Interconnected assembly factors regulate the biogenesis of mitoribosomal large subunit

Victor Tobiasson[1†], Ondřej Gahura[2†], Shintaro Aibara[1,4†], Rozbeh Baradaran[1], Alena Zíková[2,3*], Alexey Amunts[1*]

[1] Science for Life Laboratory, Department of Biochemistry and Biophysics, Stockholm University, 17165 Solna, Sweden.

[2] Institute of Parasitology, Biology Centre, Czech Academy of Sciences, Ceske Budejovice, Czech Republic

[3]Faculty of Science, University of South Bohemia, Ceske Budejovice, Czech Republic.

[4] Current address: Department of Molecular Biology, Max-Planck-Institute for Biophysical Chemistry, Göttingen, Germany.

† The authors contributed equally to this work.

* correspondence to: amunts@scilifelab.se, azikova@paru.cas.cz

Abstract

Mitoribosomes synthesize essential proteins encoded in mitochondria. They consist of RNA and protein components, coordinated assembly of which is critical for function. We used mitoribosomes with reduced RNA and increased protein mass from a mammalian parasite *Trypanosoma brucei* to provide insights into the biogenesis of mitoribosomal large subunit. Structural characterisation of a stable assembly intermediate revealed 16 assembly factors, 10 of which were not previously implicated in biogenesis. Most of the reported factors are also encoded in mammalian genomes. In the structure, the assembly factors form a protein network that spans over 180 Å, shielding the ribosomal RNA surface. The entire central protuberance and L7/L12 stalk are not assembled, and require removal of the factors and remodeling to become functional sites. The conserved mt-RbgA and mt-EngA are bound together in proximity to the peptidyl transferase center. A mitochondrial acyl-carrier protein plays a role in stably docking the L1 stalk that needs to be repositioned during maturation. Together, the extensive network of the assembly factors stabilizes the immature domains and connects between the functionally important regions.

Introduction

Mitoribosomes differ from bacterial and cytosolic ribosomes in their ribosomal RNA (rRNA), protein content, overall size, and structure. Their formation is an intricate and hierarchical process involving multiple proteins and a few RNA molecules working in coordination and under tight regulation (Pearce et al 2017). The cooperative effort involves regulation of two genomes, because rRNA is encoded by the organellar genome, and almost all the mitoribosomal proteins and assembly factors are encoded by the nuclear genome and therefore imported from the cytosol (Couvillion et al 2016). Finally, the fundamental process of the mitoribosomal assembly is complicated due to the localization of its large subunit (mtLSU) to the inner mitochondrial membrane. Therefore, stages of assembly were suggested to involve various mitochondrial *milieu* and specific kinetics (Bogenhagen et al 2014; Antonicka and Shoubridge 2015; De Silva et al 2015). The presence of different compositions is hypothesized to promote formation of defined pre-mitoribosomal complexes with as-yet-unknown organelle-specific auxiliary factors.

Mitochondria of *Trypanosoma brucei* provide a good model for studying the assembly process, because their mitoribosomes consist of over a hundred components, and the ratio of protein to rRNA is unusually high (Zikova et al 2008; Ramrath et al 2018). Since the rRNA forms a compact core of the mitoribosome, and proteins are mostly peripherally associated, an architecture based on the reduced rRNA and supernumerary mitoribosomal proteins would need additional stabilization for its assembly. Therefore, it increases the chances to characterize defined pre-mitoribosomal complexes, which are not stable enough for biochemical isolation in mitochondria of other species. Indeed, structural characterization of an assembly intermediate of the *T. brucei* mitoribosomal small subunit (mtSSU) provided insight into its assembly pathway with many additional prioteins (Saurer et al 2019). However, the assembly of the mtLSU is poorly understood.

The mtLSU accommodates the peptidyl transferase center (PTC) that forms peptide bonds between amino acids, tRNA binding sites, the L7/L12 stalk that is responsible for the recruitment of translation factors, the L1 stalk, the central protuberance (CP) that facilitates communication between various functional sites, and exit tunnel for egress of synthesized protein. In bacteria, our understanding of the LSU assembly is relatively limited (Davis and Williamson 2017). It comes primarily from characterization of the final maturation stages (Li et al 2013; Jomaa et al 2014; Ni et al 2016), studies on incomplete LSU particles as a result of protein depletion (Davis et al 2016), and *in vitro* reconstitution studies with purified ribosomal RNA and protein components (Nikolay et al 2018). These studies identified different LSU precursors with assembly factors bound to rRNA components (Davis and Williamson 2017). In mitochondria, the mtLSU is lacking many of the rRNA components involved in the canonical pathways, and higher complexity of the interactions between the mitoribosomal proteins at the functional sites has evolved (Ott et al 2016; Greber and Ban 2016). A functional mtLSU require that the PTC and the rRNA regions which bind tRNAs are folded, a flexible L1 stalk, an extended L7/L12 protrusion, and a CP built from multiple mitochondria-specific proteins. However, only the final stage of the

mtLSU assembly has been visualized (Brown et al 2017; Itoh et al 2020), and no preceding steps in the formation of the functional sites have been characterised.

To provide insight in the process of the mtLSU assembly, we determined the cryo-EM structure of a native *T. brucei* mtLSU assembly intermediate (pre-mtLSU) complexed with 16 factors. Most of the assembly factors have not been previously implicated in mitoribosomal biogenesis. The structural data suggests that the biogenesis relies on an extensive protein network formed by assembly factors that connect premature PTC, L1 and L7/L12 stalks with the CP. A homology search suggests that some of the newly identified assembly factors are also conserved in mitochondria from other species, including mammals, and therefore may represent a general principle for the mitoribosome assembly. Comparison with two bacterial assembly intermediates (Zhang et al 2014; Seffouh et al 2019) further provides insights into the conserved GTPases EngA and RbgA bound at the subunit interface.

Results and Discussion

Structural determination and composition of the native pre-mtLSU complex

We used a *T. brucei* procyclic strain grown in low-glucose medium that maintains translationally active mitochondria. Mitoribosomal complexes were rapidly purified directly from *T. brucei* mitochondria and analyzed by single-particle cryo-EM. During image processing, in addition to intact monosomes, we detected a pool of free subunits. We focused the analysis on this population and through 3D classification isolated a homogeneous subset of pre-mtLSUs that corresponded to ~3.5 % of the particles combined from five data sets.

896,263 particles were picked using Warp (Tegunov and Cramer 2019), and further processed using RELION (Kimanius et al 2016; Zivanov et al 2018). We performed reference-based 3D classification with references generated from a preliminary classification of a screening data set. This resulted in 207,788 particles corresponding to the mtLSU shape but distinct from that of a mature mtLSU of which we found 152,816 particles. Refinement of pose assignments and subsequent classification using fine-angular searches with a solvent mask identified 32,339 pre-mtLSU particles (Appendix Fig S1). To improve the angles further, the particles were subjected to masked auto-refinement. Following the CTF refinement, we obtained a reconstruction of a pre-mtLSU that likely reflects a stable premature complex. This was evidenced by the presence of densities corresponding to conserved ribosomal assembly factors.

The cryo-EM reconstruction was refined to 3.50 Å resolution. This allowed us to build a ~2.2 MDa model and assign 16 assembly factors, as well as additional mitoribosomal proteins, directly from the density (Fig 1, Fig 2). Five distinct features define the overall pre-mtLSU: 1) the rRNA domain V is well resolved and covered by newly identified mitochondria-specific assembly factors; 2) the subunit interface consists almost entirely of newly identified assembly factors and two conserved GTPases; 3) the proteinaceous CP is absent; 4) the L7/L12 stalk proteins are missing, and its rRNA platform is not folded, instead assembly factors occupy similar positions; 5) the L1 stalk is shifted inward ~30 Å and linked to the CP base by assembly

factors. Due to these features, compositional and conformational changes are required for the maturation of the pre-mtLSU. In terms of the mitoribosomal proteins, 18 previously identified proteins are missing from the structure of the pre-mtLSU. Seven of these have bacterial homologs (uL10m, uL12m, uL16m, bL27m, bL31m, bL33m and bL36m) and the rest are mitochondria specific (Fig 1, Appendix Fig S2, Appendix Fig S3). Additionally, we assigned sequences to two previously unidentified mtLSU proteins, which we named mL109 and mL110, and observed the presence of mitoribosomal protein uL14m which was not modeled in the mtLSU (Fig 2).

Following the previously identified mitoribosomal small subunit assembly factors (Saurer et al 2019) , we adopt a similar nomenclature for the mitochondria specific large subunit factors. Therefore, we reference them as *T. brucei* Large subunit Assembly Factor(s) (TbLAF1-11), while proteins with bacterial homologs are referred to as their bacterial names with the prefix "mt-". The identified assembly factors of the mitoribosome include two homologs of bacterial GTPase assembly factors mt-EngA and mt-RbgA, a homolog of the ribosome silencing factor mt-RsfS, a DEAD-box helicase (TbLAF10), two pseudouridinases (TbLAF4 and TbLAF6), a methyltransferase (TbLAF5), two copies of the mitochondrial acyl-carrier protein (mt-ACP), two LYR-motif-containing proteins (TbLAF1, TbLAF11), and other proteins with previously unassigned functions (TbLAF2, 3, 7, 8 and 9). In the model, we included only the parts for which clear secondary structure is apparent, and other regions with only partial density visible were modeled as UNK1-11 (Supplementary file 2).

GTPase mt-RbgA is structurally linked to the mitoribosomal core via specific assembly linkers

We started the structural analysis by searching for similar assembly intermediate architectures in bacterial counterparts. Particles with an absent CP were reported previously in RbgA-depleted *Bacillus subtilis* cells. RbgA was then added *in vitro* and shown to bind to the complex, which identified its role as an assembly factor (Seffouh et al 2019). RbgA is from the Ras GTPase family containing a typically low intrinsic GTPase activity, which is increased in the presence of a mature LSU subunit (Achila et al 2012). It has N-terminal GTPase domain and C-terminal helical domain, which forms a five-helix bundle (Pausch et al 2018). In the pre-mtLSU structure, we found a conserved mt-RbgA homolog. Studies in yeasts reported that deletion of mt-RbgA (yeast Mtg1) results in respiration deficiency (Barrientos et al 2003). In *B. subtilis*, where this assembly factor is essential, the LSU:RbgA complex showed that the N-terminal domain overlaps with rRNA H69 and H71, and the C-terminal helical domain interacts with H62 and H64 (Seffouh et al 2019). In this position, RbgA displaces the P-site and interacts with the surrounding rRNA, including H92 and H93. Therefore, the binding of RbgA requires contacts with rRNA.

