uS9, which are important for translational fidelity (Arora et al., 2013), form ionic and hydrogen bonding interactions with the nucleotides in the U-turn motif of the P-site ASL, and Arg130 stacks with the base of U33 (**Figure 2D**). However, the C-terminal tails of ribosomal proteins uS13 and uS19, which come from the 30S subunit head domain and are lysine-rich, are not visible in the map, suggesting they do not make specific contacts with the tRNA when the ribosome is in the unrotated state. Direct interactions between A-site tRNA and the 30S subunit are highly localized to the top of helices h44 and h18 in 16S rRNA and to helix H69 in 23S rRNA (i.e. 16S rRNA nucleotides G530, A1492, and A1493, and 23S rRNA nucleotide A1913). By contrast, contacts between the A-site tRNA and 30S subunit head domain are entirely solvent mediated (**Figure 2E**) apart from nucleotide C1054, possibly reflecting the weaker binding of tRNA to the A site (Lill et al., 1986) and the need for A-site tRNA conformational dynamics during mRNA decoding (Rodnina et al., 2017).

Contacts of protein bS21 with the 30S subunit head domain

Protein bS21, which resides near the path of mRNA on the 30S subunit platform (Held et al., 1973; Sashital et al., 2014), is essential in *E. coli* (Bubunenko et al., 2007; Goodall et al., 2018). Although partial structural models of bS21 have been determined (Fischer et al., 2015; Noeske et al., 2015), the present structure provides clear evidence for the conformation of the entire open reading frame, including 13 amino acids at the C-terminus that extend to the base of the 30S subunit head domain (**Figure 3A**). Most of the C-terminal residues are found in an alpha-helical conformation near the Shine-Dalgarno helix formed between the 3' end of 16S rRNA and the mRNA ribosome binding site (Shine and Dalgarno, 1975), while the C-terminal arginine-leucine-tyrosine (RLY) motif makes close contacts with 16S rRNA helix h37 and nucleotide A1167 (**Figure 3B**). The arginine and leucine residues pack in the minor groove of helix h37, and the terminal tyrosine stacks on A1167. Multiple alignment of bS21 sequences from distinct bacterial phyla revealed that the RLY C-terminal motif is conserved in bS21 sequences of the Gammaproteobacteria phylum while in Betaproteobacteria, the sibling group of Gammaproteobacteria, bS21 sequences possess a lysine-leucine-tyrosine (KLY) C-terminal motif instead (**Figure 3C**). Interestingly, such C-terminal extensions (RLY or KLY) are absent in other bacterial phyla (**Figure 3–figure supplement 1**). Recently, putative homologues of bS21 were identified in huge bacteriophages, which were shown to harbor genes encoding components of the translational machinery (Al-Shayeb et al., 2020). Inspection of sequence alignments reveals that some phage S21 homologues also contain KLY-like motifs (**Figure 3C**). The presence of phage S21 homologues containing KLY-like motifs is consistent with the host range of these phages (**Figure 3C, Supplementary Table 1**). Other phages with predicted hosts lacking the C-terminal motif in bS21 encode S21 homologues that also lack the motif (**Supplementary Table 1**). Although the C-terminal region of bS21 resides near the Shine-Dalgarno helix, examination of ribosome binding site consensus sequences in these bacterial clades and the predicted ribosome

binding sites in the associated phages do not reveal obvious similarities (**Supplementary Table 1**).

Post-translational and post-transcriptional modifications in the 30S subunit

Ribosomal protein uS11 is a central component of the 30S subunit platform domain, and assembles cooperatively with ribosomal proteins uS6 and uS18 (Stern et al., 1988), preceding the binding of bS21 late in 30S subunit maturation (Held et al., 1973; Sashital et al., 2014). Protein uS11 makes intimate contact with 16S rRNA residues in a 3-helix junction that forms part of the 30S subunit E site and stabilizes the 16S rRNA that forms the platform component of the P site (Stern et al., 1988). Remarkably, inspection of the 30S subunit map uncovered a previously unmodeled isoaspartyl residue in protein uS11, at the encoded residue N119 (**Figure 4**) (see Methods), marking the first identified protein backbone modification in the ribosome. While it has been known that this modification can exist in uS11 (David et al., 1999), its functional significance has remained unclear, and prior structures did not have the resolution to pinpoint its exact location. Conversion of asparagine to isoaspartate inserts an additional methylene group into the backbone (the Cβ position) and generates a methylcarboxylate side chain (Reissner and Aswad, 2003). In the present map, the shapes of the backbone density and proximal residues in the chain clearly reveal the presence of the additional methylene, allowing it to pack closely with the contacting rRNA nucleotides (**Figure 4A, 4B**).

Investigation of residues flanking the isoaspartate in uS11 reveal near-universal conservation in bacteria, chloroplasts and mitochondria (**Figure 4C, Figure 4–figure supplement 1**), suggesting that the isoaspartate may contribute to 30S subunit assembly or stability. Consistent with this idea, the isoaspartate allows high shape complementarity between

this region of uS11 and the 16S rRNA it contacts, which involves three consecutive purine-purine base pairs in bacteria (**Supplementary Table 2**), and a change in rRNA helical direction that is capped by stacking of histidine 118 in uS11 on a conserved purine (A718 in *E. coli*) (**Figure 4B, Supplementary Table 2**). Strikingly, the sequence motif in bacterial uS11 is also conserved in a domain-specific manner in archaea and eukaryotes (**Figure 4C, Figure 4–figure supplement 1**), as are the rRNA residues in close proximity to the predicted isoAsp (**Supplementary Table 2**). Remodeling the isoAsp motifs in maps from recently-published cryo-EM reconstructions of an archaeal 30S ribosomal subunit complex at 2.8 Å resolution (Nürenberg-Goloub et al., 2020) and a yeast 80S ribosome complex at 2.6 Å resolution (Tesina et al., 2020) shows that isoaspartate also seems to be present in these organisms (**Figure 4–figure supplement 2**). Taken together, the phylogenetic data and structural data indicate that the isoaspartate in uS11 is nearly universally conserved, highlighting its likely important role in ribosome assembly and function.

The *E. coli* 30S ribosomal subunit has eleven post-transcriptionally modified nucleotides in 16S rRNA, all of which can be seen in the present maps or inferred in the cases of many pseudouridines. Interestingly, two methylated nucleotides–m^7G527 and $m^6{}_2$A1519–appear hypomodified, based on the density at ~2.1 Å resolution (**Figure 1–figure supplement 5**). Loss of methylation at m^7G527, which is located near the mRNA decoding site, has been shown to confer low-level streptomycin resistance (Okamoto et al., 2007), and possibly neomycin resistance in some cases (Mikheil et al., 2012). The position of the methyl group is located in a pocket formed with ribosomal protein uS12, adjacent to the post-translationally modified Asp89, β-methylthio-Asp (Anton et al., 2008; Kowalak and Walsh, 1996), which also may be hypomodified in the present structure (**Figure 1–figure supplement 6**). Notably, loss of m^7G527 methylation is synergistic with mutations in uS12 that lead to high-level streptomycin resistance (Benítez-Páez et al., 2014; Okamoto et al., 2007). Loss of m^7G527 methylation would remove a positive charge

and open a cavity adjacent to uS12, which may contribute to resistance by shifting the equilibrium of 30S subunit conformational states to an "open" form that is thought to be hyperaccurate with respect to mRNA decoding (Ogle et al., 2002; Zaher and Green, 2010).

Within the 30S subunit platform near the P site, the two dimethylated adenosines–m^6_2A1518 and m^6_2A1519–have also been connected to antibiotic resistance. Although impacting 30S subunit assembly (Connolly et al., 2008) and ribosome function (Sharma and Anand, 2019), loss of methylation of these nucleotides also leads to kasugamycin resistance (Ochi et al., 2009). By contrast, bacteria lacking KsgA, the methyltransferase responsible for dimethylation of both nucleotides, become highly susceptible to other antibiotics including aminoglycosides and macrolides (O'Farrell and Rife, 2012; Phunpruch et al., 2013; Zou et al., 2018). In the present structure, m^6_2A1519 is singly-methylated whereas m^6_2A1518 is fully methylated (**Figure 1–figure supplement 5**). KsgA fully methylates both nucleotides in *in vitro* biochemical conditions (O'Farrell et al., 2012), but the methylation status of fully-assembled 30S subunits *in vivo* has not been determined. The loss of a single methylation of m^6_2A1519, observed here for the first time, could be a mechanism for conferring low-level antibiotic resistance to some antibiotics without appreciably affecting 30S subunit assembly or leading to sensitivity to other classes of antibiotics, a hypothesis that could be tested in the future.

Paromomycin binding in the mRNA decoding site

Aminoglycoside antibiotics (AGAs) are a widely-used class of drugs targeting the mRNA decoding site (A site) of the ribosome, making them an important focus for continued development against antibiotic resistance (Sati et al., 2019). Paromomycin, a 4,6-disubstituted 2-deoxystreptamine AGA (**Figure 5A**), is one of the best studied structurally. Structures include

paromomycin bound to an oligoribonucleotide analog of the A site at 2.5 Å resolution (Vicens and Westhof, 2001), to the small subunit of the *T. thermophilus* ribosome at 2.5 Å resolution (Kurata et al., 2008), and to the full 70S *T. thermophilus* ribosome at 2.8 Å resolution (Selmer et al., 2006). While the overall conformation of rings I-III of paromomycin is modeled in largely the same way (**Figure 5B**), the resolution of previous structures did not allow for unambiguous interpretation of ring IV, and only the oligoribonucleotide structural model of the decoding site includes some water molecules in the drug's vicinity (Vicens and Westhof, 2001). Although ring IV remains the least ordered of the four rings in the present structure, the cryo-EM map of the focus-refined 30S subunit allows high-resolution modeling of the entire molecule and the surrounding solvation for the first time (**Figure 5C, 5D**). The conformation of paromomycin in the oligonucleotide structure (Vicens and Westhof, 2001) agrees most closely with the current structure, with ring IV adopting a chair conformation with the same axial and equatorial positioning of exocyclic functional groups. However, the N6''' group in the present structure points in the opposite direction and forms multiple contacts with the backbone phosphate groups of G1489 and U1490 (**Figure 5D**). Tilting of ring IV in the present model also positions N2''' and O3''' to make contacts with the G1405 and A1406 phosphate groups, respectively. The paromomycin models in the previous structures of the 30S subunit and 70S ribosome differ further by modeling ring IV in the alternative chair conformation, which also breaks the contacts observed here. We do not see paromomycin bound in H69 of the 50S ribosomal subunit, a second known AGA binding site, consistent with prior work indicating that binding of aminoglycosides to H69 may be favored in intermediate states of ribosomal subunit rotation (Wang et al., 2012; Wasserman et al., 2015).

Ribosomal subunit interface

The ribosome undergoes large conformational changes within and between the ribosomal subunits during translation, necessitating a complex set of interactions that maintain ribosome

function. Contacts at the periphery of the subunit interface have been less resolved in many structures, likely due to motions within the ribosome populations. Additionally, some key regions involved in these contacts are too conformationally flexible to resolve in structures of the isolated subunits. In the present structure of the unrotated state of the ribosome, with tRNAs positioned in the A site and P site, improvement of maps of the individual ribosomal subunits and smaller domains within the subunits help to define these contacts more clearly. Helix H69 of the 50S subunit, which is mostly disordered in the isolated subunit, becomes better defined once the intact ribosome is formed. The 23S rRNA stem-loop closed by H69 is intimately connected to the 30S subunit at the end of 16S rRNA helix h44 near the mRNA decoding site and tRNA binding sites in the ribosome. During mRNA decoding, the loop closing H69 rearranges to form specific interactions with the A-site tRNA (Selmer et al., 2006). The stem of H69 also compresses as the 30S subunit rotates during mRNA and tRNA translocation, thereby maintaining contacts between the 30S and 50S subunits (Dunkle et al., 2011). In the present reconstructions, helix H69 seems conformationally more aligned to the 30S subunit than the 50S subunit, as the cryo-EM density for H69 is much better defined in the map of the 30S subunit compared to the map of the 50S subunit.

The loop comprising 23S rRNA nucleotides G713-A718 closing helix H34 forms an additional bridge between the 50S and 30S subunits and is also known to be dynamic in its position (Dunkle et al., 2011). In the present reconstructions this bridge is also better defined in the map of the 30S subunit compared to the 50S subunit, placing the more highly conserved arginine Arg88 in uS15 in direct contact with the RNA backbone of the H34 stem-loop rather than the less conserved Arg89 (**Figure 6A**). An additional conformationally dynamic contact between the 30S and 50S subunits involves the A-site finger (ASF, helix H38 in 23S rRNA), which is known to modulate mRNA and tRNA translocation (Komoda et al., 2006). In the present model, although

the central helical region of the ASF is only visible in low-resolution maps, loop nucleotide C888 which stacks on uS13 residues Met81 and Arg82 in the 30S subunit head domain is clearly defined (**Figure 6B**). Maps of the 30S subunit head domain and Central Protuberance of the 50S subunit also reveal clearer density defining the unrotated-state contacts between uS13 in the small subunit, uL5 in the large subunit, and bL31 (bL31A in the present structure) which spans the two ribosomal subunits.

High-resolution structural features of the 50S ribosomal subunit

The core of the 50S subunit is the most rigid part of the ribosome, which has enabled it to be modeled to higher resolution than the 30S subunit, historically and in the present structure. Recent models derived from recent cryo-EM reconstructions of the bacterial 50S ribosomal subunit reach global resolutions of 2.4 Å-2.5 Å by the map-to-model FSC criterion (Halfon et al., 2019; Stojković et al., 2020) (see Methods) (**Figure 7–figure supplement 1**). In the present 70S ribosome and 50S subunit reconstructions, which have global map-to-model resolutions of 2.06 Å and 1.95 Å respectively, the resolution of the core of the 50S subunit reaches 1.8 Å (**Figure 1, Figure 1–figure supplement 3A**), revealing unprecedented structural details of 23S rRNA, ribosomal proteins, ions and solvation (**Figure 1–figure supplement 4**, **Figure 1–figure supplement 5**, **Figure 7–figure supplement 2**). The resulting maps are also superior in the level of detail when compared to maps previously obtained by X-ray crystallography (Noeske et al., 2015; Polikanov et al., 2015), which aided in improving models of high-resolution chemical features like backbone dihedrals in much of the rRNA, non-canonical base pairs and triples, arginine side chain rotamers, and glycines in conformationally constrained RNA-protein contacts. The density also enabled modeling of thousands of water molecules, dozens of magnesium ions, and polyamines (**Figure 1–figure supplement 4, Figure 7–figure supplement 2, Table 2**). The present model now even allows for comparison of ribosome phylogenetic conservation to the level

of solvent positioning. For example, waters and ions with conserved positions in the Peptidyl Transferase Center can be seen in comparisons of different bacterial and archaeal ribosome structures, even when solvation was not included in the deposited models (Halfon et al., 2019; Polikanov et al., 2015; Schmeing et al., 2005) (**Figure 7**). The Central Protuberance (CP) of the 50S subunit, which contacts the P-site tRNA and the head of the 30S subunit, is dynamic, but is well resolved with focused refinement, here reaching a resolution of 2.29 Å in map-to-model FSC comparisons (**Figure 1–figure supplement 3**). The improved resolution of the CP aided in modeling ribosomal proteins uL5 and bL31A, as well as the CP contact to P-site tRNA.

Backbone modification in ribosomal protein uL16

Details of post-transcriptional and post-translational modifications are also clear in the 50S subunit maps (**Figure 1 - figure supplement 5**). Surprisingly, the post-translationally modified β-hydroxyarginine at position 81 in uL16 (Ge et al., 2012) is followed by unexplained density consistent with a thiopeptide bond between Met82 and Gly83 (**Figure 8A**). Adjusting the contour level of the map shows map density for the modified atom similar to that of the sulfur in the adjacent methionine and nearby phosphorus atoms in the RNA backbone, in contrast to neighboring peptide oxygen atoms. Notably, the mass for *E. coli* uL16 has been shown to be 15328.1 Da and drops to 15312.1 Da with loss of Arg81 hydroxylation (Ge et al., 2012). However, this mass is still +30.9 more than the encoded sequence (15281.2 Da, Uniprot P0ADY7). In *E. coli* uL16 is also N-terminally methylated (Brosius and Chen, 1976), leaving 16 mass units unaccounted for, consistent with the thiopeptide we observe in the cryo-EM map. We examined a high-resolution mass spectrometry bottom-up proteomics dataset (Dai et al., 2017) to find additional evidence supporting the interpretation of the cryo-EM map as a thiopeptide. Several uL16 peptides were found across multiple experiments that matched the expected mass shift closer to that of a thiopeptide's O to S conversion (+15.9772 Da) than an oxidation (+15.9949), a

common modification with a similar mass shift (**Figure 8B**). Fragmentation spectra localized the mass shift near the Met82-Gly83 bond, further supporting the presence of a thiopeptide (**Figure 8C**). Taken together, the cryo-EM map of the 50S subunit and mass spectrometry data support the model of a thiopeptide between Met82 and Gly83 in *E. coli* uL16.

The enzymes that might be responsible for insertion of the thioamide in uL16 remain to be identified. *E. coli* encodes the prototypical YcaO enzyme, which can form thiopeptides, but for which no substrate is known (Burkhart et al., 2017). A phylogenetic tree of YcaO family members shows a clear break separating YcaO proteins associated with secondary metabolism into a major branch (**Figure 8–figure supplement 1A**). A sub-grouping in the other major branch includes YcaO family members within Gammaproteobacteria (**Figure 8–figure supplement 1A**). Examination of genes in close proximity to YcaO across Gammaproteobacteria reveals three genes that form the *focA-pfl* operon involved in the anaerobic metabolism of *E. coli* (**Figure 8–figure supplement 1B**) (Sawers and Suppmann, 1992).

Discussion

High resolution cryo-EM maps are now on the cusp of matching or exceeding the quality of those generated by X-ray crystallography, opening the door to deeper understanding of the chemistry governing structure-function relationships and uncovering new biological phenomena. Questions about the ribosome, which is composed of the two most abundant classes of biological macromolecules and essential for life, reach across a diverse range of inquiry. Structural information about ribosomal components can have implications ranging from fundamental chemistry to mechanisms underlying translation and evolutionary trends across domains of life. For example, in our cryo-EM reconstructions, we observed a surprising level of detail about modifications to nucleobases and proteins that could not be seen in prior X-ray crystallographic structures. The most unexpected of these is the presence of two previously unknown post-translational modifications in the backbones of ribosomal proteins, which would be otherwise difficult to confirm without highly targeted analytical chemical approaches. Precise information about binding of antibiotics, protein-RNA contacts, and solvation are additional examples of what can be interrogated at this resolution. Beyond purely structural insights, these findings generate new questions about protein synthesis, ribosome assembly, and antibiotic action and resistance mechanisms, providing a foundation for future experiments.