In our map of the *T. brucei* pre-mtLSU, the corresponding regions forming the rRNA binding site for mt-RbgA are not observed, and its nucleotide binding site empty (Fig 3b). However, the

comparison of our structure with the *B. subtilis* LSU:RbgA complex (PDB ID 6PPK) shows conserved position of the catalytic site and nearly identical conformation of the factor on the pre-mtLSU complex (Fig EV1). This includes the weak interaction between the mt-RbgA C-terminal domain and the mitoribosomal protein uL14m is conserved (Fig 3).

We found that the conserved position of mt-RbgA in *T. brucei* is maintained through two specialized assembly linkers (Fig 3). The first linker is established between the C-terminal domain and the TbLAF5 N-terminal helix. The latter adopts a crescent shape around the C-terminal domain of mt-RbgA, forming a series of contacts with four out of its five helices (Fig 3). The second linker is provided by TbLAF4 approaching from the mitoribosomal core. It interacts with the mt-RbgA C-terminal domain and contributes a β-strand to a shared β-sheet (Fig EV2). Therefore, RbgA is anchored to the flexible core via two dedicated factors that compensate for the lack of rRNA contacts.

TbLAF4 belongs to a family of site-specific RluD pseudouridine synthases which are involved in the bacterial LSU assembly and responsible for creation of pseudouridines at positions 1911, 1915 and 1917 *(E. coli* numbering) in the H69 end-loop (Gutgsell et al 2001; Gutgsell et al 2005). In pre-mtLSU structure, TbLAF4 encircles the immature rRNA nucleotides U1017-U1021 and G1073-U1087 (Fig 4), and its active site is occupied by uridine 1017 likely present in H89-93 of the mature rRNA (Fig EV2B). The N-terminal domain is positioned at the distance of ~80 Å facing towards the L7/L12 stalk. This connection plays a key role in coordinating the factors between the different functional sites (Fig 5A).

TbLAF5 belongs to the family of SpoU RNA methyltransferases, but appears to have a closed active site that does not allow the binding of its typical S-adenosyl methionine cofactor (Fig EV2A). It is located peripherally, and bound to the mitoribosome via a C-terminal 24-residue helix interacting with rRNA H41/42, and via contacts with TbLAF8 (Fig EV2).

Together, TbLAF4 and TbLAF5/TbLAF8 perform a structural scaffolding role for binding mt-RbgA. A homology search of the assembly factors reveals that TbLAF4 and TbLAF5 are also present in vertebrates, plants, fungi and ciliates (Fig 5C). Since their genomes encode mt-RgbA as well, our data suggests that a cooperative action of the assembly factors might be conserved.

GTPase mt-EngA is stabilized via protein extension

In the subunit interface, we identified another conserved GTPase homolog, mt-EngA. It contains two GTPase domains arranged in tandem as well as a C-terminal K homology (KH) domain which is pointed towards the PTC. We could model two GTPs in the GTPase domains (Fig 3B). Its positioning is identical to bacteria, suggesting functional conservation. The assembly factor occupies the space between the PTC and the E-site (Fig EV1), and a role in chaperoning rRNA has been proposed (Zhang et al 2012) . The comparison with *E. coli* LSU:EngA complex reveals conformational differences that highlight the nature of the mitochondrial protein-rich system, and its role in the stabilization of the conserved assembly factors.

Firstly, the N-terminal GTPase domain is extended by 60 residues, with residues 101-108 stabilizing a helix-turn-helix motif (275-305), which remained unresolved in the bacterial complex (Fig EV1B). The N-terminal extension is generally present in mitochondria from other species (Fig EV3B, Appendix Figure S4). This motif is important for the stabilization of EngA, because one helix is stacked against a helix of TbLAF10, whereas the other forms a helical bundle with TbLAF11 (Fig EV3). TbLAF10 is one of the largest mitoribosomal assembly factors, which spans 110 Å through the rRNA core to the CP (Fig 3, Fig 5). It has an rRNA contact area of ~8400 Å and keeps the nucleotides U904-A917 as well as U953-A965 (putative H80-83) in a premature conformation (Fig 4).

Secondly, the N-terminal residues 72-75 of EngA stabilize a short helix (residues 367-374), which is buried within rRNA groove via Arg367 and Arg369 (Fig 3b). It disrupts the local structure of H75 and stabilizes the flipped nucleotide A894. This loop is also highly charged in the corresponding *E.coli* structure, but does not adopt the helical conformation observed here. Finally, the N-terminus forms additional contacts with five mitoribosomal proteins (bL28m, bL35m, bL19m, mL64, mL74), a stabilizing protein mass that compensates for the missing rRNA in this region. Overall, while the N-terminal GTPase domain aligns well with the bacterial EngA, its interacting partners in our structure are proteinaceous and specific to mitochondria.

The conserved globular domains of mt-EngA are associated with the pre-mtLSU core via TbLAF9. Its three helices from the N-terminus encloses the N-terminal GTPase domain helix 230-242 (Fig EV3). Here, TbLAF9 replaces the missing rRNA H82, H88 and protein L1, which binds the EngA N-terminal GTPase domain in bacteria. TbLAF9 spans over 100 Å to the top of the CP, where it also stabilizes unwound rRNA (Fig 4, Fig 5). Thus, mt-EngA is stably bound via protein extension and also associated with the protein-based scaffold of assembly factors, including the high molecular weight TbLAF9 and TbLAF10, which are connected to the CP.

The module mt-RbgA:mt-EngA coordinates maturation of interfacial rRNA

The process of the LSU assembly is dynamic with cooperative action of different assembly factors (Davis et al 2016; Davis and Williamson 2017). Although RbgA and EngA have previously been visualized separately on the bacterial LSU though deletion and reconstitution experiments, our cryo-EM structure shows both factors simultaneously associated with the mt-LSU and with each other. The presence of both factors rationalizes why rRNA domain V is better resolved than in the mature mtLSU (Fig 4, Fig 6). We were able to model 33% more nucleotides relative to the mature LSU, thereby showing that the H89-93 region does not occupy the expected bacterial position and highlighting a need for prominent remodeling during maturation (Fig 6, Appendix Fig S5).

The contacts between mt-RbgA and mt-EngA are formed between the N-terminal domain and KH domains, respectively (Fig 3b). The shared surface area is ~540 Å, and each of the domains is also associated with conserved rRNA. The formed contacts between mt-RbgA and mt-EngA include electrostatic interactions, as well as hydrophobic residues (Fig 3b). Since the structures

and positions of both factors are conserved with bacteria, and we identified homologs in representative eukaryotic species, these results indicate that the simultaneous binding and overall mode of interaction might be conserved in mitochondria of different organisms, including humans.

Progressive maturation of the L7/L12 stalk

The L7/L12 stalk is a universal mobile element that associates translational protein factors with the A-site. It typically consists of the rRNA platform that connects to the flexible factor-recruiting series of bL12m copies. The ubiquitous ribosomal proteins uL10m, uL11m and mitochondria-specific mL43 link the different components together. In our structure, most of the protein constituents of the stalk are missing and others adopted conformational changes (Fig 7A).

In the region of the platform, at least three proteins (uL16m, bL36m, mL88) associated with the rRNA in the mature mtLSU are absent. Consistently, the rRNA platform is not folded, as the folding relies on the missing mitoribosomal proteins. Instead, the N-terminal domain of TbLAF4 extends from the subunit interface to occupy part of the space left by uL16m absence (Fig 7A and B). It binds two specific assembly factors TbLAF2 and TbLAF7. TbLAF2 mediates further contacts with the core of the pre-mtLSU. TbLAF7 consists of 7-stranded beta-barrel, 12 α-helices, and 63-residue tail inserted into the mitoribosomal core. The structure suggests that both TbLAF2 and TbLAF7 need to dissociate for the missing mitoribosomal proteins to assemble (Fig 7A).

In the factor-recruiting region, instead of the terminal bL12m copies, TbLAF3 and density corresponding to UNK6 form a protrusion (Fig 7A). They comprise a protein continuum of at least 13 helices associated with each other. This assembly is the attached to the platform region through a 30-residue C-terminal tail of TbLAF3, forming a helical bundle with mL75 (Fig 7A, 7C). Overall, this protein module extends ~70 Å from the core in a similar fashion to the L7/L12 stalk, and both appear to be mutually exclusive.

The position of the uL10m N-terminus, which links the stalk to the body in the mature LSU, is occupied by TbLAF3 C-terminus. It interacts with mL43, resulting in its helix being bent by 90° (Fig 7B). This conformational change and the lack of uL10m is correlated with ~15 Å shift of uL11m to occupy the formed void (Fig 7B). Nevertheless, TbLAF3 is only peripherally associated with mL43, and it cannot be excluded that the protrusion is independently replaced by the mature L7/L12 stalk.

Together, our structure proposes a working model for the L7/L12 stalk maturation via five steps which include dismantling, remodeling and assembly (Fig EV5): 1) TbLAF4, which is extended from the subunit interface anchoring mt-RgbA, has to be removed; 2) TbLAF2:TbLAF7 is released from the ribosomal core; 3) rRNA platform is folded, and mitoribosomal proteins uL16m, bL36m, and mL88 are recruited to the rRNA platform; 4) TbLAF3:TbLAF5 is removed, and uL11m, mL43 adopt their mature conformations; 5) bL10m and bL12m stalk proteins are associated to form the functional L7/L12 stalk.

Below the L7/L12 stalk we identified three additional proteins: an ortholog of the bacterial ribosome silencing factor mt-Rsf, a LYR (leucine-tyrosine-arginine) motif containing protein TbLAF1, as well as an associated mt-ACP (Fig 1, Fig 5). The human and fungi mitochondrial counterparts of these three proteins have been shown to be involved in late assembly stages of the mitoribosome, preventing association of the mtSSU (Brown et al 2017; Itoh et al 2020). In our structure, the module is further stabilized by mL85, however it does not appear to obstruct the binding of the mature mtSSU, likely because of the reduced rRNA and different overall shape. This suggests that in order for the mtACP:TbLAF1 protein module to act according to the previously suggested mechanism, a complementary structural domain from pre-mtSSU would be required for a steric hindrance.

Assembly of the CP is linked to the subunit interface and L1 maturation via mt-ACP

The most prominent architectural feature of the pre-mtLSU complex is the absence of the CP. It is a universal ribosomal element that defines the shape of the LSU and forms bridges with the SSU head. In mitoribosomes, the CP is particularly protein-rich (Amunts et 2014; Greber et al 2014; Amunts et 2015; Greber et al 2015; Waltz et al 2020; Tobiasson and Amunts 2020). The acquisition of the CP mitochondria-specific proteins took place in the early evolution of the mitoribosome and therefore expected to be conserved (Petrov et al 2019).

In the pre-mtLSU, all the CP mitoribosomal proteins are missing and the high molecular weight assembly factors TbLAF6 (69 kDa) and TbLAF9 (67 kDa) are present (Fig 1, Fig 8A). TbLAF6 binds on the solvent side of the CP covering the otherwise exposed rRNA, which only engages in limited base pairing. This assembly factor is annotated as a putative TruD family pseudouridine synthase. However, in our structure, it displays a two-strand antiparallel β-sheet near the catalytic site protruding into the ribosomal core interacting with four mitoribosomal proteins (uL15m, bL28m, bL35m, mL74) (Fig 8B). TbLAF9 is an exclusively helical protein, comprised of at least 29 helices. It binds on top of the rRNA, providing an additional protective protein cap (Fig 8A). Two of the TbLAF9 helices interface with a four-helix bundle, which we identified as mt-ACP (mt-ACP2) with a local resolution of 3.5 Å (Fig 2). Since mt-ACP proteins are known to form interactions with Leu-Tyr-Arg (LYR)-motif proteins, we compared the mt-ACP1-TbLAF1 module from the subunit interface with the CP mt-ACP2 region (Fig 8C). The helices of the LYR-motif protein TbLAF1 aligned well with a density corresponding to three helices associated with the mt-ACP2. The interactions in both cases are mediated by the 4′-phosphopantetheine (4′-PP) modification of mt-ACP, resembling the canonical interactions between mt-ACP and the LYR-motif proteins. The 4′-PP appears to be acylated as indicated by the density but the exact length of the acyl chain cannot be unambiguously determined from the density (Fig 8C).
The presence of the 4′-PP modification, previous structural data (Zhu et al 2015; Fiedorczuk et al 2016; Brown et al 2017), and the overall shape of the associated density suggest that the interacting partner of mtACP2 is a protein from the LYR-motif family containing at least three helices. Therefore, we searched our current and previously published (Zikova et al 2008) mass

spectrometry data using ScanProsite (de Castro et al 2006). The hits were subjected to the secondary structure prediction and fitting to the density map. The analysis singled out the protein Tb927.9.14050 (UniProt ID Q38D50), which we named consistently with the adopted nomenclature TbLAF11. The local resolution of 3.5–4.0 Å in this region (Fig 2) allowed for 164 N-terminal residues to be built (Supplementary File 2), which includes the three helices associated with the mtACP2 in proximity to the L1 stalk, and two helices interacting with EngA. The importance of the mt-ACP2:TbLAF11 protein module in our structure is of twofold. First, it directly binds the L1 stalk and locks it in an inward facing nonfunctional conformation (Fig 1). Second, it is involved in mt-EngA stabilization by forming a U-shaped continuum from TbLAF9 on the solvent side of the CP to the subunit interface (Fig 1, Fig 5). Therefore, it contributes to the protein network connecting between the various functional sites. Our data suggests that mt-ACP2 is a principal partner in the mitoribosomal assembly pathway. In the pre-mtSSU, mt-ACP was also identified as one of the assembly factors (Saurer et al 2019). In addition, ACPs in mitochondria act as subunits of the electron transport chain (Zhu et al 2015; Fiedorczuk et al 2016) and Fe-S cluster assembly complexes (Van Vranken et al 2016). This further supports the proposed concept that mt-ACPs could be signaling molecules in an intramitochondrial metabolic state sensing circuit (Masud et al 2019).

At the CP, the assembly factors cooperatively bind unwound rRNA nucleotides U934-U953 (H83) (Fig 5, Fig 8A). Remarkably, the rRNA forms a loop 25 Å in diameter that contains the TbLAF6 β-sheet and mL64 C-terminal helix, both inserted from the opposite directions (Fig 8B). The conserved helix of mL64 is shifted in our pre-mtLSU structure ~30 Å from the final location on the mature LSU, where it aligns the E-site. Interestingly, this is also one of the most conserved mitoribosomal proteins across species (Petrov et al 2019). To switch to the conserved and mature position, the extended C-tail of mL64 has to liberate from the rRNA loop and then undergo a large conformational shift towards the L1 stalk. Subsequently, the C-tail is inserted to its mature position, where it contacts CP components absent from the assembly intermediate. Since the L1 stalk is also shifted, the maturation towards a mature LSU is likely to occur in a concerted manner upon the release of the mt-ACP2:TbLAF11 module.

Conclusions

Our structure of pre-mtLSU in complex with 16 assembly factors reveals that high molecular weight assembly factors shield the rRNA and form a network that spans over 180 Å, which connects the subunit interface with the progressive maturation of the L7/L12 stalk, L1 stalk, and the assembly of the CP. The tight binding of the mt-ACP2 with partner proteins emphasizes its role in forming the connections. The factors mt-RbgA and mt-EngA are held together by protein linkers and interact with each other. The analysis shows that most of the assembly factors are also found in mammals. Overall, the protein-based communication allows the formation of the defined pre-mitoribosomal complex with newly identified factors that provides a conceptual model how mtLSU can progress through the biogenesis pathway.

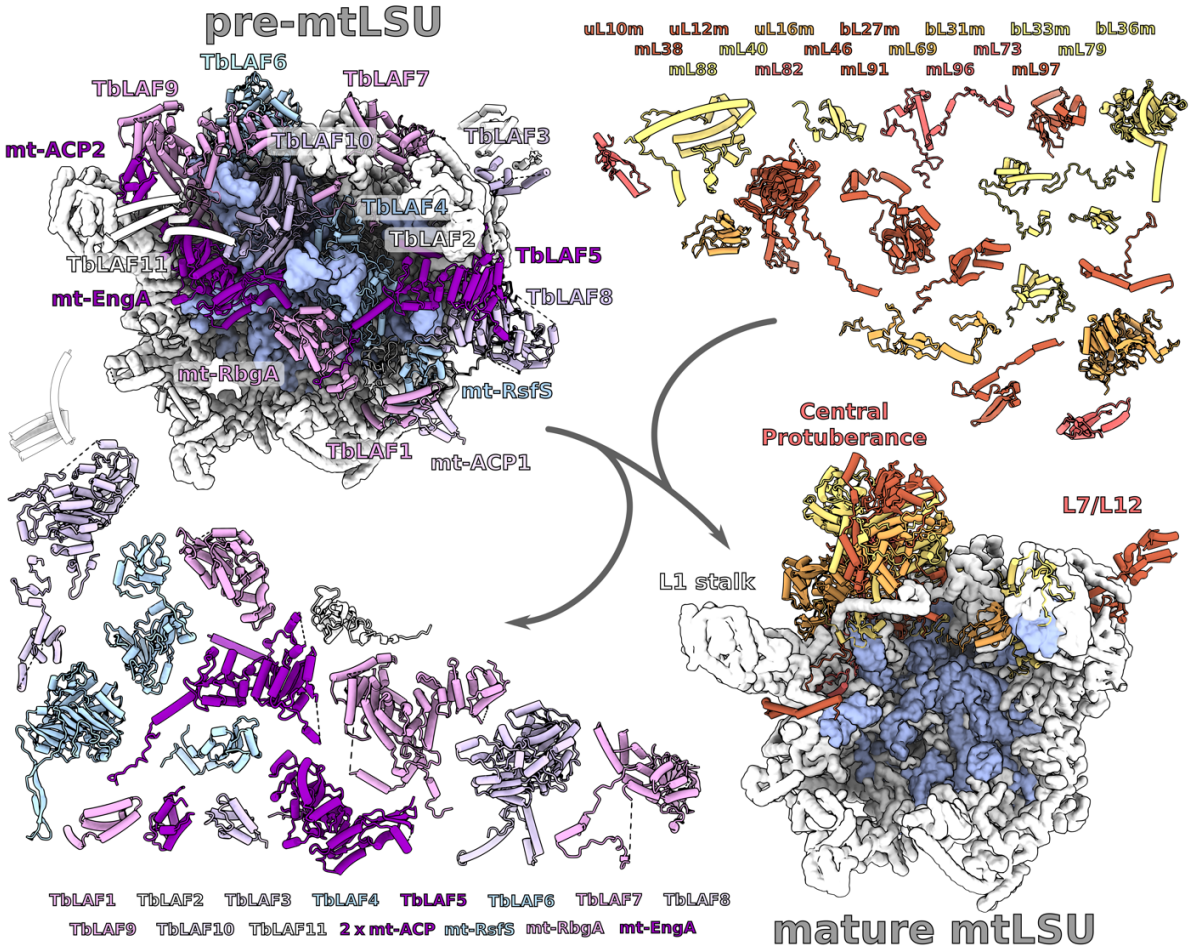


Figure 1. Structure of *T. brucei* pre-mtLSU with 16 assembly factors. Left, the overall modeled structure of the pre-mtLSU (rRNA blue) with models of assembly factors (shades of purple and white) covering the subunit interface, CP, L7/L12 stalk and connecting to the L1 stalk. Right, structure of the mature mtLSU (PDB ID 6HIX) with 18 additional mitoribosomal proteins (shades of orange) absent from pre-mtLSU.