The remarkable finding of a thioamide modification in protein uL16, only the second such example in a protein (Mahanta et al. 2019), is a perfect example of the power of working at <2 Å resolution. The difference in bond length of a thiocarbonyl compared to a typical peptide carbonyl is ~0.4 Å, with otherwise unchanged geometry, and is too subtle to identify at lower resolution. More conspicuously, the difference in electrostatic potential is clear at <2 Å. Analysis of previously published mass spectrometry data with sufficient mass accuracy to differentiate O to S modifications from a more common +O oxidation event (Dai et al., 2017) (**Figure 8**) corroborates

the finding. The possible role for the thiopeptide linkage in the *E. coli* ribosome remains to be shown, along with the mechanism by which its formation is catalyzed. One candidate enzyme for this purpose is *E. coli* protein YcaO, an enzyme known to carry out thioamidation and other amide transformations (Burkhart et al., 2017). Although this enzyme has been annotated as possibly participating with RimO in modification of uS12 Asp89 (Strader et al., 2011), genetic evidence for a specific YcaO function is lacking. For example, *E. coli* lacking YcaO are cold-sensitive, and have phenotypes most similar in pattern to those observed with knockout of UspG, universal stress protein 12 (Nichols et al., 2011). Furthermore, knockout of YcaO has phenotypes uncorrelated with those of knockout of YcfD, the β-hydroxylase for Arg81 in uL16 adjacent to the thioamide (Nichols et al., 2011). *YcaO*-like genes in Gammaproteobacteria genomes colocalize with the *focA-pfl* operon, a common set of genes involved in anaerobic and formate metabolism (**Figure 8–figure supplement 1**) (Sawers and Suppmann, 1992). Since the ribosomes used here were obtained from aerobically grown cultures and the *focA-pfl* operon is transcribed independently of the *YcaO* gene in *E. coli (Sawers, 2005)* it is likely that the *YcaO* gene and the *focA-pfl* operon encode proteins with unrelated functions. Interestingly, the clear phylogenetic separation between the *YcaO* gene in Gammaproteobacteria and the *YcaO* genes known to be involved in secondary metabolism in the phylogenetic tree suggests that, if YcaO is responsible for uL16 thioamidation, this modification may only be conserved in Gammaproteobacteria.

The maps of the 30S subunit, resolved to a slightly lower resolution of 2.1-2.2 Å by map-to-model FSC, enabled the identification of the only known isopeptide bond in a ribosomal protein, an isoAsp at position 119 in uS11. While isoaspartyl residues have been hypothesized to mainly be a form of protein damage requiring repair, previous work identified the existence of isoAsp in uS11 at near stoichiometric levels, suggesting it might be functionally important (David et al., 1999). Certain hotspots in protein sequences are known to be especially prone to isoaspartate

formation (Reissner and Aswad, 2003), including Asn-Gly, as encoded in nearly all bacterial uS11 sequences (**Figure 4C**). However, the half-life of the rearrangement is on the timescale of days (Robinson and Robinson, 2001; Stephenson and Clarke, 1989). In archaea and eukaryotes, formation of the isoaspartate would require dehydration of the encoded aspartate, which occurs even more slowly than deamidation of asparagine (Stephenson and Clarke, 1989). Importantly, the residue following the aspartate is nearly always serine in eukaryotes, and is enriched for glycine, serine and threonine in archaea (**Figure 4C**, **Figure 4–figure supplement 1**), consistent with the higher rates of dehydration that occur when aspartate is followed by glycine and serine in peptide models (Stephenson and Clarke, 1989). These results suggest that the isoAsp modification may be nearly universally conserved in all domains of life. Concordant with this hypothesis, isoAsp modeling provides a better fit to cryo-EM maps of uS11 in archaeal and eukaryotic ribosomes (**Figure 4–figure supplement 2**). Although it is possible that isoaspartate formation could be accelerated in specific structural contexts (Reissner and Aswad, 2003), it is not clear if the isoAsp modification in uS11 occurs spontaneously or requires an enzyme to catalyze the reaction. *O*-methyl-transferase enzymes have been identified that install a β-peptide in a lanthipeptide (Acedo et al., 2019) or serve a quality control function to remove spontaneously formed isoaspartates (David et al., 1999). Deamidases that catalyze isoAsp formation from asparagine are not well described in the literature, although examples have been identified in viral pathogens, possibly repurposing host glutamine amidotransferases (Zhao et al., 2016). Future work will be needed to identify the mechanisms by which the isoAsp in uS11 is generated in cells. Its biological significance, whether in 30S subunit assembly or other steps in translation, also remains to be defined.

The resolution achieved here also has great potential for better informing structure activity relationships in future antibiotic research, particularly because the ribosome is so commonly

targeted. For example, we were able to identify hypomodified bases in 16S rRNA (m^7G527 and $m^6{}_2$A1519) and possible hypomodification of Asp89 (β-methylthio-Asp) in uS12 (**Figure 1–figure supplement 5, Figure 1–figure supplement 6**). These hypomodifications could confer antibiotic resistance in some cases. Furthermore, we were also able to see more clearly the predominant position of paromomycin ring IV in the decoding site of the 30S subunit (**Figure 5**). The proposed primary role of ring IV has been to increase the positive charge of the drug to promote binding (Hobbie et al., 2006), in line with its ambiguous modeling in previous structures (Kurata et al., 2008; Selmer et al., 2006; Vicens and Westhof, 2001). While ring IV's features in the current map are weaker relative to those of rings I-III, we were able to identify interactions of ring IV with surrounding 16S rRNA nucleotides and ordered solvent molecules that were not previously modeled. Importantly, the observed interactions between the N6''' amino group and the phosphate backbone of nucleotides G1489-U1490 in particular are likely responsible for known susceptibility of PAR to N6''' modification (Sati et al., 2017). While the same loss of interactions is expected for neomycin, which differs from paromomycin only by the presence of a 6'-hydroxy rather than a 6'-amine in ring I, the penalty of modifying N6''' in neomycin is likely compensated for by the extra positive charge and stronger hydrogen bonding observed with neomycin ring I (Sati et al., 2017). The level of detail into modes of aminoglycoside binding that can now be obtained using cryo-EM thus should aid the use of chemical biology to advance AGA development.

The cryo-EM maps of the 30S subunit also revealed new structural information about protein bS21 at lower resolution, particularly at its C terminus. The location of bS21 near the ribosome binding site suggests it may play a role in translation initiation. The conservation of the RLY (or KLY) motif and its contacts to the 30S subunit head domain also suggests bS21 may have a role in modulating conformational dynamics of the head domain relative to the body and platform of the 30S subunit. Rearrangements of the 30S subunit head domain are seen in every

stage of the translation cycle (Javed and Orlova, 2019). Although we could align putative S21 homologues from huge phages (Al-Shayeb et al., 2020) with specific bacterial clades, and show that many also possess KLY-like motifs, there were no clear relationships between the predicted consensus ribosome binding sites in these bacteria and these phages. It is possible that bS21 and the phage homologues interact with nearby mRNA sequences 5' of the Shine-Dalgarno helix, affecting translation initiation in this way. Taken together, the structural and phylogenetic information on bS21 and the phage S21 homologues raise new questions about their role in translation and the phage life cycle, i.e. whether they contribute to specialized translation and/or help phage evade bacterial defenses.

Radiation damage to nucleic acids has been extensively studied because of its clinical relevance. Experiments using high-energy X-ray or electron pulses showed damage to be caused principally by low-energy secondary electrons rather than as a direct result of irradiation (Ma et al., 2019). In our structure evidence for such damage appears as a clear break in the ring density for numerous riboses and is generally surrounded by otherwise strong density for bases, phosphates, and amino acids. Breaks appeared to be most commonly found between atoms C1' and C2' or between C3' and C4', and less frequently to and from O4' (**Figure 1 - figure supplement 7**), although the exact number of each was not quantified. While all examples shown correspond to breaks in guanosine nucleotides, the base-dependence of ribose damage will require a more detailed analysis. The appearance of these breaks across varying bond positions in a structure of many averaged images suggests a preference for cleavage to occur at a specific site in different contexts, but this observation, and the underlying chemistry of ribose damage under the electron beam, remain to be fully explored.

With regard to resolution of cryo-EM maps, in this work we have emphasized the use of map-to-model FSC curves in addition to the "gold-standard" FSC. In general, the consensus is that there is no substitute for visual inspection of the map to determine its quality. However, as higher resolution becomes more commonplace it is necessary to make use of standards that are relevant on the scale of the questions being answered. Sub-Ångstrom differences in nominal resolution as reported by half-map FSCs have significant bearing on chemical interactions at face value but may lack usefulness if map correlation with the final structural model is not to a similar resolution. While map-to-model FSC carries intrinsic bias from the model's dependence on the map, in a high resolution context it does provide additional information about the overall confidence with which to interpret the model, not captured in half-map FSCs. This is perhaps a reason why it appears in recent work more focused on methods and tool development (Nakane et al., 2020; Terwilliger et al., 2020a, 2020b).

Aside from global high resolution, the conformational heterogeneity of the ribosome also calls attention to the tools used for working with complexes that display variable resolution. Methods for refining heterogeneous maps have proliferated, including multi-body refinement (Nakane et al., 2018) and 3D variability analysis (Punjani and Fleet, 2020) among others, but ways to work with and create a model from many maps of the same complex have not yet been standardized and require substantial manual intervention. Additionally, as it tends to be most convenient to develop tools with the highest resolution possible, it is common for new methods to utilize very high resolution maps of apoferritin as a standard, which is highly symmetric and well-ordered. Specimens that do not share these characteristics may require tools that violate others' assumptions, for instance, flat solvent outside the domain of interest. For the present 70S map, the half-map FSC ≥ 0.97 up to 3.3 Å resolution, which is estimated to represent a theoretical correlation with a "perfect" map up to 0.99 (Terwilliger et al., 2020a). Structural information with

certainty up to so-called "near-atomic" resolution has potential use in benchmarking newer tools and may specifically make our results valuable in addressing issues with focused or multi-body refinement. This structure also has potential use for aiding the future development of *de novo* RNA modeling tools, which are historically less developed compared to similar tools for proteins, and often rely on information generated from lower-resolution RNA structures (Watkins and Das, 2019). Finally, our micrographs uploaded to the Electron Microscopy Public Image Archive (EMPIAR) (Iudin et al., 2016) should serve as a resource for ribosome structural biologists and the wider cryo-EM community to build on the present results.

Methods

Biochemical preparation

E. coli 70S ribosome purification (Travin et al., 2019) and tRNA synthesis, purification, and charging (Ad et al., 2019) were performed as previously described. Briefly, 70S ribosomes were purified from *E. coli* MRE600 cells using sucrose gradients to isolate 30S and 50S ribosomal subunits, followed by subunit reassociation and a second round of sucrose gradient purification. Transfer RNAs were transcribed from PCR DNA templates using T7 RNA polymerase, and purified by phenol-chloroform extraction, ethanol precipitation, and column desalting. Flexizyme ribozymes were used to charge the P-site $tRNA^{fMet}$ with either pentafluorobenzoic acid or malonate methyl ester and the A-site $tRNA^{Val}$ with valine (Goto et al., 2011). Ribosome-mRNA-tRNA complexes were formed non-enzymatically by incubating 10 μM P-site tRNA, 10 μM mRNA, and 100 μM paromomycin with 1 μM ribosomes for 15 min at 37 °C in Buffer AC (20 mM Tris pH 7.5, 100 mM NH_4Cl, 15 $MgCl_2$, 0.5 mM EDTA, 2 mM DTT, 2 mM spermidine, 0.05 mM spermine). Then, 10 μM A-site tRNA was added and the sample was incubated for an additional 15 min at 37 °C. Complexes were held at 4 °C and diluted to 100 nM ribosome concentration in the same buffer immediately before grid preparation. The mRNA of sequence 5'-GUAUAA**GGAGG**UAAAA*AUGGUA*UAACUA-3' was chemically synthesized (IDT) and was resuspended in water without further purification. The Shine-Dalgarno sequence is shown in bold, and the Met-Val codons are in italics.

Cryo-EM sample preparation

300 mesh R1.2/1.3 UltraAuFoil grids from Quantifoil with an additional amorphous carbon support layer were glow discharged in a Pelco sputter coater. 4 μL of each sample was deposited onto grids and incubated for 1 minute, then washed in Buffer AC with 20 mM NH_4Cl rather than

100 mM NH_4Cl. Grids were plunge-frozen in liquid ethane using a Vitrobot Mark IV with settings: 4 °C, 100% humidity, blot force 6, blot time 3.

Data Acquisition

Movies were collected on a 300-kV Titan Krios microscope with a GIF energy filter and Gatan K3 camera. Super-resolution pixel size was 0.355 Å, for a physical pixel size of 0.71 Å. SerialEM (Schorb et al., 2019) was used to correct astigmatism, perform coma-free alignment, and automate data collection. Movies were collected with the defocus range -0.6 to -1.5 µm and the total dose was 39.89 $e^-/Å^2$ split over 40 frames. One movie was collected for each hole, with image shift used to collect a series of 3x3 holes for faster data collection (Cheng et al., 2018).

Image processing

Datasets of 70S ribosome complexes with the two differently charged P-site tRNAs were initially processed separately. Movies were motion-corrected and binned to the recorded physical pixel size (0.71 Å) within RELION 3.0 (Scheres, 2012) using MotionCor2 (Zheng et al., 2017). CTF estimation was done with CTFFind4 (Rohou and Grigorieff, 2015), and micrographs with poor CTF fit as determined by visual inspection were rejected. Particles were auto-picked with RELION's Laplacian-of-Gaussian method. 2D classification was performed in RELION. Particles were separated into 3D classes in cryoSPARC heterogeneous refinement (Punjani et al., 2017), using an initial model generated from PDB 1VY4 with A- and P-site tRNAs (Polikanov et al., 2014) and keeping particles that were sorted into well-resolved 70S ribosome classes. Masked 3D classification without alignment was used in RELION to further sort particles based on A-site tRNA occupancy. CTF Refinement and Bayesian polishing were performed in RELION 3.1 before pooling the two datasets together, with nine optics groups defined based on the 3x3 groups for

image shift-based data collection. The resulting 70S ribosome reconstruction was used as input for focused refinements of the 50S and 30S subunits. Focused refinement of the Central Protuberance was performed starting from the 50S subunit-focused refinement reconstruction, and head- and platform-focused refinements started from the 30S subunit focused refinement reconstruction.

Modeling

The previous high-resolution structure of the *E. coli* 70S ribosome (Noeske et al., 2015) was used as a starting model. We used the "Fit to Map" function in Chimera (Pettersen et al., 2004) to adjust the voxel size of the cryo-EM map of the 50S ribosomal subunit generated here to maximize correlation, resulting in a pixel size of 0.7118 Å rather than the recorded 0.71 Å. Focused-refined maps were transformed into the frame of reference of the 70S ribosome for modeling and refinement, using the "Fit to Map" function in Chimera, and resampling the maps on the 70S ribosome grid. The 50S and 30S subunits were refined separately into their respective focused-refined maps using PHENIX real-space refinement (RSR) (Liebschner et al., 2019). Protein and rRNA chains were visually inspected in Coot (Casañal et al., 2020) and manually adjusted where residues did not fit well into the density. Focused-refined maps on smaller regions were used to make further manual adjustments to the model, alternating with PHENIX RSR. Some parts of the 50S subunit, including H69, H34 and the tip of the A-site finger, were modeled based on the 30S subunit focused-refined map. The A-site and P-site tRNAs were modeled as follows: anticodon stem-loops, 30S subunit focused-refined map; P-site tRNA body, 50S subunit focused-refined map, with a B-factor of 20 $Å^2$ applied; A-site tRNA body, 30S subunit focused-refined map and 50S subunit focused-refined map with B-factors of 20 $Å^2$ applied; tRNA-ACCA 3' ends, 50S subunit focused-refined map with B-factors of 20-30 $Å^2$ applied. Alignments of uS15 were generated using BLAST (Altschul et al., 1997b) with the *E. coli* sequence as reference. The model

for bL31A (*E. coli* gene rpmE) was manually built into the CP and 30S subunit head domain focused-refined maps before refinement in PHENIX.

A model for paromomycin was manually docked into the 30S subunit focused-refined map, followed by real-space refinement in Coot and in PHENIX. Comparisons to prior paromomycin structural models (PDB codes 1J7T, 2VQE, and 4V51) (Kurata et al., 2008; Selmer et al., 2006; Vicens and Westhof, 2001) used least-squares superposition of paromomycin in Coot. Although ring IV is in different conformations in the various paromomycin models, the least-squares superposition is dominated by rings I–III, which are in nearly identical conformations across models.

Ribosome solvation including waters, magnesiums, and polyamines was modeled using a combination of PHENIX (phenix.douse) and manual inspection. The phenix.douse feature was run separately on individual focused-refined maps, and the resulting solvent models were combined into the final 30S subunit and 50S subunit models. Due to the fact that the solvent conditions used here contained ammonium ions and no potassium, no effort was made to systematically identify monovalent ion positions. The numbers of various solvent molecules are given in **Table 2**.

Modeling of isoAsp residues in uS11

Initial real-space refinement of the 30S subunit against the focused-refined map using PHENIX resulted in a single chiral volume inversion involving the backbone of N119 in ribosomal protein uS11, indicating that the L-amino acid was being forced into a D-amino acid chirality, as

reported by phenix.real_space_refine. Inspection of the map in this region revealed clear placement for carbonyl oxygens in the backbone, and extra density consistent with an inserted methylene group, as expected for isoAsp. The model of isoAsp was refined into the cryo-EM map using PHENIX RSR. IsoAsp was also built and refined into models of archaeal and eukaryotic uS11 based on cryo-EM maps of an archaeal 30S ribosomal subunit complex (PDB 6TMF) (Nürenberg-Goloub et al., 2020) and a yeast 80S ribosome complex (PDB 6T4Q) (Tesina et al., 2020). These models were refined using PHENIX RSR, and real-space correlations by residue calculated using phenix.model_map_cc.