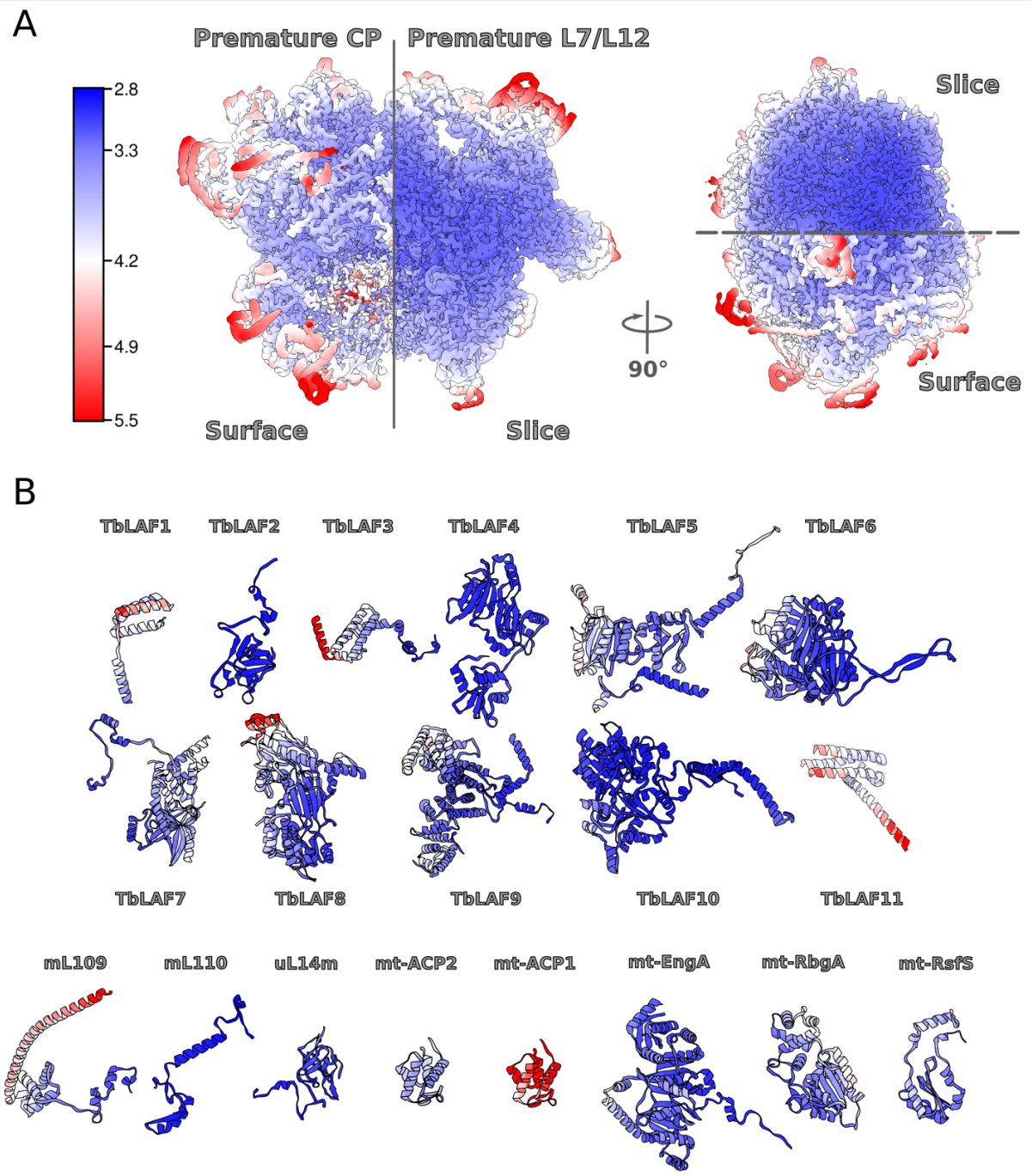


Figure 2. Cryo-EM data quality. (**A**) Final map colored by local resolution. (**B**) Models for individual assembly factors and newly identified proteins colored by local resolution of the density in the corresponding regions.

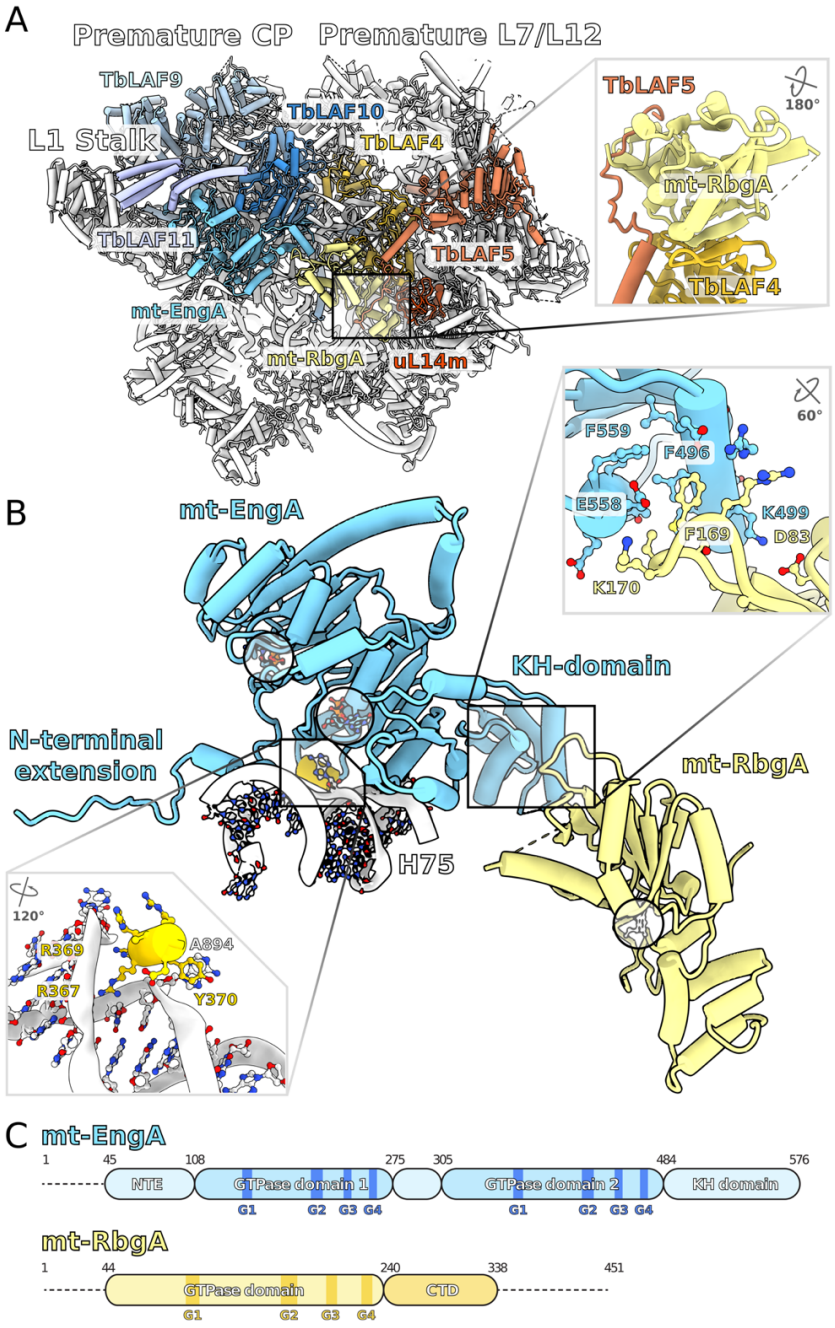


Figure 3. Binding of the mt-RgbA and mt-EngA to the subunit interface. (**A**) mt-RgbA is bound to TbLAF4 and TbLAF5, which are connected to the L7/L12 stalk; mt-EngA is associated with TbLAF9 and TbLAF10, which are connected to the CP. (**B**) A short helix of mt-EngA (yellow) interacts with a flipped A894 nucleotide from H75 (white). Two GTPs in their binding sites on mt-EngA are shown as sticks. A superimposed GTP in its binding site on mt-RgbA is shown in as white sticks. The residues forming interactions between mt-EngA and mt-RbgA are shown in the top right inset. (**C**) Schematic representation of mt-EngA and mt-RbgA indicating the positions of the conserved GTP binding motifs.

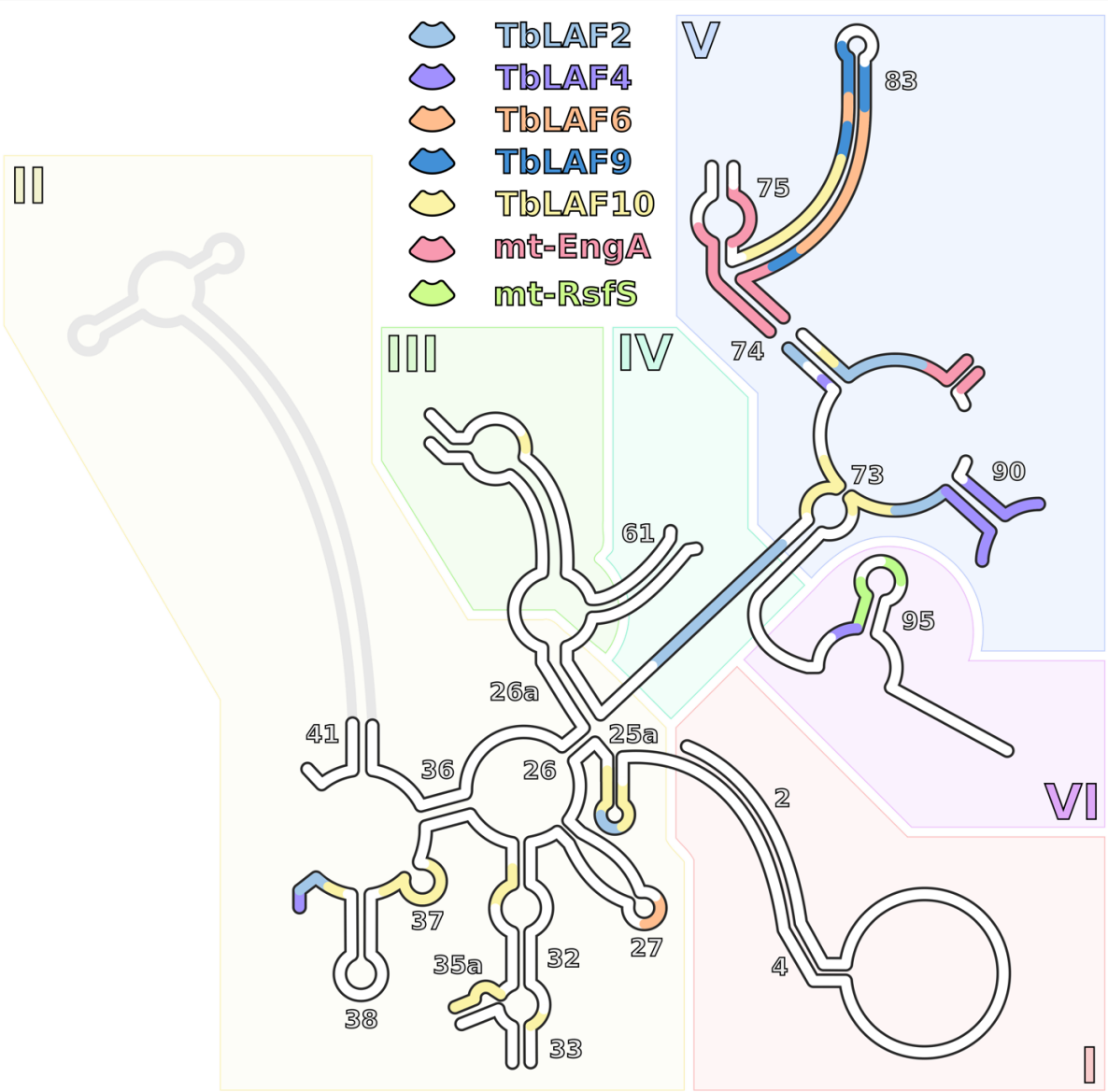


Figure 4. Schematic representation of assembly factors' binding to rRNA mapped on the secondary structure diagram. The rRNA regions contacting individual assembly factors are represented by different colors. Bound regions of at least 3 nucleotides are shown. For regions where more than one factor is bound, only a factor with higher local binding is shown. Unbound rRNA is white, unmodeled rRNA is grey.

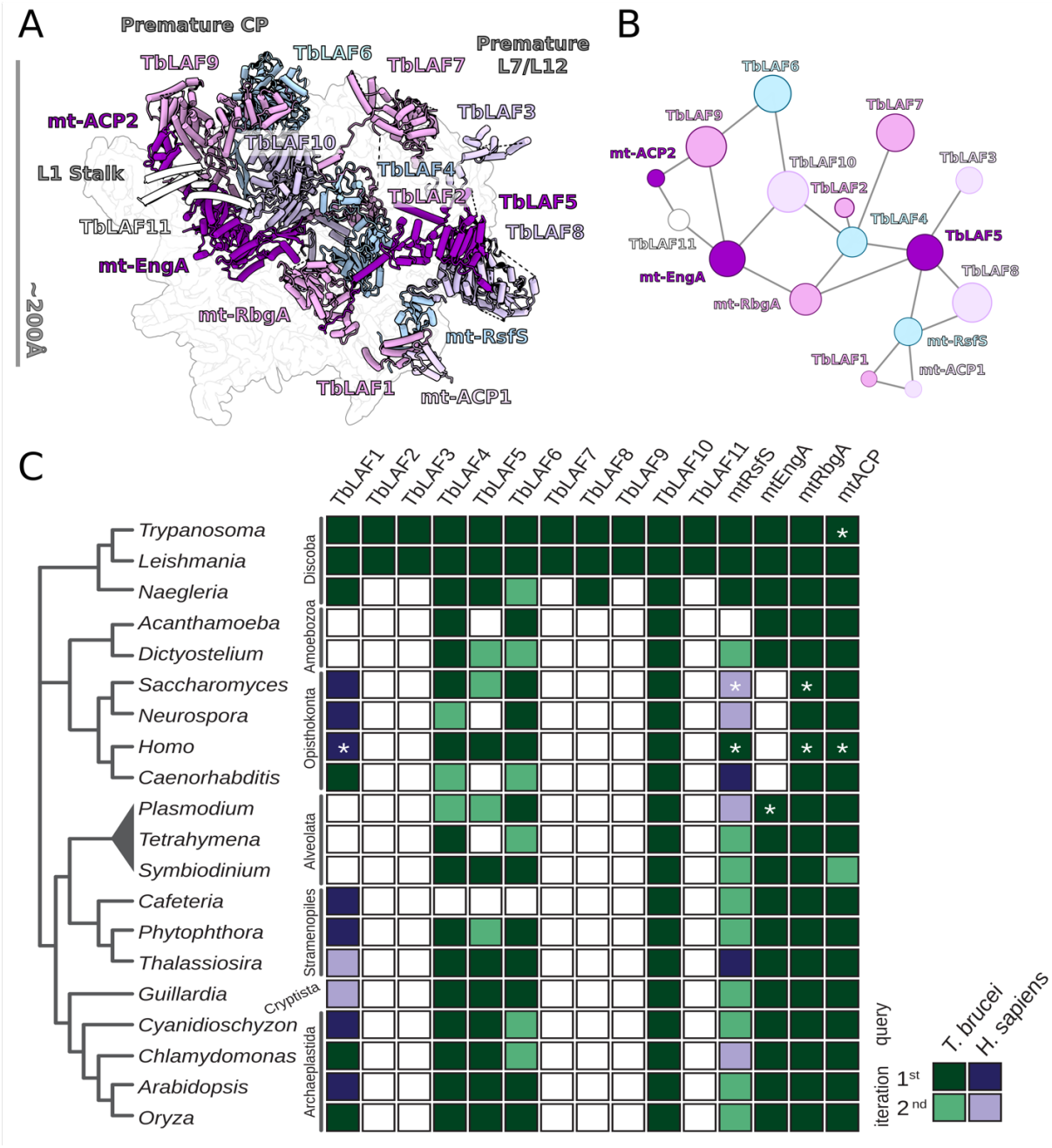
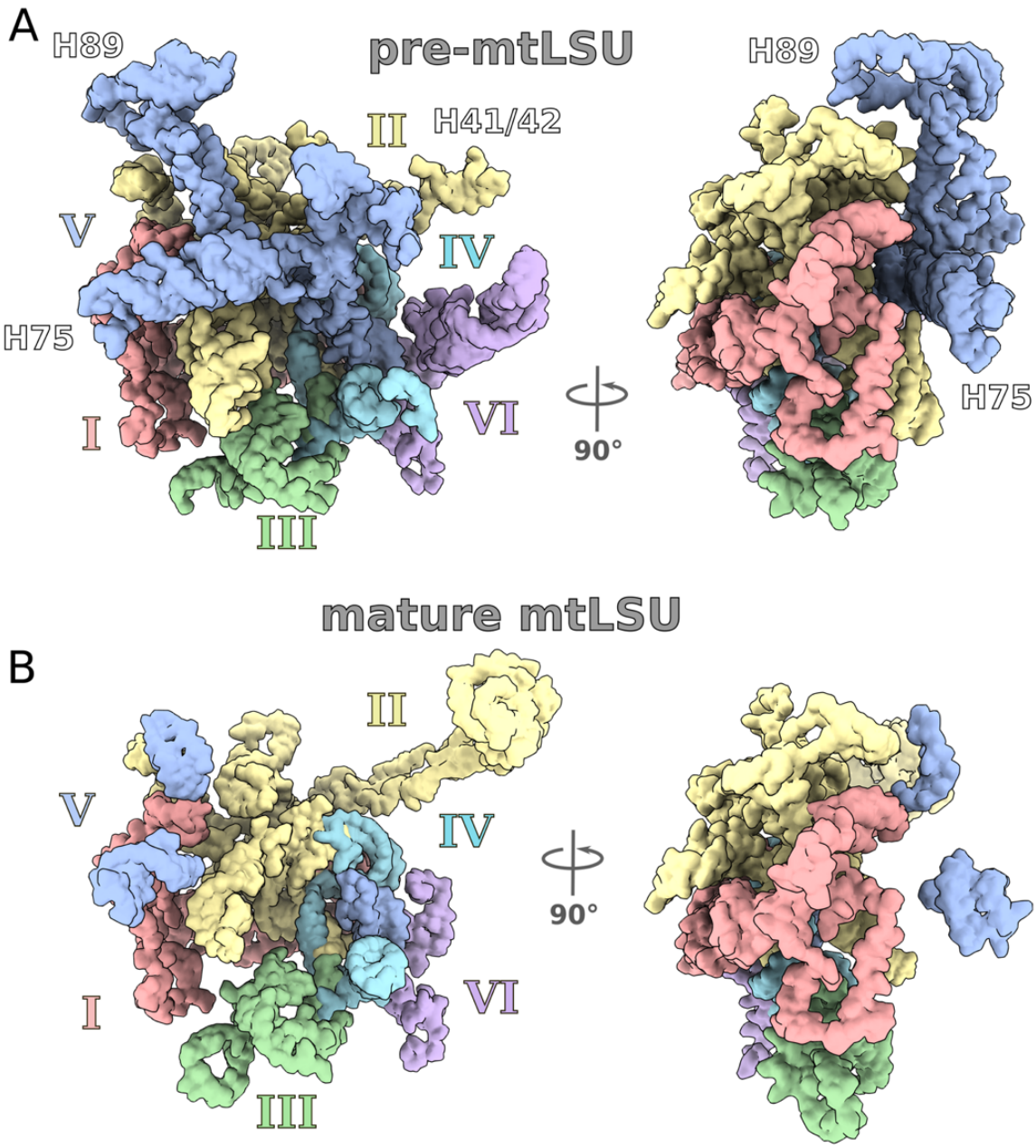


Figure 5. Network of interactions between the assembly factors in pre-mtLSU. (**A**) Assembly factors shown on the background of the pre-mtLSU density map, featuring the interconnection. (**B**) Schematic of protein-protein network. The node size represents the relative molecular mass of the protein. All the assembly factors are interconnected with each other. (**C**) Homology search of the assembly factors. Colored squares indicate identified homologs/orthologs using *T. brucei* (green) or human (purple) assembly factors as queries. White squares indicate not-identified homologs/orthologs. The stars mark proteins, for which experimental data has been reported.