Phylogenetic analysis of uS11 and its rRNA contacts

All archaeal genomes were downloaded from the NCBI genome database (2,618 archaeal genomes, last accessed September 2018). Due to the high number of bacterial genomes available in the NCBI genome database, only one bacterial genome per genus (2,552 bacterial genomes) was randomly chosen based on the taxonomy provided by the NCBI (last accessed in December 2017). The eukaryotic dataset comprises nuclear, mitochondrial and chloroplast genomes of 10 organisms (*Homo sapiens, Drosophila melanogaster, Saccharomyces cerevisiae, Acanthamoeba castellanii, Arabidopsis thaliana, Chlamydomonas reinhardtii, Phaeodactylum tricornutum, Emiliania huxleyi, Paramecium aurelia* and *Naegleria gruberi*).

Genome completeness and contamination were estimated based on the presence of single-copy genes (SCGs) as described in (Anantharaman et al., 2016). Only genomes with completeness >70% and contamination <10% (based on duplicated copies of the SCGs) were kept and were further de-replicated using dRep at 95% average nucleotide identity (version

v2.0.5) (Olm et al., 2017). The most complete genome per cluster was used in downstream analyses.

Ribosomal uS11 genes were detected based on matches to the uS11 Pfam domain (PF00411) (Punta et al., 2012) using hmmsearch with an E-value below 0.001 (Eddy, 1998). Amino acid sequences were aligned using the MAFFT software (version v7.453) (Katoh and Standley, 2016). Alignment was further trimmed using Trimal (version 1.4.22) (--gappyout option) (Capella-Gutiérrez et al., 2009). Tree reconstruction was performed using IQ-TREE (version 1.6.12) (Nguyen et al., 2015), using ModelFinder (Kalyaanamoorthy et al., 2017) to select the best model of evolution, and with 1000 ultrafast bootstrap (Hoang et al., 2018). The tree was visualized with iTol (version 4) (Letunic and Bork, 2019) and logos were made using the weblogo server (Crooks et al., 2004).

16S and 18S rRNA genes were identified from the prokaryotic and eukaryotic genomes using the method based on hidden Markov model (HMM) searches using the cmsearch program from the Infernal package (Nawrocki et al., 2009) and fully described in (Brown et al., 2015). The sequences were aligned using the MAFFT software.

Phylogenetic analysis of bS21 and phage S21 homologues

S21 sequences were retrieved from the huge phage database described in (Al-Shayeb et al., 2020). Cd-hit was run on the set of S21 sequences to reduce the redundancies (Fu et al., 2012) (default parameters) (version 4.8.1). Non redundant sequences were used as a query against the database of prokaryotic genomes used for uS11 above using BLASTP (version

2.10.0+) (e-value 1e-20) (Altschul et al., 1997a). Alignment and tree reconstruction were performed as described for uS11 except that we did not perform the alignment trimming step.

Phylogenetic analysis of YcaO genes

Similarly to uS11, the YcaO sequences were identified in prokaryotic genomes based on its PFAM accession (PF02624) (Punta et al., 2012) using hmmsearch with an E-value below 0.001 (Eddy, 1998). Amino acid sequences were aligned using the MAFFT software (version v7.453) (Katoh and Standley, 2016). Alignment was further trimmed using Trimal (version 1.4.22) (--gappyout option) (Capella-Gutiérrez et al., 2009). Tree reconstruction was performed using IQ-TREE (version 1.6.12) (Nguyen et al., 2015), using ModelFinder (Kalyaanamoorthy et al., 2017) to select the best model of evolution, and with 1000 ultrafast bootstraps (Hoang et al., 2018). The tree was visualized with iTol (version 4) (Letunic and Bork, 2019). The three genes downstream and upstream of each *YcaO* gene were identified and annotated using the PFAM (Punta et al., 2012) and the Kegg (Kanehisa et al., 2016) databases.

Map-to-model FSC calculations

Refined PDB coordinates for the 70S ribosome, individual ribosomal subunits or domains (30S subunit head, 50S subunit Central Protuberance) were used for comparisons to the 70S ribosome map or focused-refined maps. Masks for each map were generated in Chimera (Pettersen et al., 2004) using the relevant PDB coordinates as follows. A 10 Å resolution map from the coordinates was calculated using molmap, and the surface defined at 1 standard deviation was used to mask the high-resolution map. The effective global resolution of a given map is given at the FSC cutoff of 0.5 in **Table 2** and **Figure 1–figure supplements 2-3**.

Map-to-model comparisons for other 50S subunit reconstructions (emd_20353, emd_10077)

For the recent *E. coli* 50S subunit structure (Stojković et al., 2020), we used Chimera to resize the deposited map (emd_20353) to match the dimensions of the maps presented here. Briefly, our atomic coordinates for the 50S subunit were used with the "Fit to Map" function and the voxel size of the deposited map was adjusted to maximize correlation. The resulting voxel size changed from 0.822 Å to 0.8275 Å in linear dimension. After rescaling the deposited map, we used phenix.model_map_cc to compare the map with rescaled atomic coordinates deposited in the PDB (6PJ6) or to the present 50S model, yielding a map-to-model FSC of 0.5 at ~2.5 Å. Similar comparisons of the structure of the *Staphylococcus aureus* 50S subunit to the deposited map (PDB 6S0Z, emd_10077) (Halfon et al., 2019) yielded a map-to-model FSC of 0.5 at 2.43 Å, accounting for a change in voxel linear dimension from 1.067 Å to 1.052 Å.

Analysis of uL16 mass spectrometry datasets

Previously published *E. coli* tryptic peptide mass spectrometry (MS/MS) raw data was used for the analysis (Dai et al., 2017) (MassIVE accession: MSV000081144). Peptide searches were performed with MSFragger (Kong et al., 2017) using the default parameters for a closed search with the following exceptions: additional variable modifications were specified on residues R (hydroxylation, Δ mass: 15.9949) and M (thioamide, Δ mass: 15.9772), maximum modifications per peptide set to four, and multiple modifications on a residue were allowed. Spectra were searched against a database of all *E. coli* proteins plus common contaminants concatenated to a decoy database with all original sequences reversed. Results were analyzed using TPP (Deutsch et al., 2015) and Skyline (Pino et al., 2020).

Figure preparation

Cryo-EM maps were supersampled in Coot for smoothness. Figure panels showing structural models were prepared using Pymol (Schrödinger) and ChimeraX (Goddard et al., 2018). Sequence logo figures made with WebLogo 3.7.4 (Crooks et al., 2004). Phylogenetic trees were visualized with iTol (version 4) (Letunic and Bork, 2019) and multiple alignments were visualized with geneious 9.0.5 (https://www.geneious.com).

Data deposition

Ribosome coordinates have been deposited in the Protein Data Bank (entry ___), maps in the EM Database (entries ____), and micrographs in EMPIAR (entry ___).

Acknowledgments

We thank Dan Toso and Paul Tobias for assistance with cryo-EM data collection, Kyle Hoffman and Dieter Söll for supplying tRNAs, Andrew Cairns and Aaron Featherstone for monomer synthesis, Dieter Söll and Nikolay Aleksashin for comments on the manuscript, and Douglas Mitchell for discussions on the thioamide linkage. This work was funded primarily by the Center for Genetically Encoded Materials (NSF No. CHE-2021739) with additional contributions from NIH No. GM R01-114454 (F.W. support). O.A. was supported in part by Agilent Technologies as an Agilent Fellow. R.M. and J.F.B. were supported by the Innovative Genomics Institute at Berkeley and the Chan Zuckerberg Biohub.

Figures

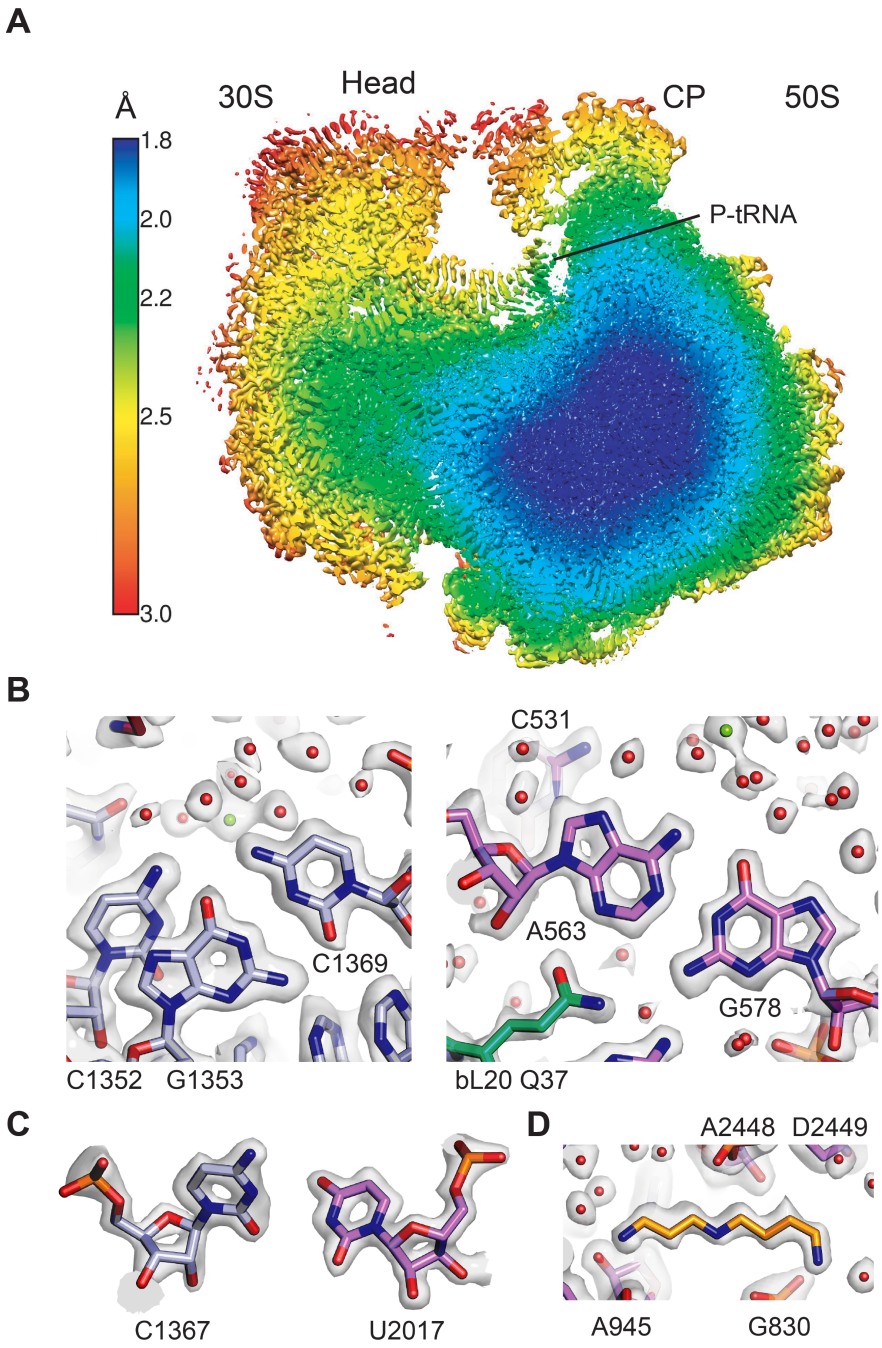


Figure 1. Overall structure of the 70S ribosome and cryo-EM map quality. (**A**) Cutaway view through the local resolution map of the 70S ribosome reconstruction. (**B**) Base pair density in the cores of the 50S and 30S ribosomal subunits. Examples demonstrate overall high resolution of base pairs and nearby solvation and Mg^{2+} sites. B-factors of -15 $Å^2$ and -10 $Å^2$ were applied to the RELION post-processed 50S subunit- and 30S subunit head-focused maps, respectively. (**C**) Nucleotide riboses in the core of 50S subunit and 30S subunit. A B-factor of -10 $Å^2$ was applied to 30S density after post-processing. (**D**) Cryo-EM density of the 50S subunit showing the polyamine spermidine.

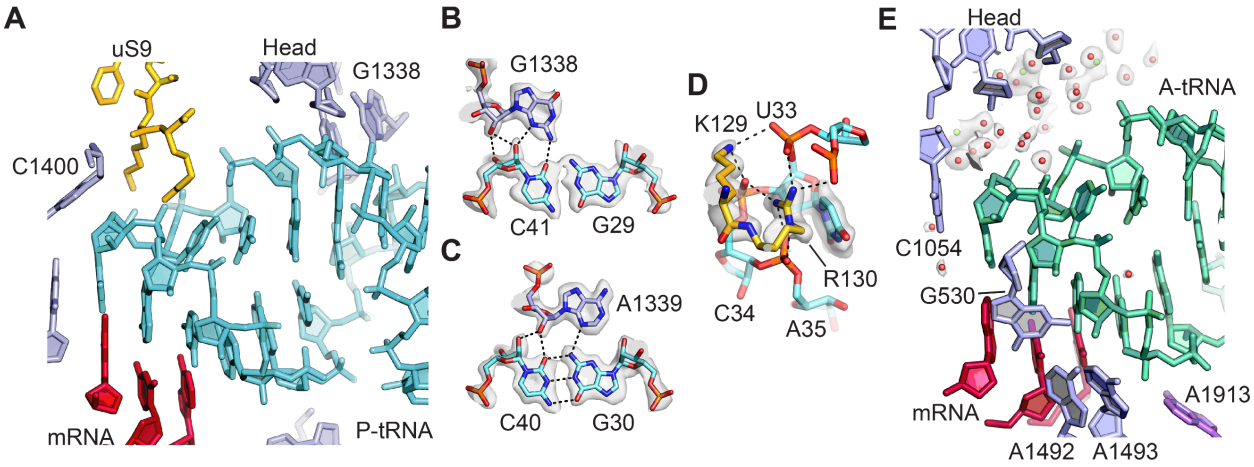
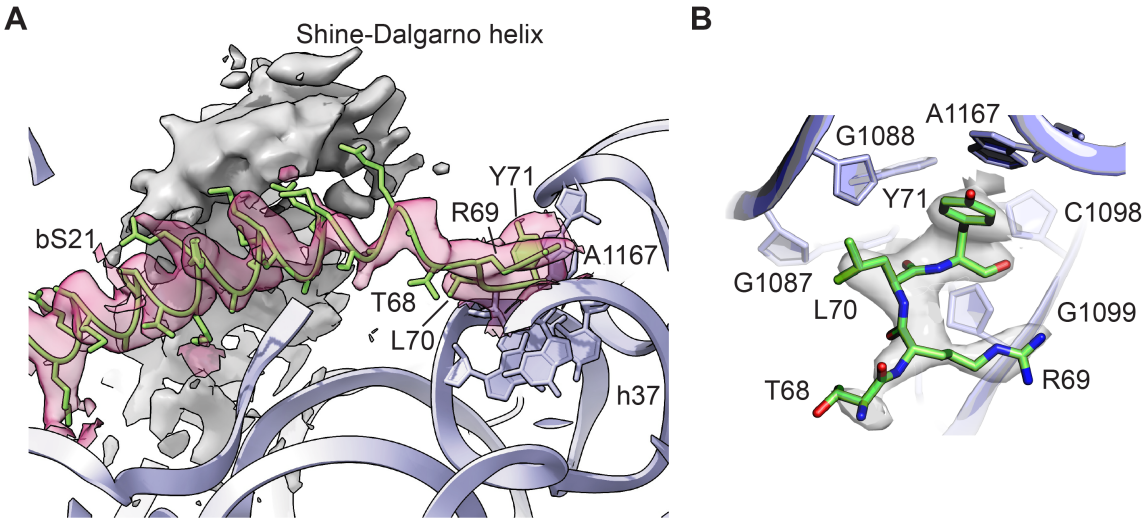


Figure 2. tRNA binding to the 30S ribosomal subunit. (**A**) Overall view of P-site tRNA anticodon stem-loop (ASL, cyan), mRNA (red), 16S rRNA nucleotides (light purple), and uS9 residues (gold). (**B**) Interactions between 30S subunit head nucleotide G1338 with P-tRNA ASL. (**C**) Interactions between 30S subunit head nucleotide A1339 with P-site ASL. (**D**) Interactions between P-tRNA ASL and protein uS9. Arg130 is observed stacking with nucleotide U33 of the ASL and forming hydrogen bonds with backbone phosphate groups. (**E**) Solvation of A-site tRNA near the 30S subunit head domain. A-site tRNA ASL in green, 16S rRNA in light purple, mRNA in purple-red. Water oxygen atoms in red spheres and Mg^{2+} in green spheres. Maps shown in panels **B**-**E** are from 30S subunit head-focused refinement.



C

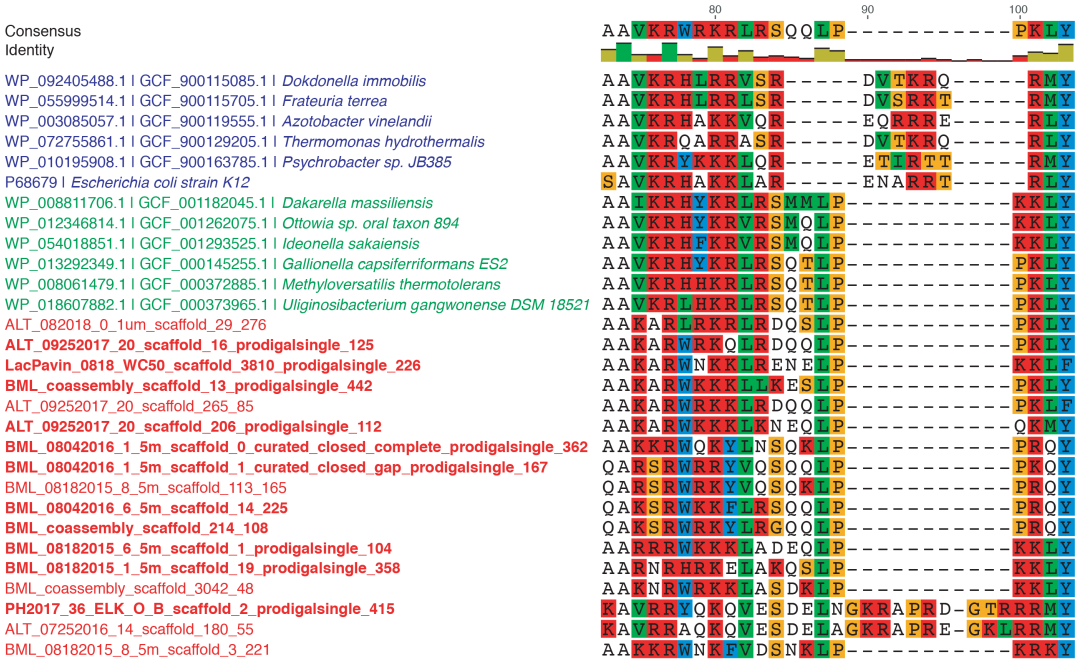


Figure 3. Protein bS21 interactions with 30S ribosomal subunit head domain. (**A**) bS21 C-terminal structure in the 30S subunit head-focused map, with Shine-Dalgarno helical density shown in grey and density for bS21 in rose. Low-pass filtering to 3.5 Å resolution was applied to clarify helical density. 16S rRNA shown in light purple ribbon and bS21 shown in light green. 16S rRNA bases that interact with RLY motif are shown in stick representation. (**B**) Closeup of RLY motif of bS21 and contacts with 16S rRNA h37 and A1167. (**C**) Protein bS21 sequence alignment near the C-terminus, along with associated phage S21 sequences. Gammaproteobacteria (blue), Betaproteobacteria (green), and phage (red) are shown.