385

386 **Figure 6. Tertiary structure of rRNA in pre-mtLSU (A) and mature mtLSU (B)**. Shown
387 from the subunit interface (left) and sideview (right). Two views of rRNA related by 90° are
388 shown with each domain in a different color. Domain V is more structured in pre-mtLSU, and
389 H89-93 adopt a different conformation. Domain II that is responsible for L7/L12 stalk is largely
390 disordered.

391

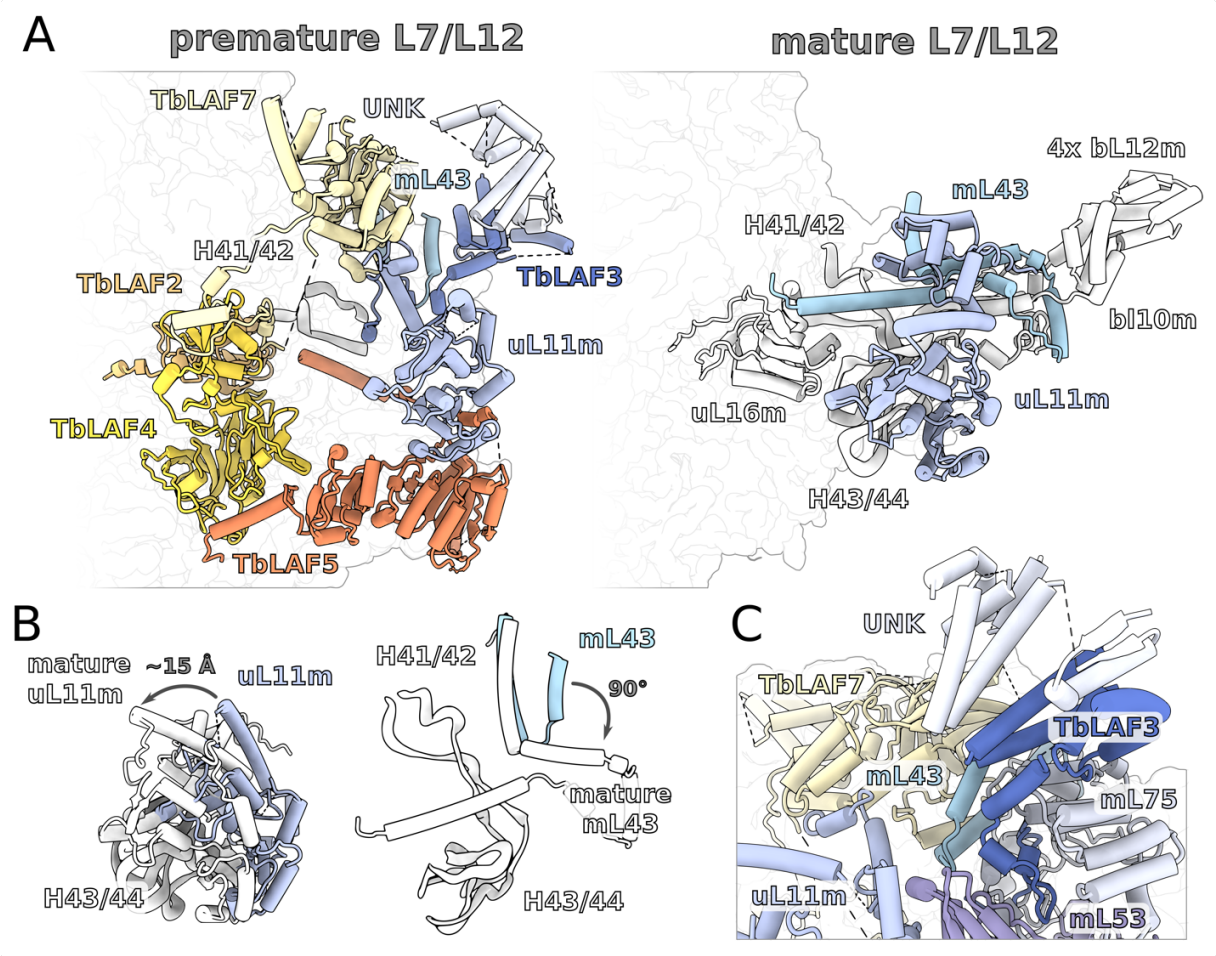


Figure 7. Assembly of the L7/L12 stalk. (**A**) In pre-mtLSU, TbLAF4 extends from the subunit interface to occupy the position of uL16m in the mature mtLSU. TbLAF2 and TbLAF7 are bound at the stalk base to the unfolded rRNA H41/42. TbLAF3 and additional protein form a protrusion similar to bL10m:bL12m. Other mitoribosomal protein removed for clarity. (**B**) Conformational changes from pre-mtLSU (green) to mature mtLSU (white) include mL43 and uL11m. (**C**) TbLAF3, mL75, and UNK protein form continuum of at least 13 helices that is peripherally associated. (**D**) Model for the L7/L12 stalk maturation.

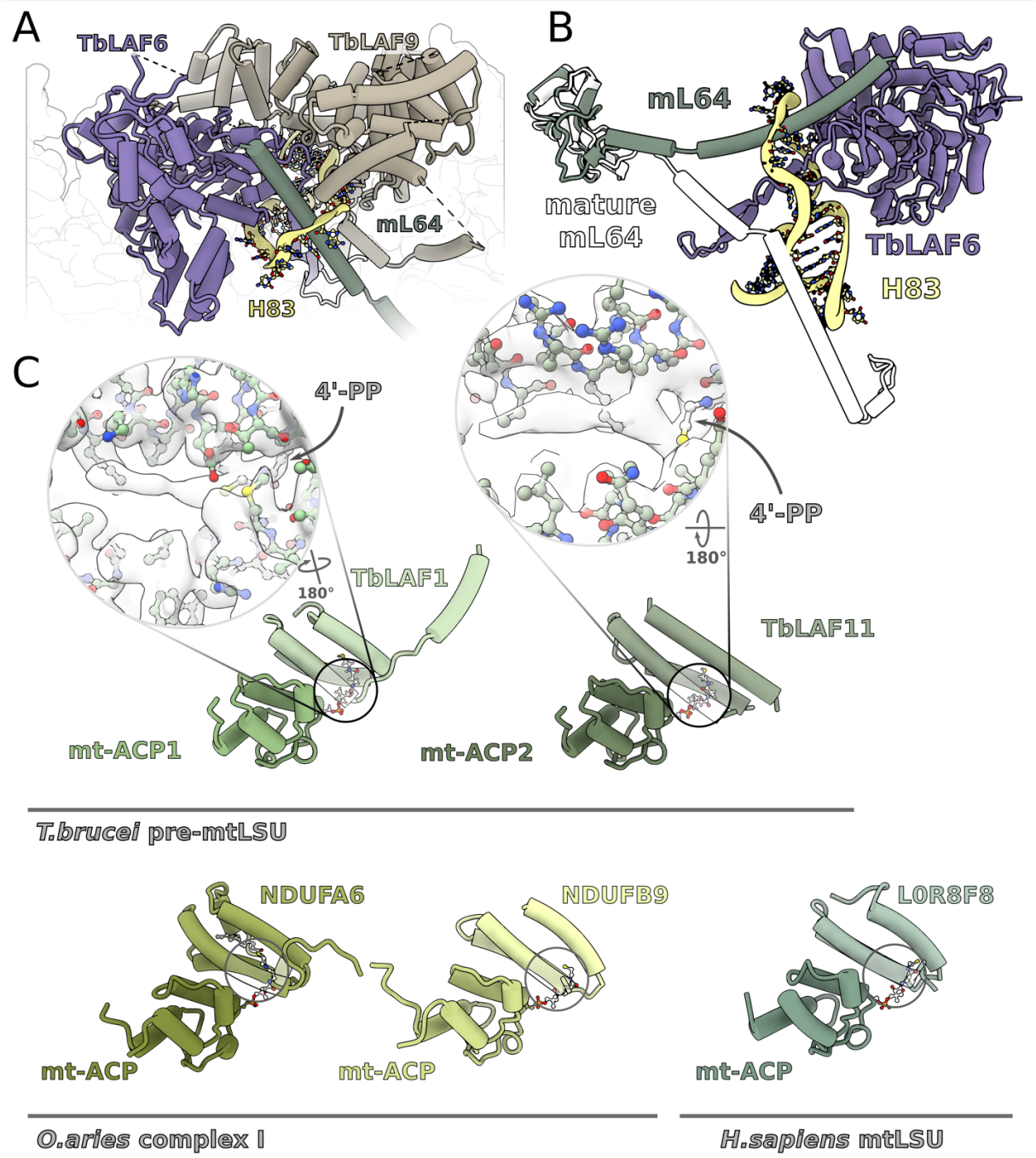


Figure 8. The CP assembly intermediate. (**A**) TbLAF6 and TbLAF9 form the CP in the pre-mtLSU. (**B**) TbLAF6 and mL64 elements are inserted through rRNA loop corresponding to H83. Conformational change of mL64 from pre-mtLSU to mature mtLSU (white) is indicated. (**C**) Comparison between the mt-ACP1:TbLAF1 (left) and the CP mt-ACP2:TbLAF11 region (right). The density (white) for acylated 4′-PP is indicated. Bottom panel, comparison with mt-ACP and associated LYR-motif proteins from complex I (PDB ID 5LNK) and human mitoribosome (PDB ID 5OOM) shows the canonical interactions.

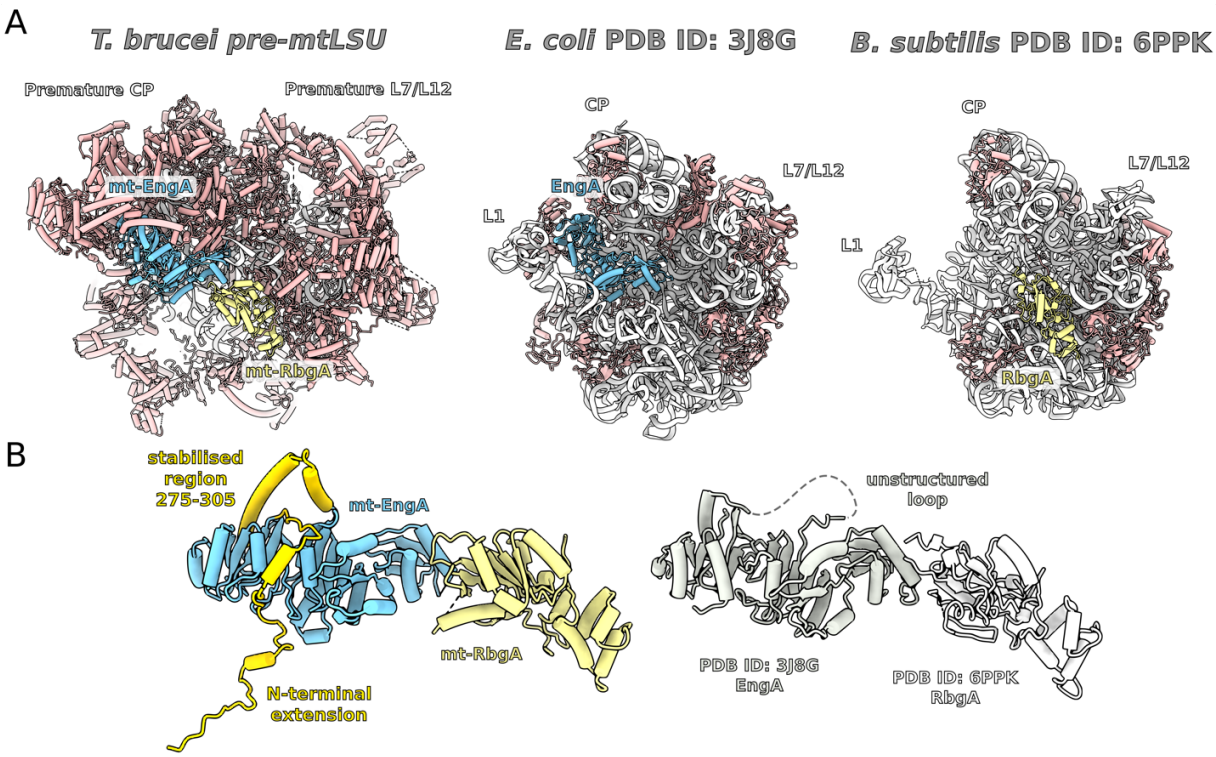


Figure EV1. The binding of mt-RbgA and mt-EngA at the mtLSU interface. (**A**) Comparison between pre-mtLSU and bacterial counterparts *E. coli* 50S:EngA (PDB ID 3J8G) and *B. subtilis* 45S:RbgA (PDB ID 6PPK) shows nearly identical positions of the factors on their ribosomal complexes. (**B**) Comparison between mt-RbgA:mt-EngA module from the pre-mtLSU and superimposed bacterial counterparts combined from the two structures from (A) shows nearly identical conformations of the factors. The N-terminal extension of mt-EngA (orange) is buried in the mitoribosomal core and stabilizes the binding, as well as 275-305 region (orange).

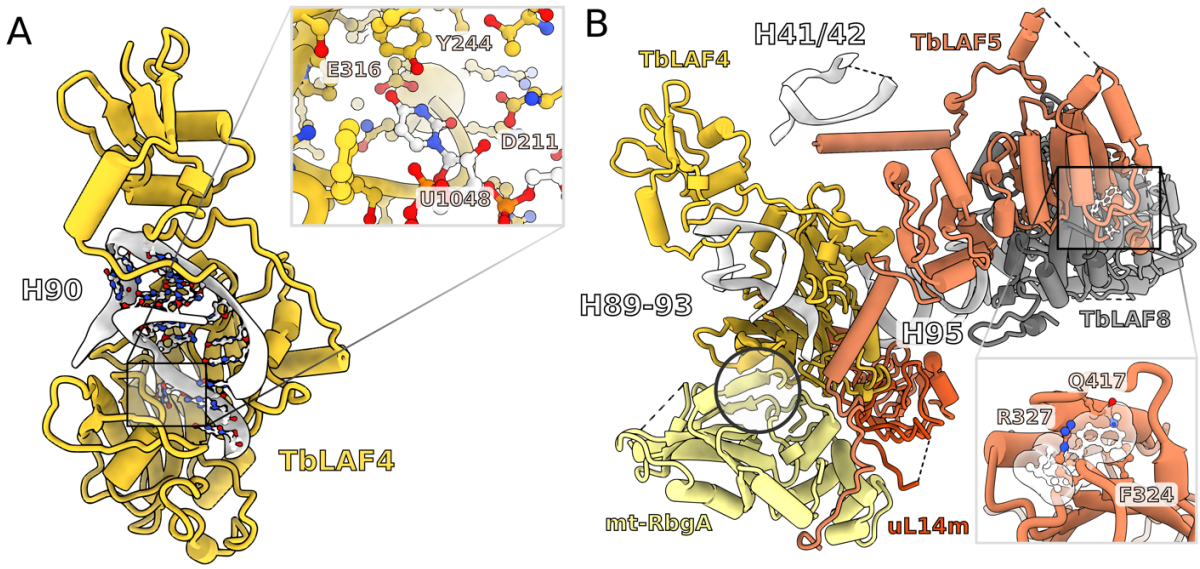


Figure EV2. (**A**) The active site of TbLAF4 (yellow) is occupied by uridine 1048. (**B**) The factor TbLAF4 (yellow) binds mt-RbgA (pale-yellow) via a shared β-sheet (circled). The methyltransferase site of TbLAF5 does not allow for the binding of S-adenosyl methionine cofactor (white sticks and surface) due to clashes with the protein residues (red).

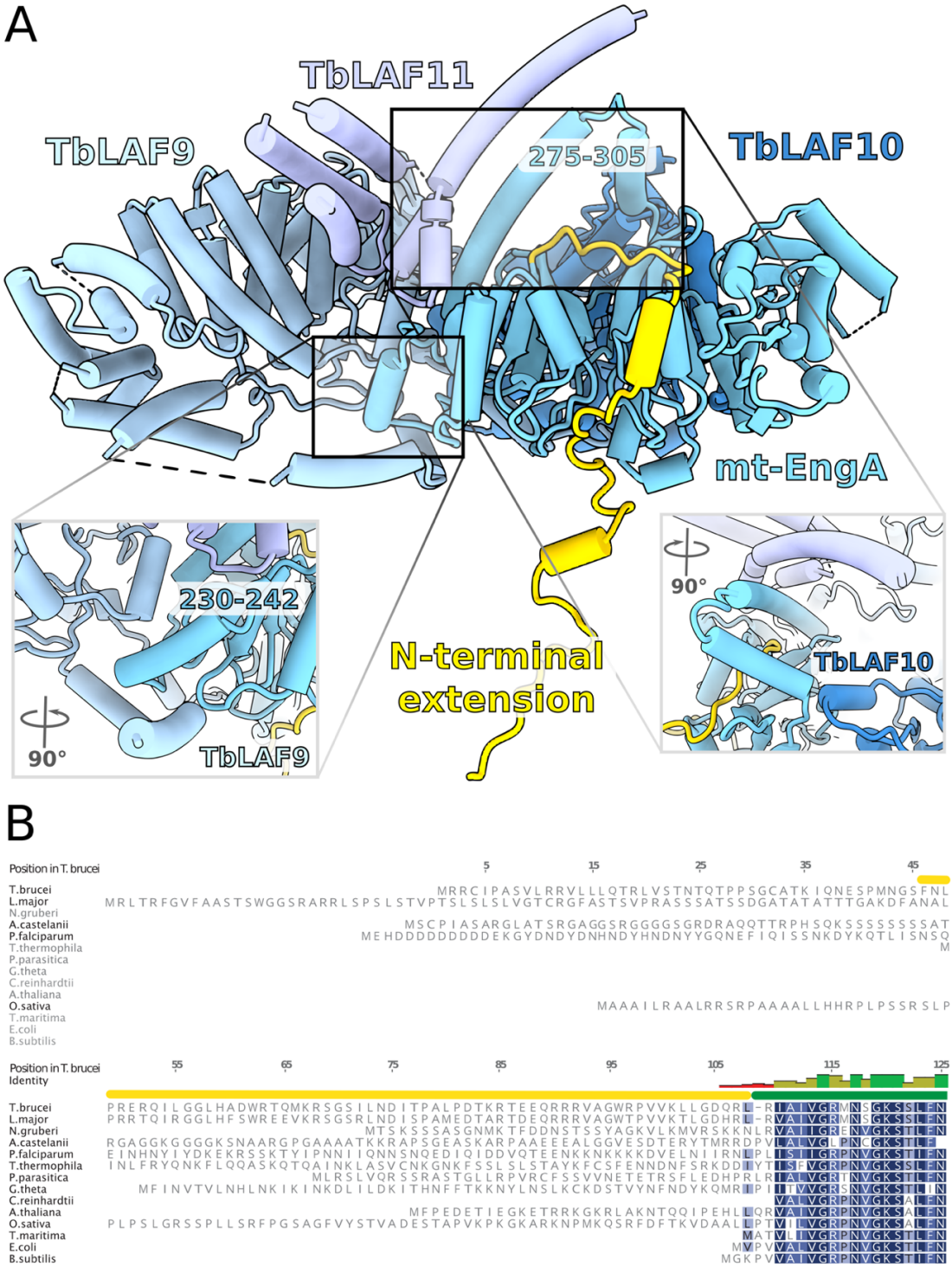


Figure EV3. (**A**) N-terminal extension (yellow) of mt-EngA stabilizes helix-turn-helix (275-305), which forms interaction with TbLAF10 on the other side (bottom right panel), and a helical bundle with TbLAF11 that is in contact with TbLAF9. (**B**) Sequence alignment of the N-terminus of mt-EngA shows presence of the extension in different organisms.