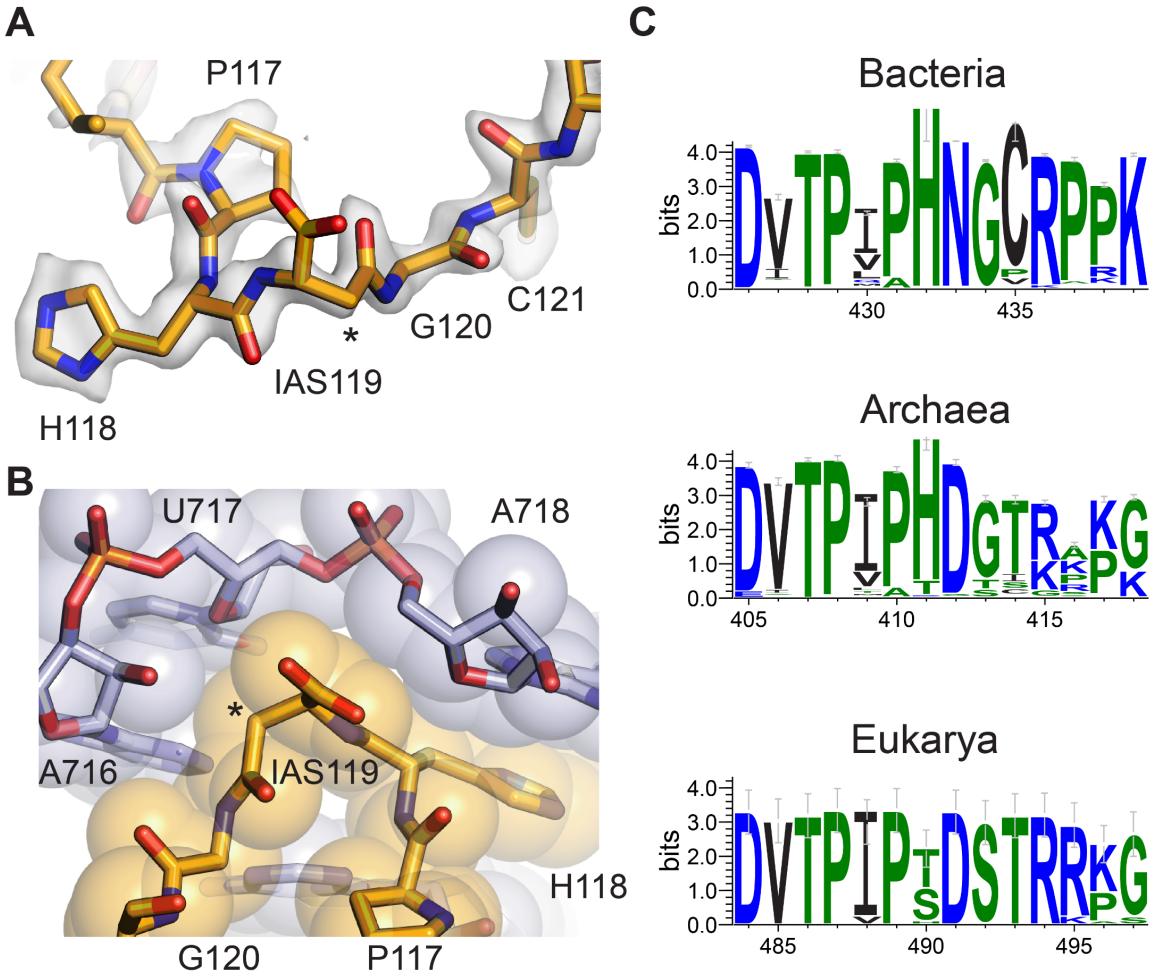


Figure 4. Isoaspartyl residue in protein uS11. (**A**) Model of isoAsp at residue position 119 in uS11, with nearby residues and cryo-EM density from the 30S subunit platform-focused refinement. Weak density for the carboxylate is consistent with effects of damage from the electron beam. The asterisk indicates the position of the additional backbone methylene group. (**B**) Shape complementarity between uS11 and 16S rRNA nucleotides surrounding IsoAsp119. 16S rRNA is shown in light purple and uS11 in orange, with atomistic coloring for the stick model. (**C**) Sequence logos of conserved amino acids spanning the putative isoAsp residue in all three domains of life.

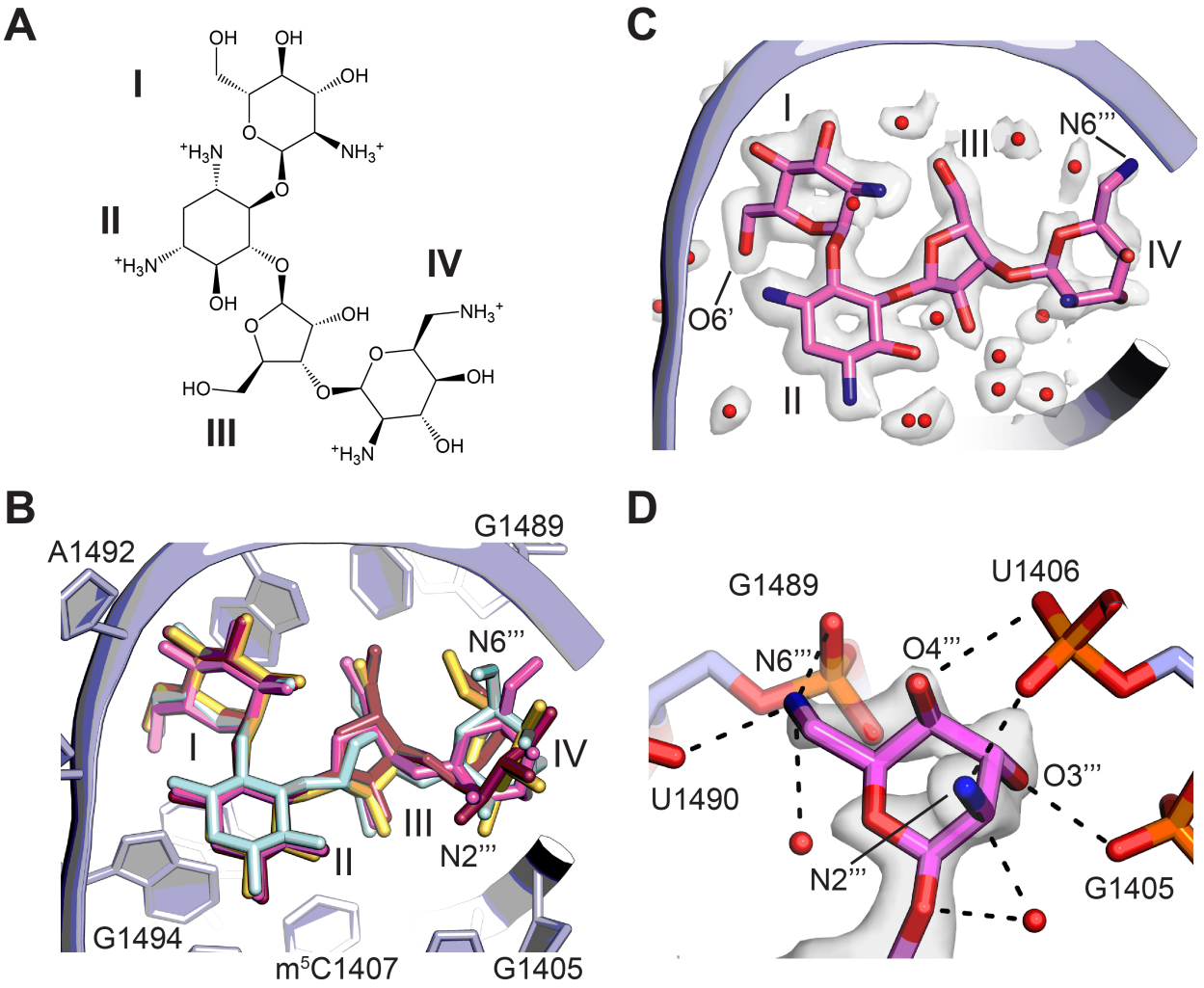


Figure 5. Binding of paromomycin to the mRNA decoding site in the 30S subunit. (**A**) Chemical structure of paromomycin (PAR) with ring numbering. (**B**) Comparison of paromomycin conformations in different structures. Paromomycin from three prior structural models (Kurata et al., 2008; Selmer et al., 2006; Vicens and Westhof, 2001), shown in yellow, dark pink, and light blue, respectively, superimposed with the present model of paromomycin, shown in pink. The binding pocket formed by 16S rRNA is shown in light purple. (**C**) Overall positioning of PAR within the binding site including solvation. (**D**) Paromomycin ring IV contacts to the phosphate backbone in 16S rRNA helix h44. Dashed lines denote contacts within hydrogen-bonding distance. Map was blurred with a B-factor of 10 Å^2.

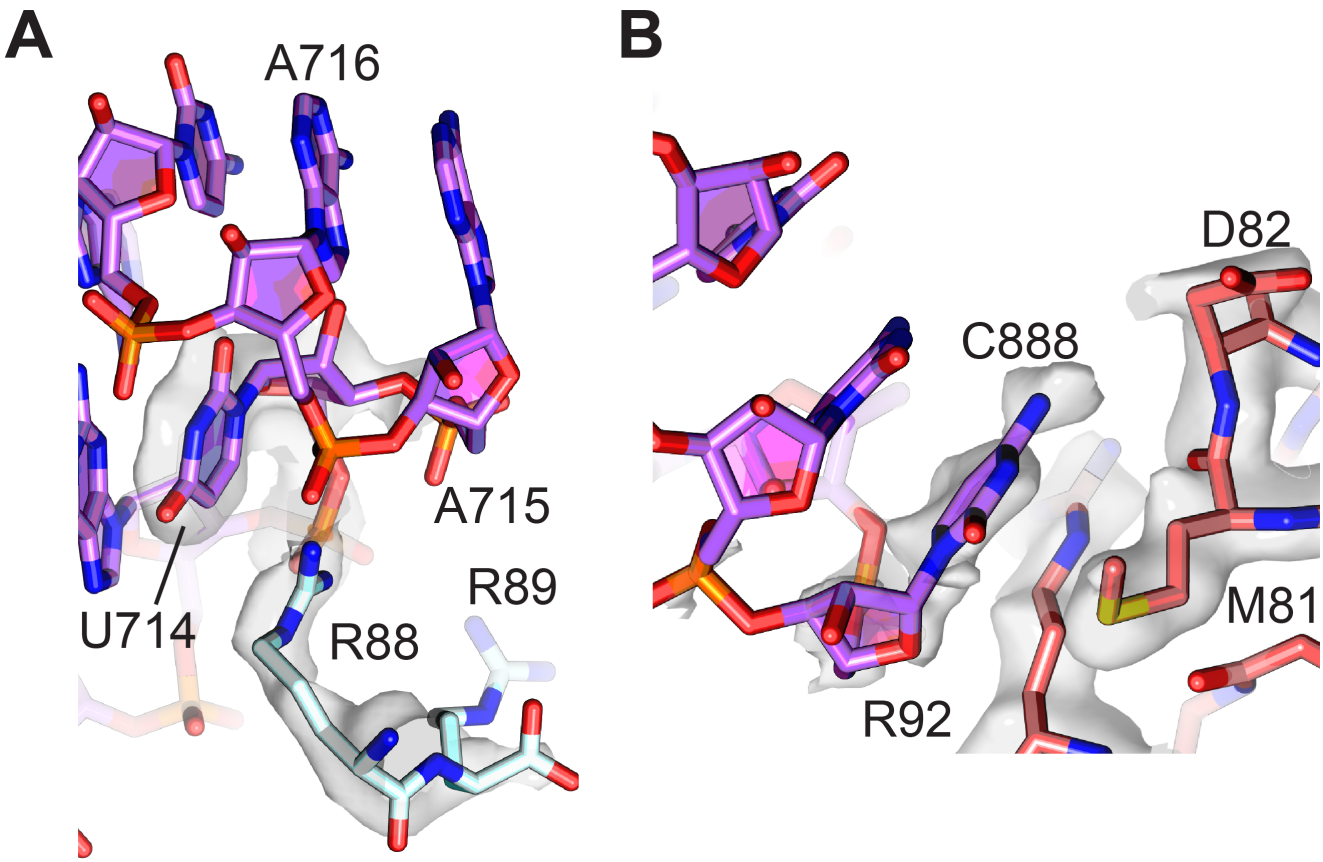


Figure 6. Peripheral contacts between the 30S and 50S subunits. (**A**) Interaction of the C-terminus of uS15 (light blue) with 23S rRNA nucleotides 713-715 (purple). The 30S subunit cryo-EM map is shown with a B-factor of 20 Å^2 applied. (**B**) Interaction between uS13 (salmon) and the A-Site Finger hairpin loop nucleotide C888 in the 50S subunit (purple). A B-factor of 10 Å^2 was applied to the head-focused map.

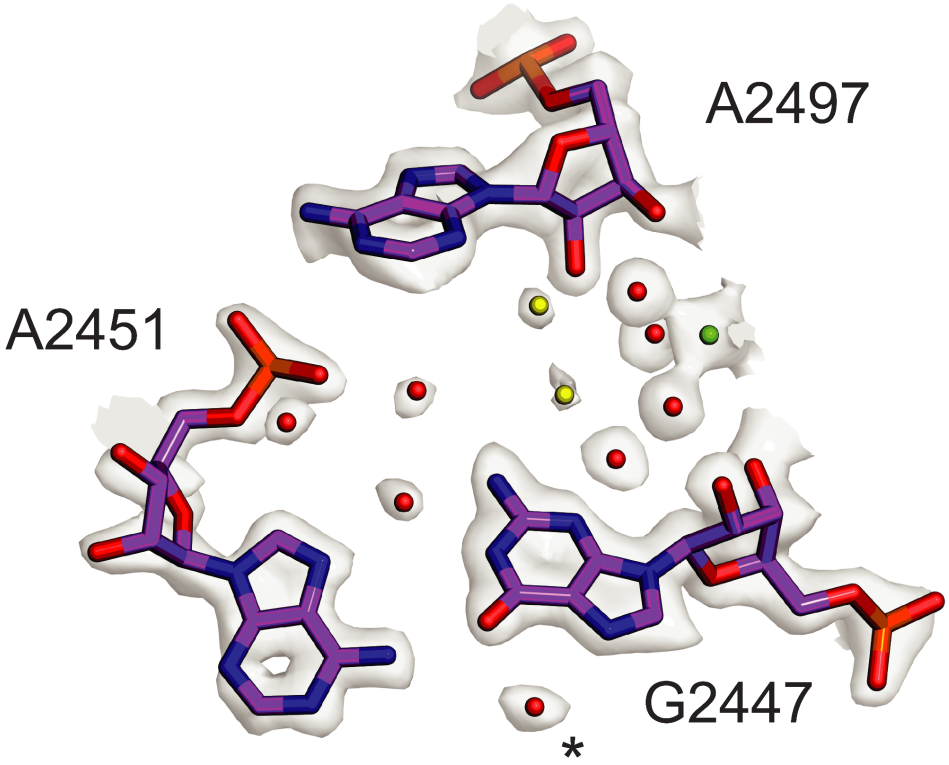
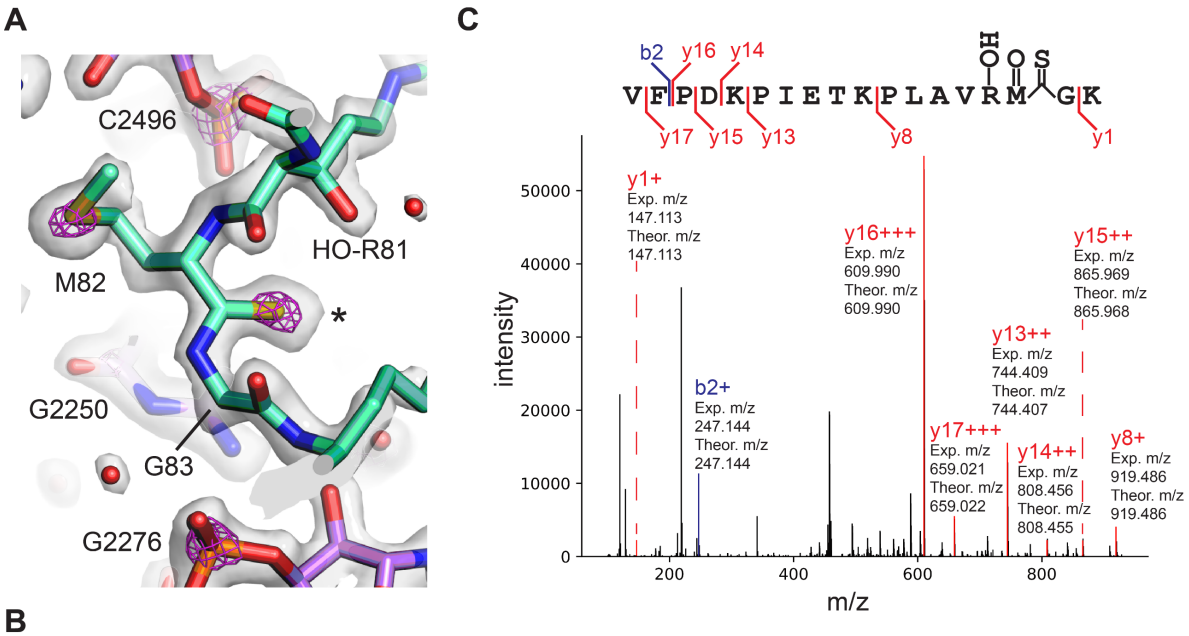


Figure 7. Conserved solvation in the PTC in the 50S ribosomal subunit. Comparison of solvation in the PTC near *E. coli* nucleotide G2447 to that in phylogenetically diverse 50S subunits. Solvent molecules conserved in bacterial ribosomes from *E. coli*, *S. aureus*, and *T. thermophilus* (Halfon et al., 2019; Polikanov et al., 2015) and in the archaeal 50S subunit from *H. marismortui* (Schmeing et al., 2005) are colored red. Waters conserved in 3 of 4 structures are colored yellow. Mg^{2+} is shown in green. Asterisk (*) denotes density modeled as K^{+} in the *H. marismortui* 50S subunit structure.



B

Peptide	Experimental Mass (Da)	Theoretical Mass (Da)	Mass Diff. (ppm)
VFPDKPITEKPLAVR[O]M[O,O->S]GK	2073.1106	2073.1113	-0.3377
VFPDKPITEKPLAVR[O]	1724.9823	1724.9823	0.0
VFPDKPITEKPLAVR	1708.9865	1708.9874	-0.5266
VFPDKPITEKPLAVRM[O,O->S]GK	2057.1174	2057.1165	0.4375
VFPDKPITEKPLAVR[O]M[O,O]GK (hypothetical)	2073.1106	2073.1290	**-8.8756**

Figure 8. Thioamide modification in protein uL16. (**A**) Structural model of thioamide between Met82 and Gly83 in uL16 (mint), with cryo-EM density contoured at two levels to highlight sulfur and phosphorus atoms. The lower contour level is shown as a grey surface and the higher contour level is shown as fuchsia mesh. 23S rRNA is shown in purple. (**B**) LC-MS/MS data supporting the presence of a thioamide bond between M82 and G83 of uL16 (Dai et al., 2017). Shown are selected uL16 peptides with designated modifications found in the spectral search and their associated experimental masses, theoretical masses, and mass differences. All peptides were found in multiple fractions and replicates of the experiment. The final row shows a hypothetical peptide identical to the first row, except carrying an oxidation modification instead of O to S replacement. (**C**) Annotated fragmentation spectra from the LC-MS/MS experiment showing a uL16 peptide with a thioamide bond. Peptide is assigned modifications of: oxidation on M, oxidation on R, and a thiopeptide between M and G. Fragmentation ions are annotated with experimental and theoretical m/z ratios.

Table 1. Data collection and processing.

Magnification	109,160
Voltage (kV)	300
Spherical aberration (mm)	2.7
Electron exposure (e–/Å2)	39.89
Defocus range (μm)	-0.6/ -1.5
Pixel size (Å)	0.7118
Symmetry imposed	C1
Initial particle images (no.)	874,943
Final particle images (no.)	307,495
Map resolution (Å)	2.02
FSC threshold (half map comparison)	0.143

Table 2. Model refinement statistics.

Model component	70S ribosome	30S subunit	50S subunit
Model resolution (Å)	2.06	2.17	1.95
FSC threshold (map-vs.-model)	0.5	0.5	0.5
Map sharpening *B* factor (Å^2)	-33.5	-37.6	-30.6
Model composition			
non-hydrogen atoms	149356	54550	91592
Mg2+ ions	309	93	218
Zn2+ ions	2	0	2
polyamines	17	2	15
waters	7248	2413	4835
ligands (paromomycin)	1	1	0
B factors (Å^2)			
RNA	23.83	28.09	20.9
protein	24.42	28.91	20.95
waters	20.66	17.94	22.02
other	29.29	20.35	33.12
R.m.s. deviations from ideal values			
Bond (Å)	0.006	0.005	0.006
Angle (°)	0.952	0.838	0.997
Molprobity all-atom clash score	7.34	7.12	7.02
Ramachandran plot			
Favored (%)	96	95.66	96.26
Allowed (%)	3.87	4.17	3.65
Outliers (%)	0.13	0.17	0.1
RNA validation			
Angles outliers (%)	0.02	0.009	0.02
Sugar pucker outliers (%)	0.46	0.39	0.39
Average suiteness	0.579	0.586	0.583

Supplementary Information

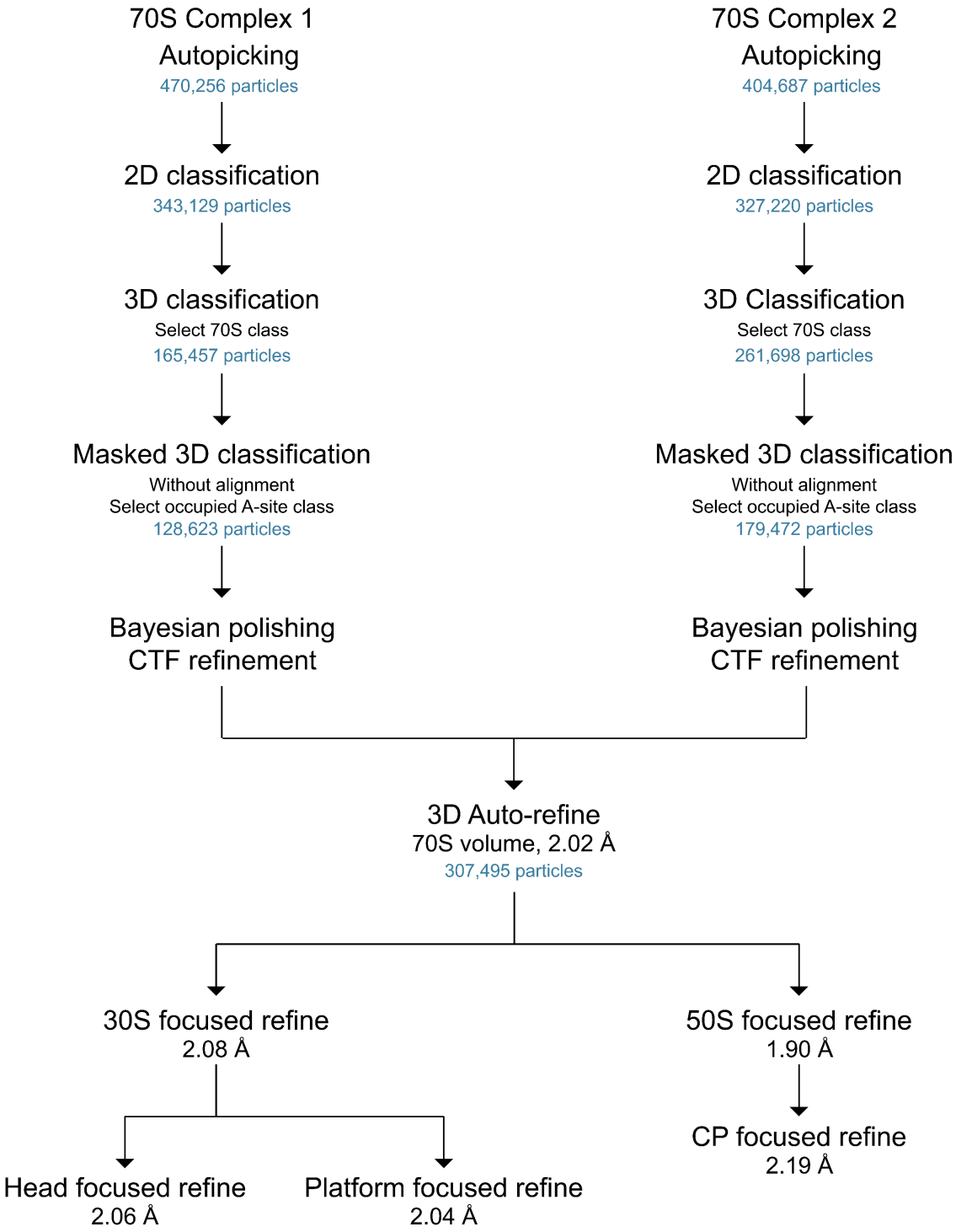


Figure 1–figure supplement 1. General scheme of cryo-EM data processing workflow. The number of particles at each stage are shown, as well as the final resolutions as defined by half-map FSCs (cutoff of 0.143).

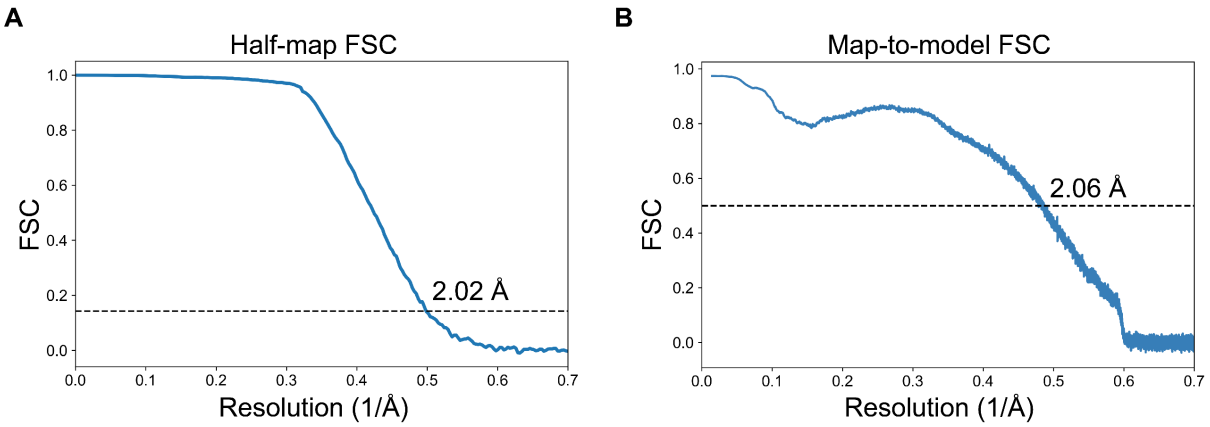


Figure 1–figure supplement 2. Fourier Shell Correlations for cryo-EM maps of the 70S ribosome. (**A**) Half-map FSC curve for the 70S ribosome is shown. The “gold-standard” cutoff value for resolution (0.143) is indicated. (**B**) The map-to-model FSC curve for the 70S ribosome, with an overall resolution indicated at an FSC value of 0.5.

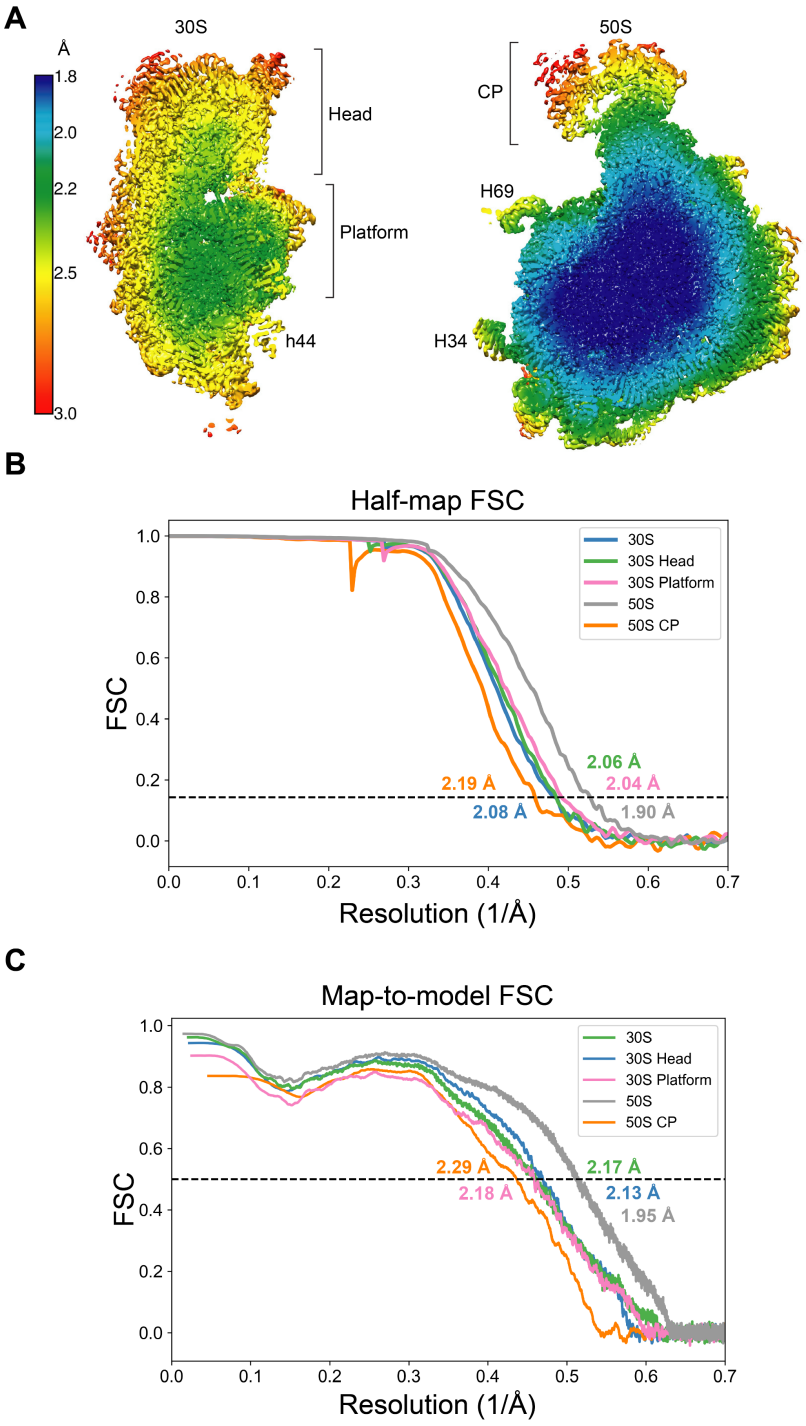


Figure 1–figure supplement 3. Resolution of maps of the 30S and 50S ribosomal subunits. (**A**) Local resolution of the 30S and 50S subunits. Brackets indicate regions masked for further focused refinement. 16S rRNA helix h44 and 23S rRNA helices H34 and H69 are also labeled. (**B**) Half-map FSC curves for focused-refined maps of the 50S subunit, 30S subunit, 50S Central Protuberance, 30S subunit head domain, and 30S subunit platform are shown. The "gold-standard" cutoff value for resolution (0.143) is indicated in each graph. (**C**) Map-to-model FSC

curves for the focused-refined maps of the 50S subunit, 30S subunit, 50S Central Protuberance, 30S subunit head domain, and 30S subunit platform are shown.

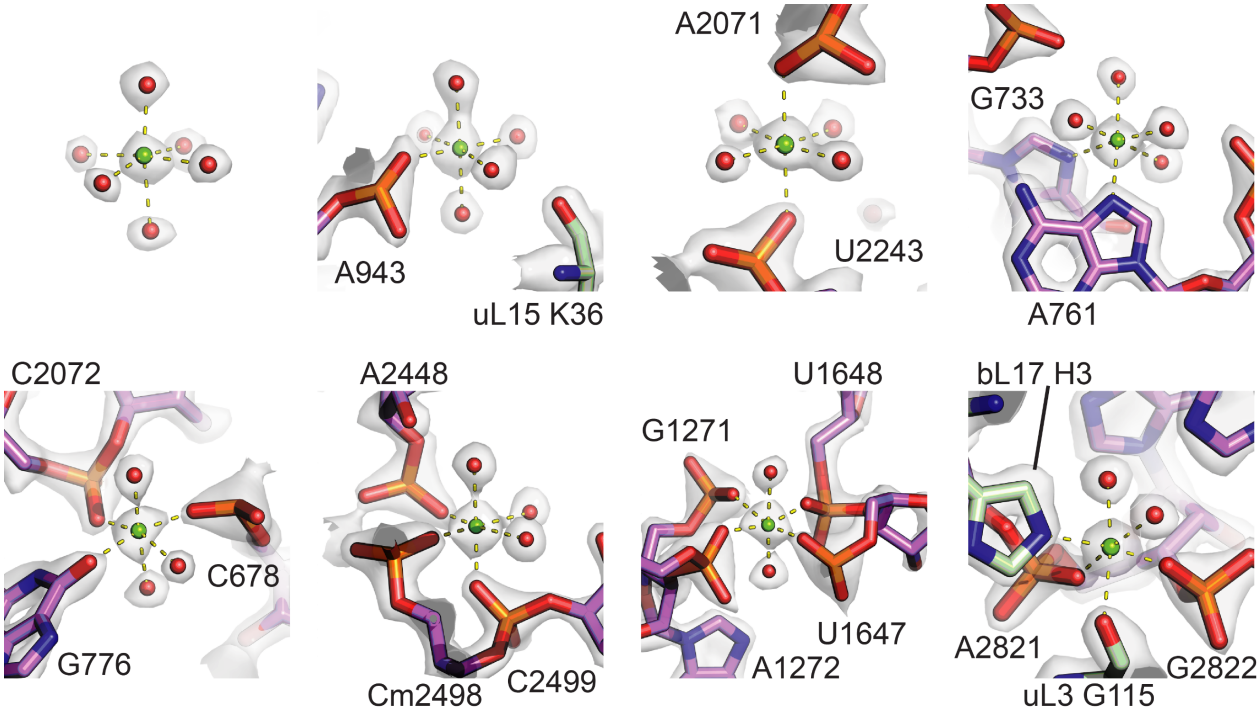


Figure 1–figure supplement 4. Gallery of Mg^{2+} coordination states observed in the 50S subunit. Examples include magnesium ions with different levels of direct coordination to the rRNA. The fully-hydrated Mg^{2+} is located near 23S rRNA nucleotide A973.

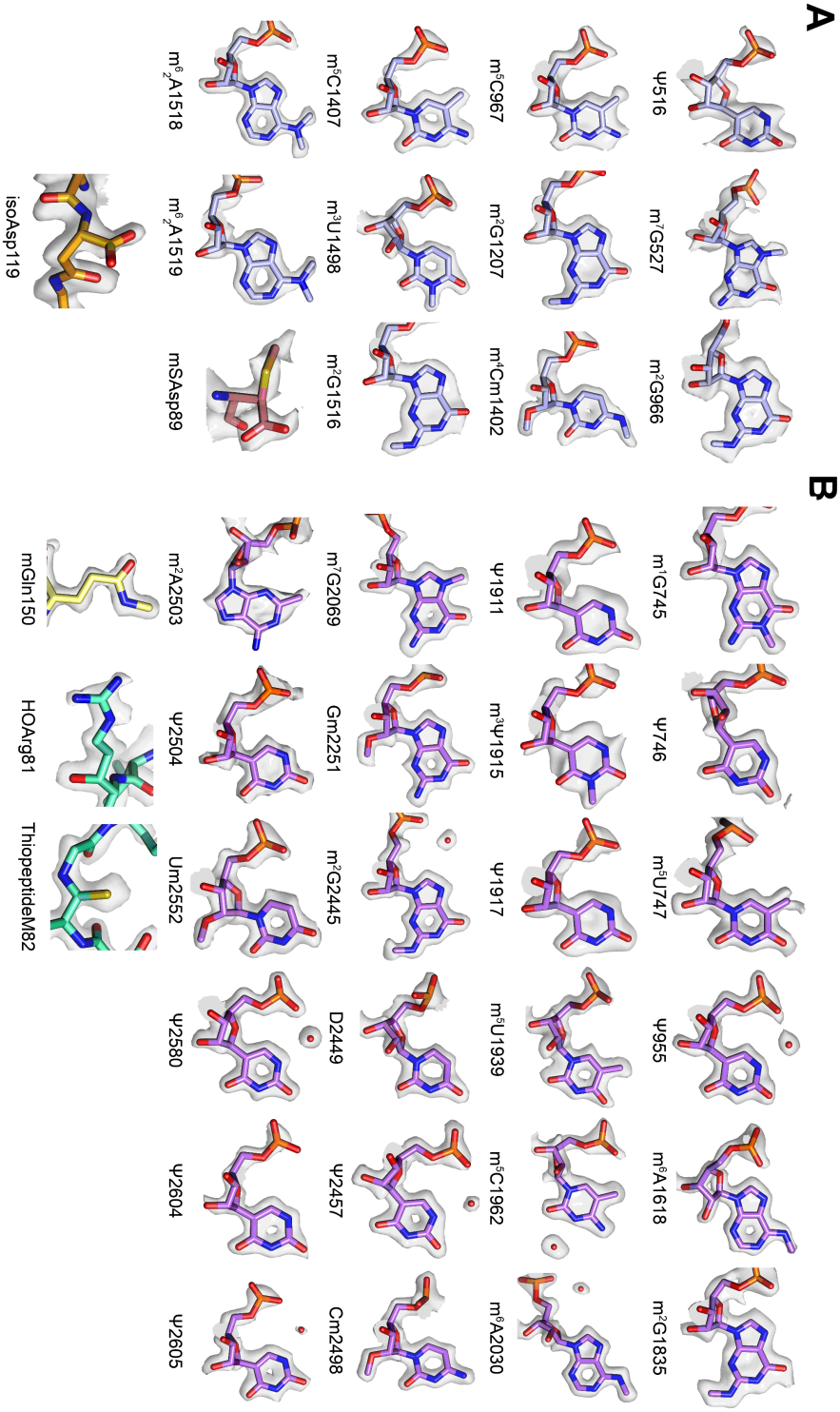


Figure 1–figure supplement 5. Gallery of post-transcriptionally modified nucleotides and post-translationally modified amino acids. (**A**) Post-transcriptionally modified nucleotides and amino acids in the 30S subunit. (**B**) Post-transcriptionally modified nucleotides and amino acids in the 50S subunit. In panel **A**, nucleotides m^7G527 and $m^6{}_2$A1519 appear hypomodified, when compared to m^7G2069 in 23S rRNA and $m^6{}_2$A1518 in 16S rRNA, respectively.

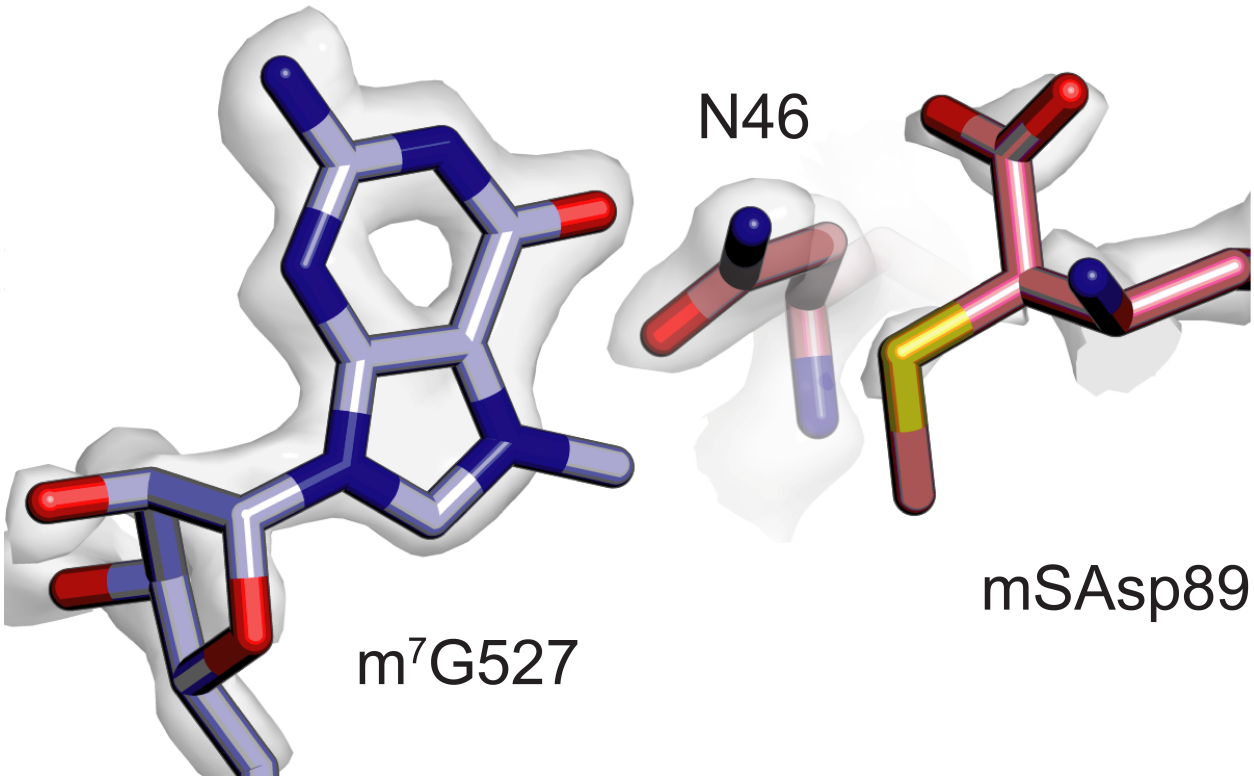


Figure 1–figure supplement 6. Close proximity of m^7G527 in 16S rRNA and β-methylthio-Asp89 in uS12. Both m^7G527 (light purple) and β-methylthio-Asp89 in uS12 (pink) appear hypomodified based on the cryo-EM map contour level required to enclose adjacent atoms and residues.

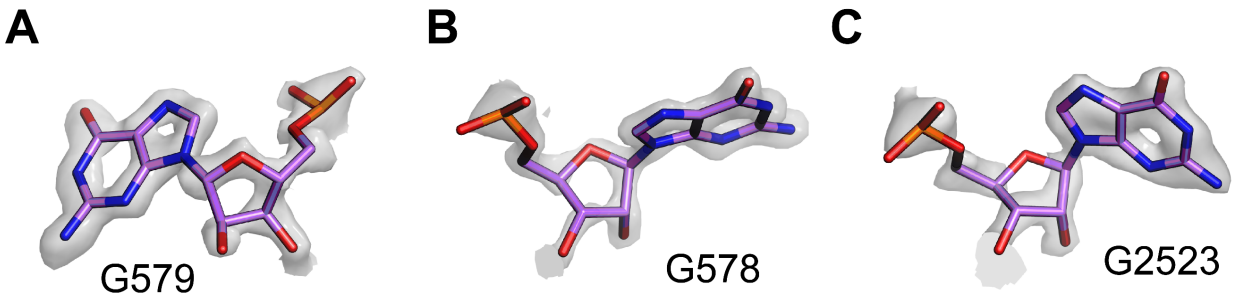


Figure 1–figure supplement 7. **Possible damage to ribose rings at high electron dose visible in cryo-EM maps.** (**A**) Break in the electrostatic potential map between ribose C1' and C2' atoms, in map of the 50S ribosomal subunit. (**B**) Break in the electrostatic potential map between ribose C3' and C4' atoms. (**C**) Break in the electrostatic potential map near the ribose O4' atom. In panel **A**, the focused-refined map of the 50S subunit was sharpened with a B-factor of -15 Å^2. In panel **B**, the map was not sharpened, and in panel **C**, the map was blurred with a B-factor of 5 Å^2.

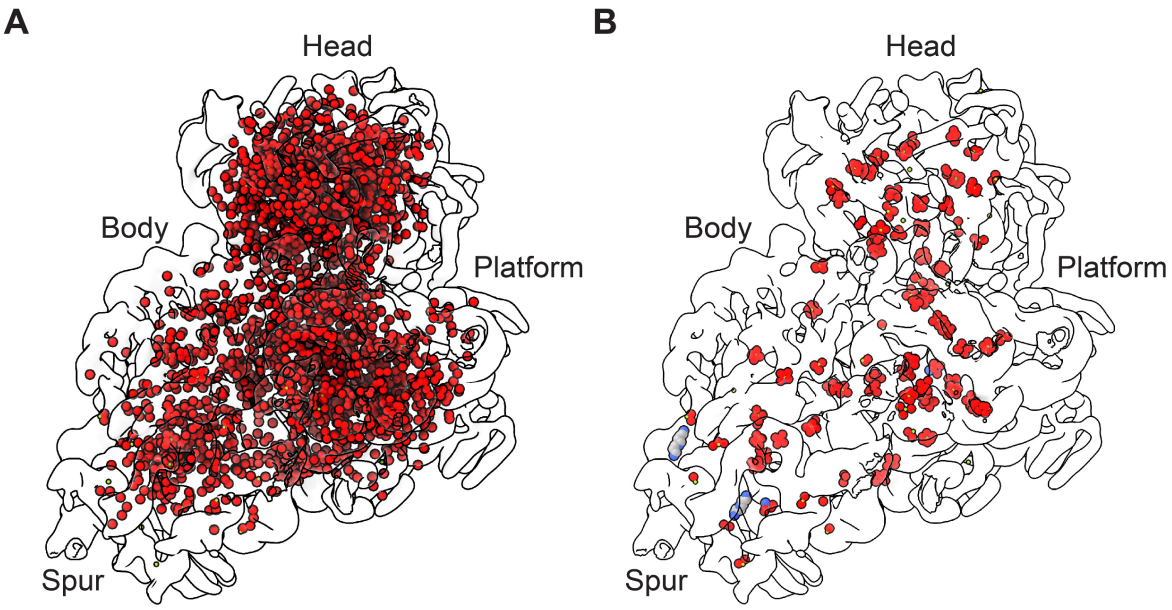


Figure 2–figure supplement 1. Solvation of the 30S ribosomal subunit. (**A**) Solvation of the 30S subunit with solvent oxygen atoms shown as red spheres. (**B**) Polyamine (grey carbon) and Mg^{2+} (green) sites in the 30S subunit, shown with metal-coordinating atoms (oxygen in red, nitrogen in blue). Cryo-EM density is shown with a low-pass filter of 8 Å.

Figure 3–figure supplement 1. Phage S21 and bacterial host bS21 sequence alignments. Sequence alignments of phage S21 homologues (in red) and associated host bacterial clades from (**A**) Bacteroidetes (orange), (**B**) The Candidate Phyla Radiation (magenta), (**C**) Betaproteobacteria (green) and Gammaproteobacteria (blue), (**D**) Spirochetes (bluegreen) and (**E**) Firmicutes (black). Phage sequences in bold indicate phages with a predicted bacterial host.

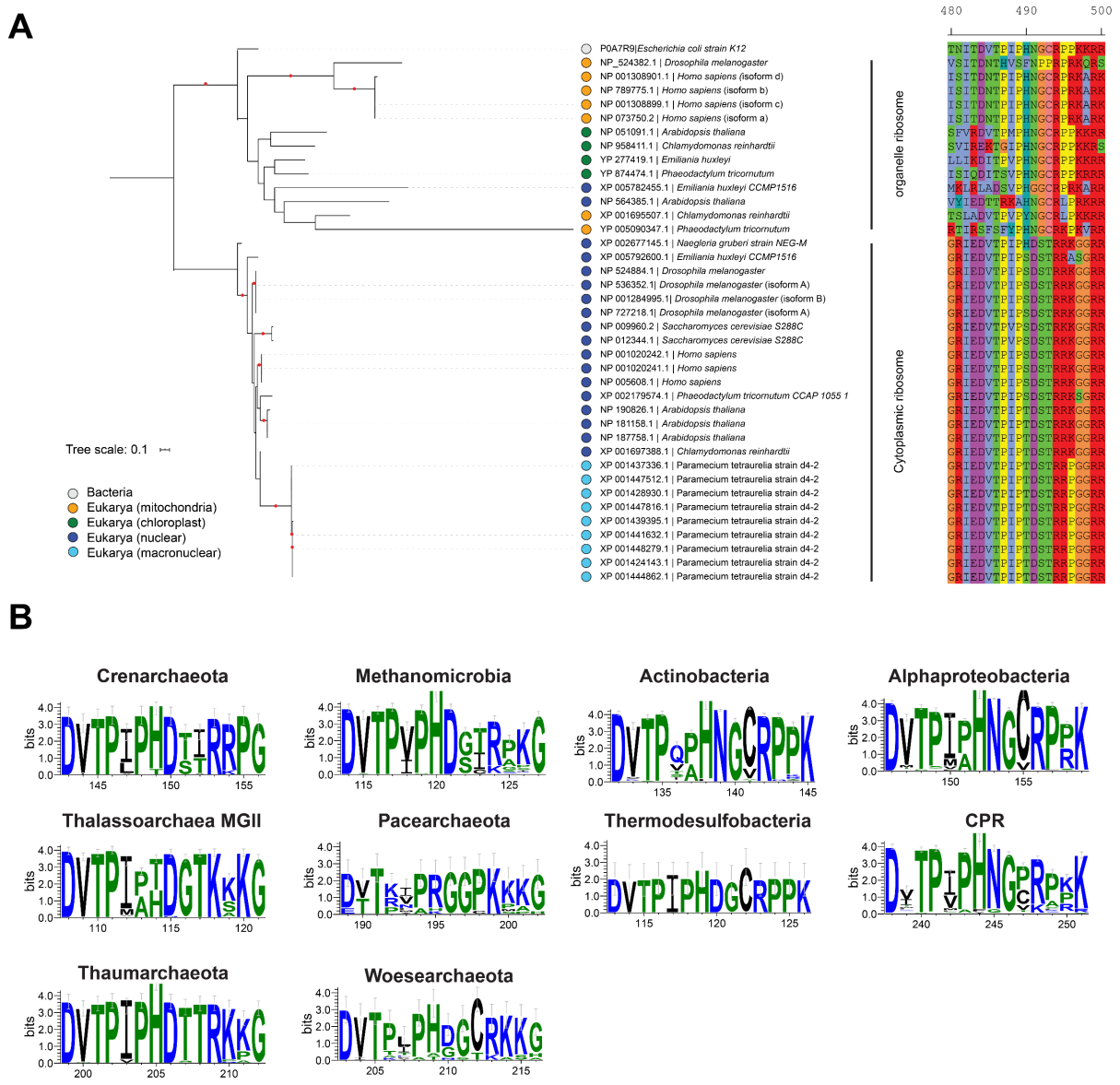


Figure 4–figure supplement 1. Conservation of residues near the isoAsp residue in uS11 homologues. (**A**) Phylogenetic tree of the uS11 ribosomal proteins in Eukaryotes, including both cytoplasmic and organelle examples, along with *Escherichia coli* and the amino acids around the PHNG motif. The maximum likelihood tree was constructed under an LG+G4 model of evolution. Nodes with bootstrap values ≥ 85 are indicated by red circles. Scale bar indicates the average substitutions per site. (**B**) Sequence conservation in uS11 homologues in different bacterial and archaeal phyla.

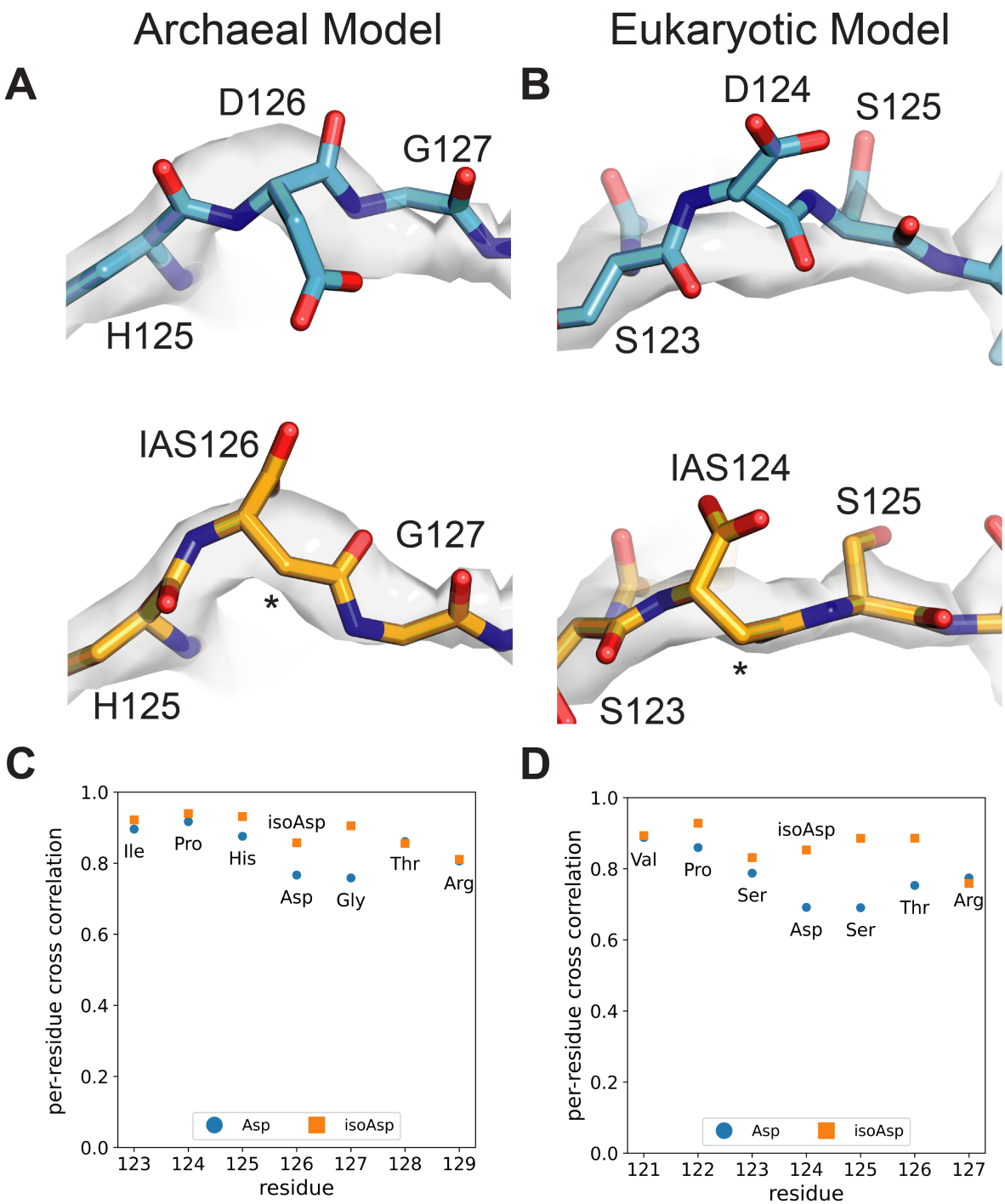


Figure 4–figure supplement 2. Structural models for isoaspartate in archaeal and eukaryotic ribosomes. Comparisons of published uS11 models (blue) in the archaeal (Nürenberg-Goloub et al., 2020) (**A**) and eukaryotic (Tesina et al., 2020) ribosomes (**B**) with models incorporating the IsoAsp modification (yellow-orange), real-space refined into the corresponding published maps. (**C**) Real-space correlations of models refined into the archaeal 30S subunit cryo-EM map, on a per-residue basis. (**D**) Real-space correlations of models refined into the eukaryotic ribosome cryo-EM map.

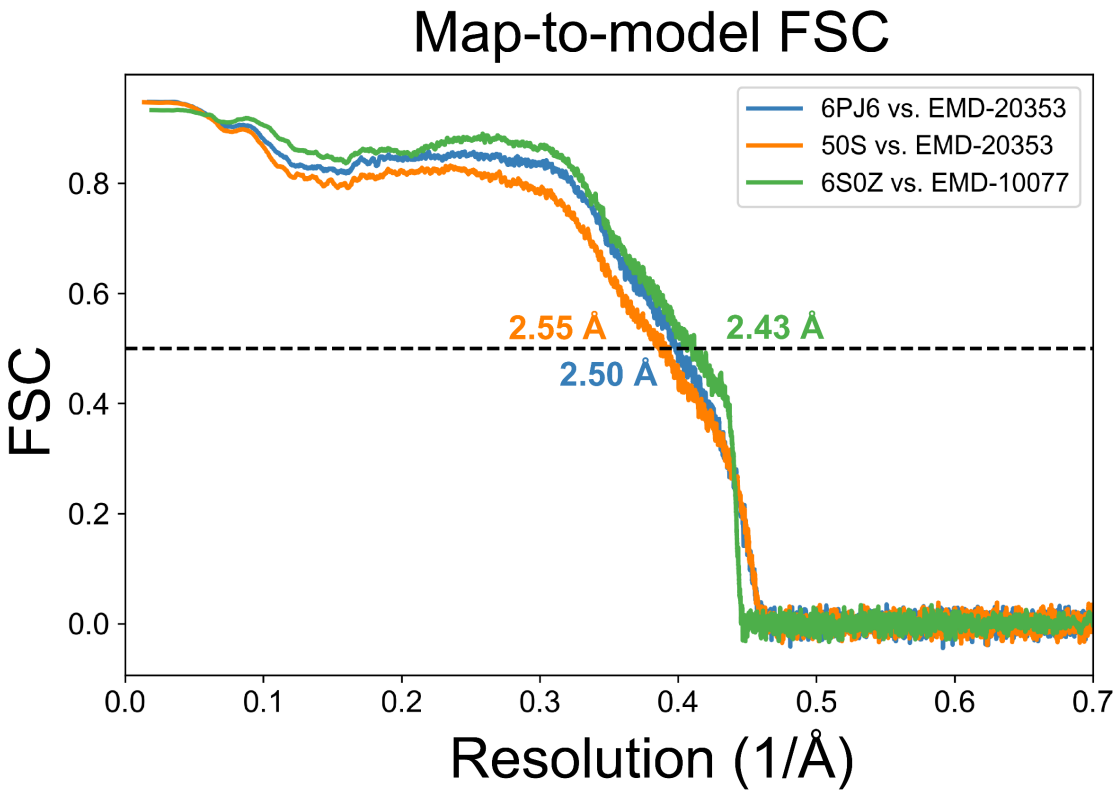


Figure 7–figure supplement 1. Map-to-model resolution estimates for deposited 50S subunit structures. Map-to-model FSC curves for deposited 50S subunit coordinates against the deposited 50S subunit maps, scaled to the present 50S subunit model.

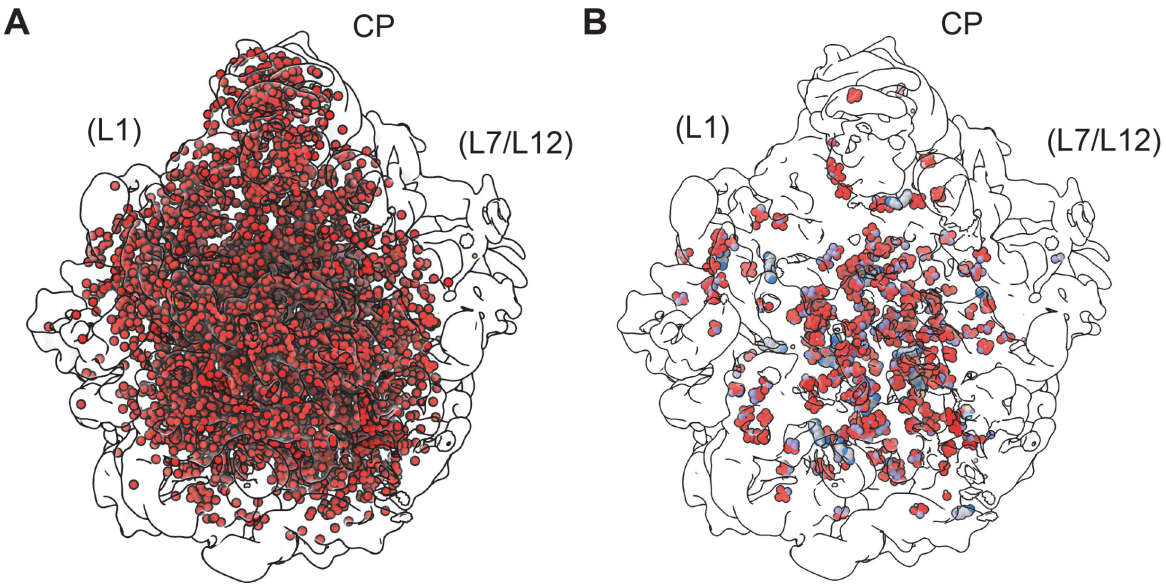


Figure 7–figure supplement 2. Solvation in the 50S ribosomal subunit. (**A**) Waters modeled in the 50S subunit are shown as red spheres (oxygen atoms) in outline of the map. View is from the 30S subunit interface side, with approximate locations of uL1 and bL12 (disordered in the structure) shown. (**B**) Polyamine (grey carbons) and Mg^{2+} (green) sites in the 50S subunit, shown with metal-coordinating atoms (oxygen in red, nitrogen in blue, and carbons in violet, ivory, cyan). Cryo-EM density is shown with a low-pass filter of 8 Å.

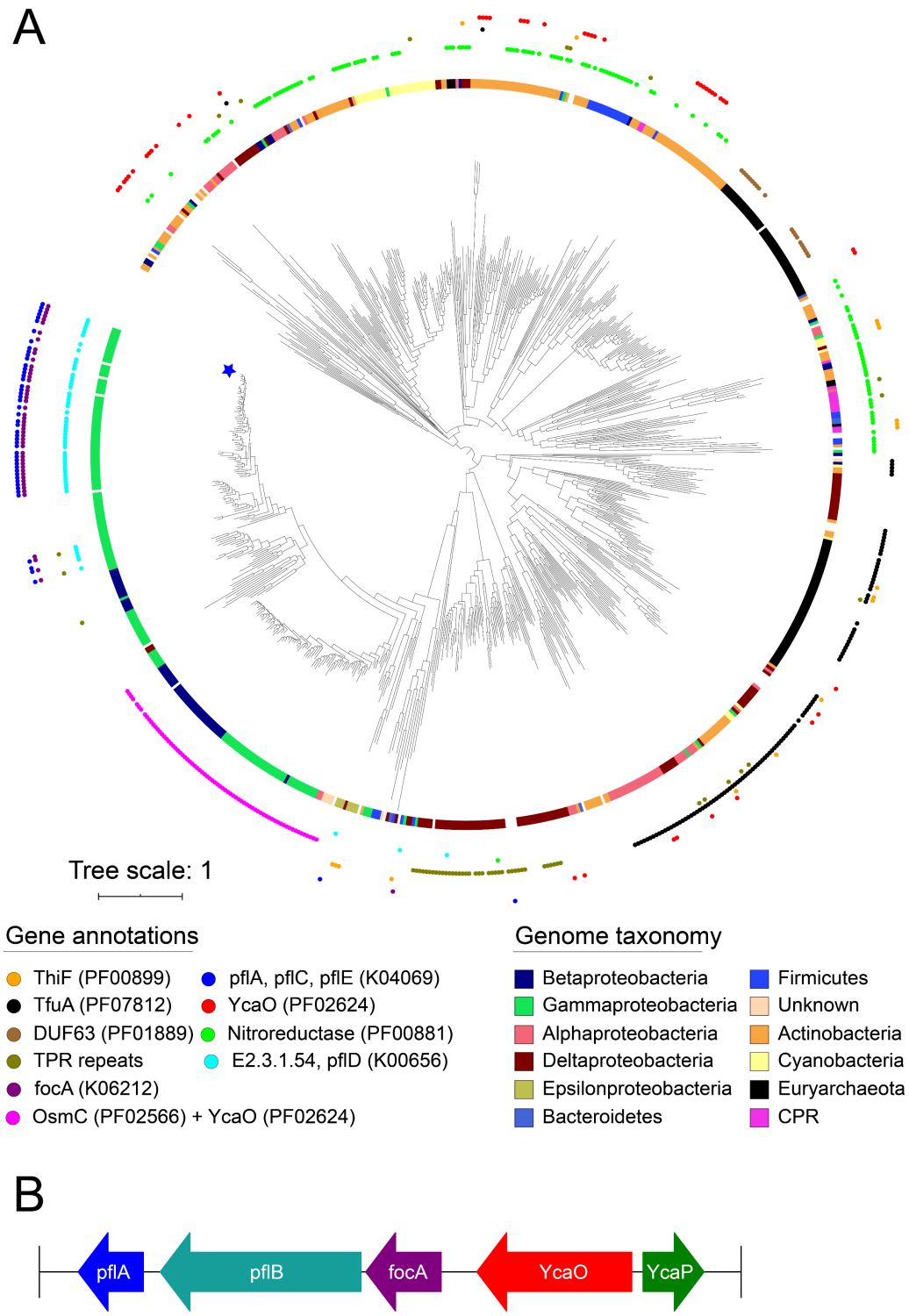


Figure 8–figure supplement 1. Phylogeny of YcaO family members. (**A**) Phylogenetic tree of YcaO family members. The maximum likelihood tree was constructed under an LG+F+G4 model of evolution. Scale bar indicates the average substitutions per site. The inner circle shows organismal distribution, while the outer circle shows neighboring gene patterns. Star indicates

location of *Escherichia coli* str K-12 substr MG1655. (**B**) Genes neighboring YcaO in *E. coli* K-12 substr MG1655.

Supplementary Table 1. Phages encoding S21 homologs. Tabs include phages encoding S21 homologs with predicted bacterial hosts, along with ribosome binding sites for the phages, Betaproteobacteria, Firmicutes, CPR bacteria, Spirochaetes, and Bacteroidetes.

Supplementary Table 2. Phylogenetic analysis of rRNA contacts near the uS11 isoAsp residue. Tabs include 16S base pair statistics for prokaryotes, bacteria, archaea, 16S rRNA genome information for prokaryotes, 18S base pair statistics for eukaryotes, 18S rRNA genome information for eukaryotes, and nucleotide statistics for position 718. All 16S rRNA base pairs and position 718 are with *E. coli* numbering. 18S rRNA base pairs are with *S. cerevisiae* numbering.

References

Acedo JZ, Bothwell IR, An L, Trouth A, Frazier C, van der Donk WA. 2019. O-Methyltransferase-Mediated Incorporation of a β-Amino Acid in Lanthipeptides. *J Am Chem Soc* **141**:16790–16801.

Ad O, Hoffman KS, Cairns AG, Featherston AL, Miller SJ, Söll D, Schepartz A. 2019. Translation of Diverse Aramid- and 1,3-Dicarbonyl-peptides by Wild Type Ribosomes in Vitro. *ACS Cent Sci* **5**:1289–1294.

Al-Shayeb B, Sachdeva R, Chen L-X, Ward F, Munk P, Devoto A, Castelle CJ, Olm MR, Bouma-Gregson K, Amano Y, He C, Méheust R, Brooks B, Thomas A, Lavy A, Matheus-Carnevali P, Sun C, Goltsman DSA, Borton MA, Sharrar A, Jaffe AL, Nelson TC, Kantor R, Keren R, Lane KR, Farag IF, Lei S, Finstad K, Amundson R, Anantharaman K, Zhou J, Probst AJ, Power ME, Tringe SG, Li W-J, Wrighton K, Harrison S, Morowitz M, Relman DA, Doudna JA, Lehours A-C, Warren L, Cate JHD, Santini JM, Banfield JF. 2020. Clades of huge phages from across Earth's ecosystems. *Nature* **578**:425–431.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W. 1997a. Lipman. DJ 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* **25**:3389.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997b. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**:3389–3402.

Anantharaman K, Brown CT, Hug LA, Sharon I, Castelle CJ, Probst AJ, Thomas BC, Singh A, Wilkins MJ, Karaoz U, Brodie EL, Williams KH, Hubbard SS, Banfield JF. 2016. Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nature Communications*. doi:10.1038/ncomms13219

Anton BP, Saleh L, Benner JS, Raleigh EA, Kasif S, Roberts RJ. 2008. RimO, a MiaB-like enzyme, methylthiolates the universally conserved Asp88 residue of ribosomal protein S12 in *Escherichia coli*. *Proc Natl Acad Sci U S A* **105**:1826–1831.

Arenz S, Wilson DN. 2016. Bacterial Protein Synthesis as a Target for Antibiotic Inhibition. *Cold Spring Harb Perspect Med* **6**. doi:10.1101/cshperspect.a025361

Arora S, Bhamidimarri SP, Weber MHW, Varshney U. 2013. Role of the ribosomal P-site elements of m^2G966, m^5C967, and the S9 C-terminal tail in maintenance of the reading frame during translational elongation in *Escherichia coli*. *J Bacteriol* **195**:3524–3530.

Benítez-Páez A, Cárdenas-Brito S, Corredor M, Villarroya M, Armengod ME. 2014. Impairing methylations at ribosome RNA, a point mutation-dependent strategy for aminoglycoside resistance: the rsmG case. *Biomedica* **34 Suppl 1**:41–49.

Brosius J, Chen R. 1976. The primary structure of protein L16 located at the peptidyltransferase center of *Escherichia coli* ribosomes. *FEBS Lett* **68**:105–109.

Brown CT, Hug LA, Thomas BC, Sharon I, Castelle CJ, Singh A, Wilkins MJ, Wrighton KC, Williams KH, Banfield JF. 2015. Unusual biology across a group comprising more than 15% of domain Bacteria. *Nature* **523**:208–211.

Bubunenko M, Baker T, Court DL. 2007. Essentiality of ribosomal and transcription antitermination proteins analyzed by systematic gene replacement in *Escherichia coli*. *J Bacteriol* **189**:2844–2853.

Burkhart BJ, Schwalen CJ, Mann G, Naismith JH, Mitchell DA. 2017. YcaO-Dependent Posttranslational Amide Activation: Biosynthesis, Structure, and Function. *Chem Rev* **117**:5389–5456.

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**:1972–1973.

Casañal A, Lohkamp B, Emsley P. 2020. Current developments in Coot for macromolecular model building of Electron Cryo-microscopy and Crystallographic Data. *Protein Sci* **29**:1069–1078.

Cheng A, Eng ET, Alink L, Rice WJ, Jordan KD, Kim LY, Potter CS, Carragher B. 2018. High resolution single particle cryo-electron microscopy using beam-image shift. *J Struct Biol* **204**:270–275.

Connolly K, Rife JP, Culver G. 2008. Mechanistic insight into the ribosome biogenesis functions of the ancient protein KsgA. *Mol Microbiol* **70**:1062–1075.

Crooks GE, Hon G, Chandonia J-M, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Res* **14**:1188–1190.

Dai Y, Shortreed MR, Scalf M, Frey BL, Cesnik AJ, Solntsev S, Schaffer LV, Smith LM. 2017. Elucidating Escherichia coli Proteoform Families Using Intact-Mass Proteomics and a Global PTM Discovery Database. *J Proteome Res* **16**:4156–4165.

David CL, Keener J, Aswad DW. 1999. Isoaspartate in ribosomal protein S11 of *Escherichia coli*. *J Bacteriol* **181**:2872–2877.

Deutsch EW, Mendoza L, Shteynberg D, Slagel J, Sun Z, Moritz RL. 2015. Trans-Proteomic Pipeline, a standardized data processing pipeline for large-scale reproducible proteomics informatics. *Proteomics Clin Appl* **9**:745–754.

DiMaio F, Zhang J, Chiu W, Baker D. 2013. Cryo-EM model validation using independent map reconstructions. *Protein Sci* **22**:865–868.

Dunkle JA, Wang L, Feldman MB, Pulk A, Chen VB, Kapral GJ, Noeske J, Richardson JS, Blanchard SC, Cate JHD. 2011. Structures of the bacterial ribosome in classical and hybrid states of tRNA binding. *Science* **332**:981–984.

Eddy SR. 1998. Profile hidden Markov models. *Bioinformatics* **14**:755–763.

Fischer N, Neumann P, Konevega AL, Bock LV, Ficner R, Rodnina MV, Stark H. 2015. Structure of the E. coli ribosome-EF-Tu complex at <3 Å resolution by Cs-corrected cryo-EM. *Nature* **520**:567–570.

Frank J. 2017. Advances in the field of single-particle cryo-electron microscopy over the last decade. *Nat Protoc* **12**:209–212.

Frank J, Al-Ali L. 1975. Signal-to-noise ratio of electron micrographs obtained by cross correlation. *Nature* **256**:376–379.

Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**:3150–3152.

Ge W, Wolf A, Feng T, Ho C-H, Sekirnik R, Zayer A, Granatino N, Cockman ME, Loenarz C, Loik ND, Hardy AP, Claridge TDW, Hamed RB, Chowdhury R, Gong L, Robinson CV, Trudgian DC, Jiang M, Mackeen MM, Mccullagh JS, Gordiyenko Y, Thalhammer A, Yamamoto A, Yang M, Liu-Yi P, Zhang Z, Schmidt-Zachmann M, Kessler BM, Ratcliffe PJ, Preston GM, Coleman ML, Schofield CJ. 2012. Oxygenase-catalyzed ribosome hydroxylation occurs in prokaryotes and humans. *Nat Chem Biol* **8**:960–962.

Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, Ferrin TE. 2018. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* **27**:14–25.

Goodall ECA, Robinson A, Johnston IG, Jabbari S, Turner KA, Cunningham AF, Lund PA, Cole JA, Henderson IR. 2018. The Essential Genome of Escherichia coli K-12. *MBio* **9**. doi:10.1128/mBio.02096-17

Goto Y, Katoh T, Suga H. 2011. Flexizymes for genetic code reprogramming. *Nat Protoc* **6**:779–790.

Halfon Y, Matzov D, Eyal Z, Bashan A, Zimmerman E, Kjeldgaard J, Ingmer H, Yonath A. 2019. Exit tunnel modulation as resistance mechanism of *S. aureus* erythromycin resistant mutant. *Sci Rep* **9**:11460.

Harauz G, van Heel M. 1986. Exact filters for general geometry three dimensional reconstruction. *Optik* **73**:146–156.

Held WA, Mizushima S, Nomura M. 1973. Reconstitution of *Escherichia coli* 30 S ribosomal subunits from purified molecular components. *J Biol Chem* **248**:5720–5730.

Henderson R, Sali A, Baker ML, Carragher B, Devkota B, Downing KH, Egelman EH, Feng Z, Frank J, Grigorieff N, Jiang W, Ludtke SJ, Medalia O, Penczek PA, Rosenthal PB, Rossmann MG, Schmid MF, Schröder GF, Steven AC, Stokes DL, Westbrook JD, Wriggers W, Yang H, Young J, Berman HM, Chiu W, Kleywegt GJ, Lawson CL. 2012. Outcome of the first electron microscopy validation task force meeting. *Structure* **20**:205–214.

Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*. doi:10.1093/molbev/msx281

Hobbie SN, Pfister P, Bruell C, Sander P, François B, Westhof E, Böttger EC. 2006. Binding of neomycin-class aminoglycoside antibiotics to mutant ribosomes with alterations in the A site of 16S rRNA. *Antimicrob Agents Chemother* **50**:1489–1496.

Iudin A, Korir PK, Salavert-Torres J, Kleywegt GJ, Patwardhan A. 2016. EMPIAR: a public archive for raw electron microscopy image data. *Nat Methods* **13**:387–388.

Javed A, Orlova EV. 2019. Unravelling Ribosome Function Through Structural Studies. *Subcell Biochem* **93**:53–81.

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* **14**:587–589.

Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2016. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* **44**:D457–62.

Katoh K, Standley DM. 2016. A simple method to control over-alignment in the MAFFT multiple sequence alignment program. *Bioinformatics* **32**:1933–1942.

Komoda T, Sato NS, Phelps SS, Namba N, Joseph S, Suzuki T. 2006. The A-site finger in 23 S rRNA acts as a functional attenuator for translocation. *J Biol Chem* **281**:32303–32309.

Kong AT, Leprevost FV, Avtonomov DM, Mellacheruvu D, Nesvizhskii AI. 2017. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat Methods* **14**:513–520.

Kowalak JA, Walsh KA. 1996. Beta-methylthio-aspartic acid: identification of a novel posttranslational modification in ribosomal protein S12 from *Escherichia coli*. *Protein Sci* **5**:1625–1632.

Kurata S, Weixlbaumer A, Ohtsuki T, Shimazaki T, Wada T, Kirino Y, Takai K, Watanabe K, Ramakrishnan V, Suzuki T. 2008. Modified uridines with C5-methylene substituents at the first position of the tRNA anticodon stabilize U.G wobble pairing during decoding. *J Biol Chem* **283**:18801–18811.

Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* **47**:W256–W259.

Liebschner D, Afonine PV, Baker ML, Bunkóczi G, Chen VB, Croll TI, Hintze B, Hung LW, Jain S, McCoy AJ, Moriarty NW, Oeffner RD, Poon BK, Prisant MG, Read RJ, Richardson JS, Richardson DC, Sammito MD, Sobolev OV, Stockwell DH, Terwilliger TC, Urzhumtsev AG, Videau LL, Williams CJ, Adams PD. 2019. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75**:861–877.

Lill R, Robertson JM, Wintermeyer W. 1986. Affinities of tRNA binding sites of ribosomes from *Escherichia coli*. *Biochemistry* **25**:3245–3255.

Mahanta N, Szantai-Kis DM, Petersson EJ, Mitchell DA. 2019. Biosynthesis and Chemical Applications of Thioamides. *ACS Chem Biol* **14**:142–163.

Ma J, Denisov SA, Adhikary A, Mostafavi M. 2019. Ultrafast Processes Occurring in Radiolysis of Highly Concentrated Solutions of Nucleosides/Tides. *Int J Mol Sci* **20**. doi:10.3390/ijms20194963

Melnikov S, Ben-Shem A, Garreau de Loubresse N, Jenner L, Yusupova G, Yusupov M. 2012. One core, two shells: bacterial and eukaryotic ribosomes. *Nat Struct Mol Biol* **19**:560–567.

Mikheil DM, Shippy DC, Eakley NM, Okwumabua OE, Fadl AA. 2012. Deletion of gene encoding methyltransferase (gidB) confers high-level antimicrobial resistance in Salmonella. *J Antibiot* **65**:185–192.

Munro JB, Sanbonmatsu KY, Spahn CMT, Blanchard SC. 2009. Navigating the ribosome's metastable energy landscape. *Trends Biochem Sci* **34**:390–400.

Nakane T, Kimanius D, Lindahl E, Scheres SH. 2018. Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. *Elife* **7**. doi:10.7554/eLife.36861

Nakane T, Kotecha A, Sente A, McMullan G, Masiulis S, Patricia M G, Grigoras IT, Malinauskaite L, Malinauskas T, Miehling J, Yu L, Karia D, Pechnikova EV, de Jong E, Keizer J, Bischoff M, McCormack J, Tiemeijer P, Hardwick SW, Chirgadze DY, Murshudov G, Radu Aricescu A, Scheres SHW. 2020. Single-particle cryo-EM at atomic resolution. *bioRxiv*. doi:10.1101/2020.05.22.110189

Nawrocki EP, Kolbe DL, Eddy SR. 2009. Infernal 1.0: inference of RNA alignments. *Bioinformatics* **25**:1335–1337.

Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* **32**:268–274.

Nichols RJ, Sen S, Choo YJ, Beltrao P, Zietek M, Chaba R, Lee S, Kazmierczak KM, Lee KJ, Wong A, Shales M, Lovett S, Winkler ME, Krogan NJ, Typas A, Gross CA. 2011. Phenotypic landscape of a bacterial cell. *Cell* **144**:143–156.

Noeske J, Wasserman MR, Terry DS, Altman RB, Blanchard SC, Cate JHD. 2015. High-resolution structure of the *Escherichia coli* ribosome. *Nat Struct Mol Biol* **22**:336–341.

Nürenberg-Goloub E, Kratzat H, Heinemann H, Heuer A, Kötter P, Berninghausen O, Becker T, Tampé R, Beckmann R. 2020. Molecular analysis of the ribosome recycling factor ABCE1 bound to the 30S post-splitting complex. *EMBO J* **39**:e103788.

Ochi K, Kim J-Y, Tanaka Y, Wang G, Masuda K, Nanamiya H, Okamoto S, Tokuyama S, Adachi Y, Kawamura F. 2009. Inactivation of KsgA, a 16S rRNA methyltransferase, causes vigorous emergence of mutants with high-level kasugamycin resistance. *Antimicrob Agents Chemother* **53**:193–201.

O'Farrell HC, Musayev FN, Scarsdale JN, Rife JP. 2012. Control of substrate specificity by a single active site residue of the KsgA methyltransferase. *Biochemistry* **51**:466–474.

O'Farrell HC, Rife JP. 2012. Staphylococcus aureus and Escherichia coli have disparate dependences on KsgA for growth and ribosome biogenesis. *BMC Microbiol* **12**:244.

Ogle JM, Murphy FV, Tarry MJ, Ramakrishnan V. 2002. Selection of tRNA by the ribosome requires a transition from an open to a closed form. *Cell* **111**:721–732.

Okamoto S, Tamaru A, Nakajima C, Nishimura K, Tanaka Y, Tokuyama S, Suzuki Y, Ochi K. 2007. Loss of a conserved 7-methylguanosine modification in 16S rRNA confers low-level streptomycin resistance in bacteria. *Mol Microbiol* **63**:1096–1106.

Olm MR, Brown CT, Brooks B, Banfield JF. 2017. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J* **11**:2864–2868.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**:1605–1612.

Phunpruch S, Warit S, Suksamran R, Billamas P, Jaitrong S, Palittapongarnpim P, Prammananan T. 2013. A role for 16S rRNA dimethyltransferase (ksgA) in intrinsic clarithromycin resistance in *Mycobacterium tuberculosis*. *Int J Antimicrob Agents* **41**:548–551.

Pichkur EB, Paleskava A, Tereshchenkov AG, Kasatsky P, Komarova ES, Shiriaev DI, Bogdanov AA, Dontsova OA, Osterman IA, Sergiev PV, Polikanov YS, Myasnikov AG, Konevega AL. 2020. Insights into the improved macrolide inhibitory activity from the high-resolution cryo-EM structure of dirithromycin bound to the *E. coli* 70S ribosome. *RNA* **26**:715–723.

Pino LK, Searle BC, Bollinger JG, Nunn B, MacLean B, MacCoss MJ. 2020. The Skyline ecosystem: Informatics for quantitative mass spectrometry proteomics. *Mass Spectrom Rev* **39**:229–244.

Polikanov YS, Melnikov SV, Söll D, Steitz TA. 2015. Structural insights into the role of rRNA modifications in protein synthesis and ribosome assembly. *Nat Struct Mol Biol* **22**:342–344.

Polikanov YS, Steitz TA, Innis CA. 2014. A proton wire to couple aminoacyl-tRNA accommodation and peptide-bond formation on the ribosome. *Nat Struct Mol Biol* **21**:787–793.

Punjani A, Fleet DJ. 2020. 3D Variability Analysis: Directly resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM images. *bioRxiv*. doi:10.1101/2020.04.08.032466

Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**:290–296.

Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer ELL, Eddy SR, Bateman A, Finn RD. 2012. The Pfam protein families database. *Nucleic Acids Res* **40**:D290–301.

Reissner KJ, Aswad DW. 2003. Deamidation and isoaspartate formation in proteins: unwanted alterations or surreptitious signals? *Cell Mol Life Sci* **60**:1281–1295.

Robinson NE, Robinson AB. 2001. Molecular clocks. *Proc Natl Acad Sci U S A* **98**:944–949.

Rodnina MV. 2013. The ribosome as a versatile catalyst: reactions at the peptidyl transferase center. *Curr Opin Struct Biol* **23**:595–602.

Rodnina MV, Fischer N, Maracci C, Stark H. 2017. Ribosome dynamics during decoding. *Philos Trans R Soc Lond B Biol Sci* **372**. doi:10.1098/rstb.2016.0182

Rohou A, Grigorieff N. 2015. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* **192**:216–221.

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.

Sashital DG, Greeman CA, Lyumkis D, Potter CS, Carragher B, Williamson JR. 2014. A combined quantitative mass spectrometry and electron microscopy analysis of ribosomal 30S subunit assembly in E. coli. *Elife* **3**. doi:10.7554/eLife.04491

Sati GC, Sarpe VA, Furukawa T, Mondal S, Mantovani M, Hobbie SN, Vasella A, Böttger EC, Crich D. 2019. Modification at the 2'-Position of the 4,5-Series of 2-Deoxystreptamine Aminoglycoside Antibiotics To Resist Aminoglycoside Modifying Enzymes and Increase Ribosomal Target Selectivity. *ACS Infect Dis* **5**:1718–1730.

Sati GC, Shcherbakov D, Hobbie SN, Vasella A, Böttger EC, Crich D. 2017. N6', N6"', and O4' Modifications to Neomycin Affect Ribosomal Selectivity without Compromising Antibacterial Activity. *ACS Infect Dis* **3**:368–377.

Sawers G, Suppmann B. 1992. Anaerobic induction of pyruvate formate-lyase gene expression is mediated by the ArcA and FNR proteins. *J Bacteriol* **174**:3474–3478.

Sawers RG. 2005. Evidence for novel processing of the anaerobically inducible dicistronic focA-pfl mRNA transcript in *Escherichia coli*. *Mol Microbiol* **58**:1441–1453.

Scheres SHW. 2012. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* **180**:519–530.

Schmeing TM, Huang KS, Kitchen DE, Strobel SA, Steitz TA. 2005. Structural insights into the roles of water and the 2' hydroxyl of the P site tRNA in the peptidyl transferase reaction. *Mol Cell* **20**:437–448.

Schorb M, Haberbosch I, Hagen WJH, Schwab Y, Mastronarde DN. 2019. Software tools for automated transmission electron microscopy. *Nat Methods* **16**:471–477.

Selmer M, Dunham CM, Murphy FV 4th, Weixlbaumer A, Petry S, Kelley AC, Weir JR, Ramakrishnan V. 2006. Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313**:1935–1942.

Sharma H, Anand B. 2019. Ribosome assembly defects subvert initiation Factor3 mediated scrutiny of bona fide start signal. *Nucleic Acids Res* **47**:11368–11386.

Shine J, Dalgarno L. 1975. Determinant of cistron specificity in bacterial ribosomes. *Nature* **254**:34–38.

Stephenson RC, Clarke S. 1989. Succinimide formation from aspartyl and asparaginyl peptides as a model for the spontaneous degradation of proteins. *J Biol Chem* **264**:6164–6170.

Stern S, Powers T, Changchien LM, Noller HF. 1988. Interaction of ribosomal proteins S5, S6, S11, S12, S18 and S21 with 16 S rRNA. *J Mol Biol* **201**:683–695.

Stojković V, Myasnikov AG, Young ID, Frost A, Fraser JS, Fujimori DG. 2020. Assessment of the nucleotide modifications in the high-resolution cryo-electron microscopy structure of the *Escherichia coli* 50S subunit. *Nucleic Acids Res* **48**:2723–2732.

Strader MB, Costantino N, Elkins CA, Chen CY, Patel I, Makusky AJ, Choy JS, Court DL, Markey SP, Kowalak JA. 2011. A proteomic and transcriptomic approach reveals new insight

into beta-methylthiolation of *Escherichia coli* ribosomal protein S12. *Mol Cell Proteomics* **10**:M110.005199.

Subramaniam S, Earl LA, Falconieri V, Milne JL, Egelman EH. 2016. Resolution advances in cryo-EM enable application to drug discovery. *Curr Opin Struct Biol* **41**:194–202.

Terwilliger TC, Ludtke SJ, Read RJ, Adams PD, Afonine PV. 2020a. Improvement of cryo-EM maps by density modification. *bioRxiv*. doi:10.1101/845032

Terwilliger TC, Sobolev OV, Afonine PV, Adams PD, Read RJ. 2020b. Density modification of cryo-EM maps. *bioRxiv*. doi:10.1101/2020.05.13.091884

Tesina P, Lessen LN, Buschauer R, Cheng J, Wu CC-C, Berninghausen O, Buskirk AR, Becker T, Beckmann R, Green R. 2020. Molecular mechanism of translational stalling by inhibitory codon combinations and poly(A) tracts. *EMBO J* **39**:e103365.

Travin DY, Watson ZL, Metelev M, Ward FR, Osterman IA, Khven IM, Khabibullina NF, Serebryakova M, Mergaert P, Polikanov YS, Cate JHD, Severinov K. 2019. Structure of ribosome-bound azole-modified peptide phazolicin rationalizes its species-specific mode of bacterial translation inhibition. *Nat Commun* **10**:4563.

Vicens Q, Westhof E. 2001. Crystal structure of paromomycin docked into the eubacterial ribosomal decoding A site. *Structure* **9**:647–658.

Wang L, Pulk A, Wasserman MR, Feldman MB, Altman RB, Cate JHD, Blanchard SC. 2012. Allosteric control of the ribosome by small-molecule antibiotics. *Nat Struct Mol Biol* **19**:957–963.

Wasserman MR, Pulk A, Zhou Z, Altman RB, Zinder JC, Green KD, Garneau-Tsodikova S, Cate JHD, Blanchard SC. 2015. Chemically related 4,5-linked aminoglycoside antibiotics drive subunit rotation in opposite directions. *Nat Commun* **6**:7896.

Watkins AM, Das R. 2019. FARFAR2: Improved de novo Rosetta prediction of complex global RNA folds. *bioRxiv*. doi:10.1101/764449

Yusupova G, Yusupov M. 2017. Crystal structure of eukaryotic ribosome and its complexes with inhibitors. *Philos Trans R Soc Lond B Biol Sci* **372**. doi:10.1098/rstb.2016.0184

Zaher HS, Green R. 2010. Hyperaccurate and error-prone ribosomes exploit distinct mechanisms during tRNA selection. *Mol Cell* **39**:110–120.

Zhao J, Li J, Xu S, Feng P. 2016. Emerging Roles of Protein Deamidation in Innate Immune Signaling. *J Virol* **90**:4262–4268.

Zheng SQ, Palovcak E, Armache J-P, Verba KA, Cheng Y, Agard DA. 2017. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**:331–332.

Zou J, Zhang W, Zhang H, Zhang XD, Peng B, Zheng J. 2018. Studies on Aminoglycoside Susceptibility Identify a Novel Function of KsgA To Secure Translational Fidelity during Antibiotic Stress. *Antimicrob Agents Chemother* **62**. doi:10.1128/AAC.00853-18