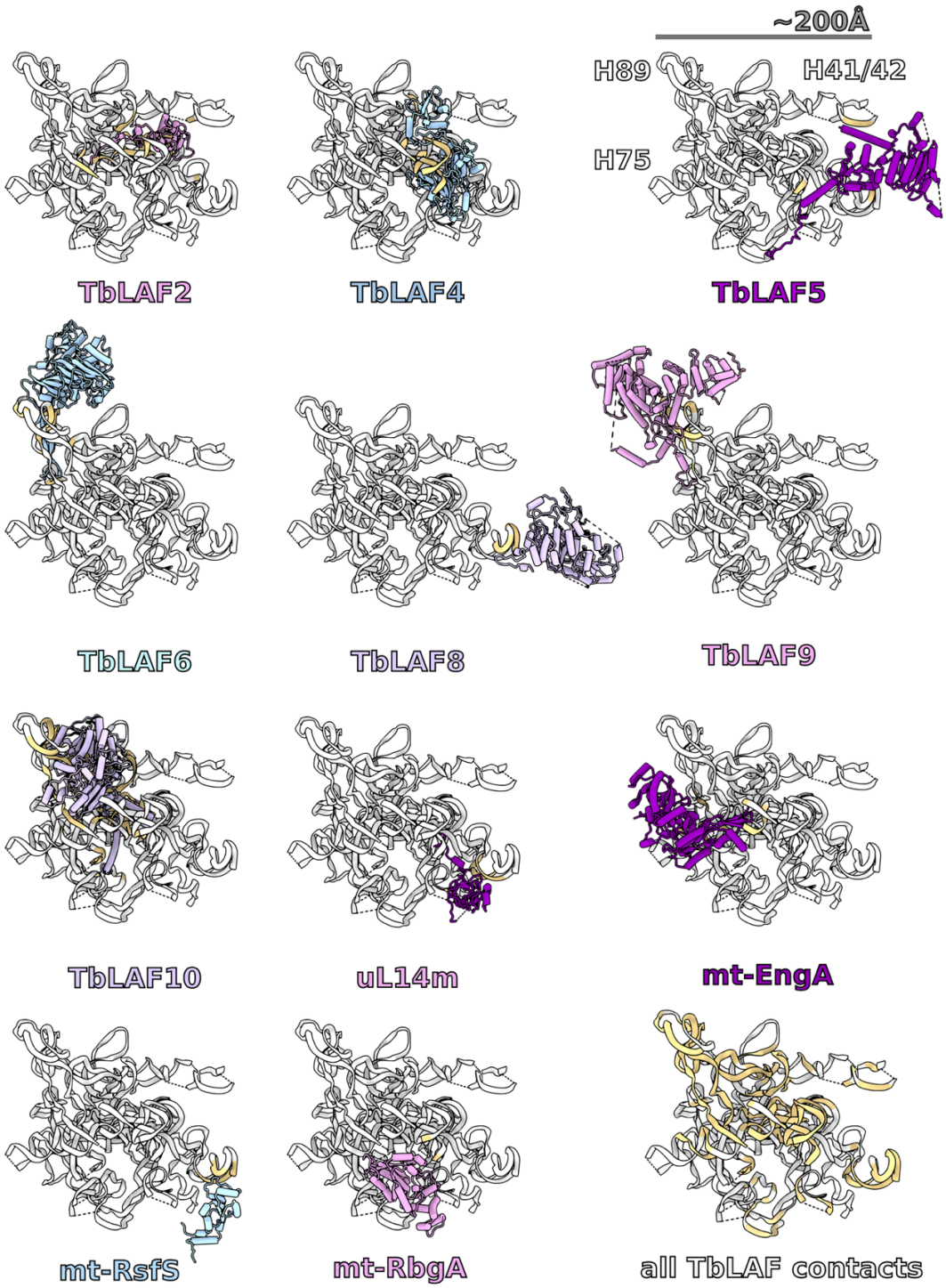
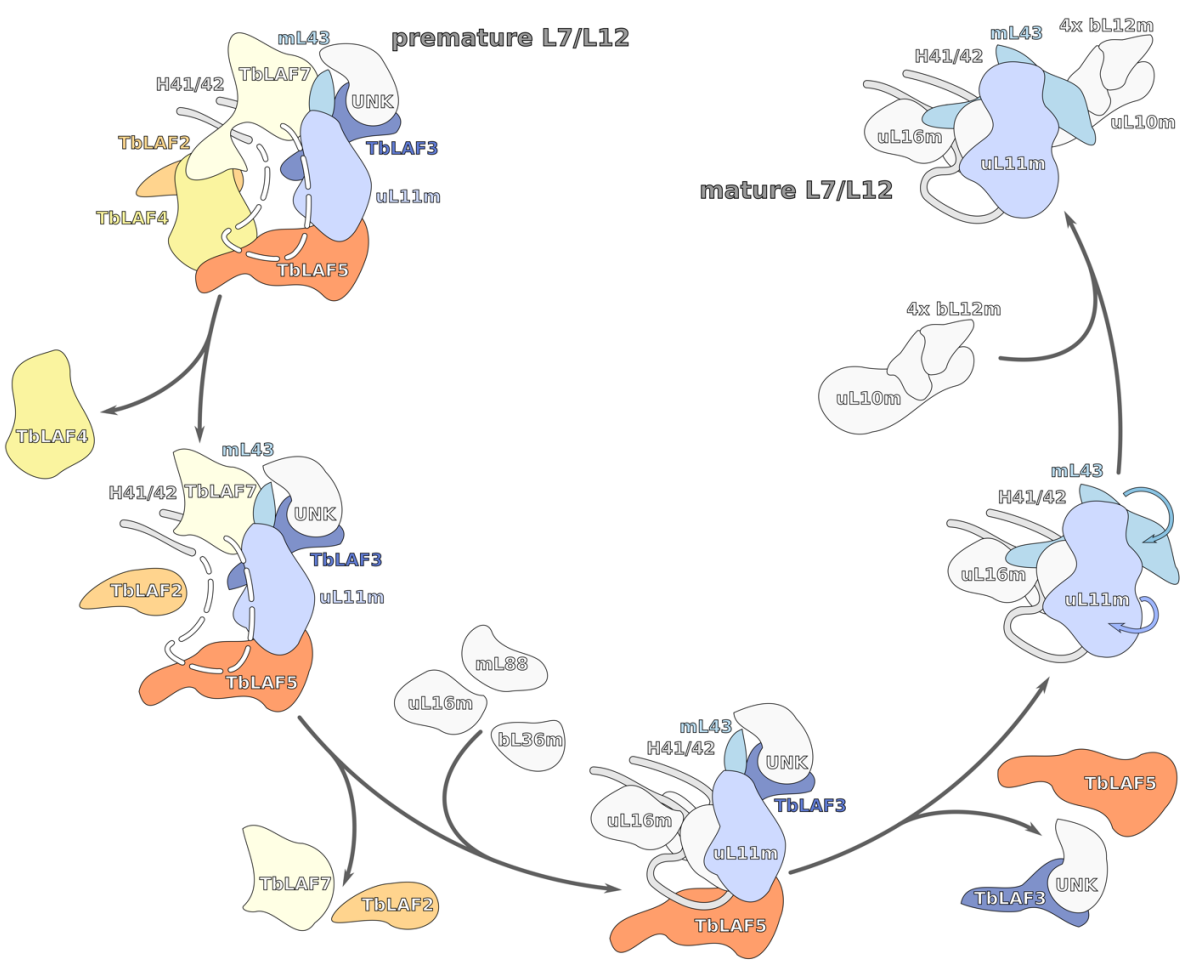


Figure EV4. Binding of assembly factors to rRNA. For each panel, rRNA is shown with an individual protein characterized in the structure, which have not been reported in the mature LSU. Bottom right panel illustrates the total RNA that is involved in the interactions (yellow) with the assembly factors. Regions and nucleotides of respective rRNA domains are also presented in Supplementary File 3.



437

438 **Figure EV5. Proposed model for the L7/L12 stalk maturation.** The series of steps starts with
439 dismantling the assembly factors from the unfolded rRNA (white dashes) that triggers rRNA
440 folding (white line), binding of the mitoribosomal proteins (grey) and conformational changes
441 (arrows).

442

Materials and Methods

Strains and growth conditions

T. brucei procyclic Lister strain 427 was grown in SDM-80 medium supplemented with 10% fetal bovine serum. Mitochondria were isolated as described earlier Schneider (2007). $1.5x10^{11}$ cells were harvested, washed in 20 mM sodium phosphate buffer pH 7.9 with 150 mM NaCl and 20 mM glucose, resuspended in 1 mM Tris-HCl pH 8.0, 1 mM EDTA, and disrupted by 10 strokes in 40 ml Dounce homogenizer. The hypotonic lysis was stopped by immediate addition of 1/6 volume of 1.75 M sucrose. Crude mitochondria were pelleted (15 min at 16000 xg, 4°C), resuspended in 20 mM Tris-HCl pH 8.0, 250 mM sucrose, 5 mM $MgCl_2$, 0.3 mM $CaCl_2$ and treated with 5 µg/ml DNase I for 60 min on ice. DNase I treatment was stopped by addition of one volume of the STE buffer (20 mM Tris-HCl pH 8.0, 250 mM sucrose, 2 mM EDTA) followed by centrifugation (15 min at 16000 xg, 4°C). The pellet was resuspended in 60% Percoll in STE and loaded on the bottom of six 10-35% Percoll gradient in STE in polycarbonate tubes for SW28 rotor (Beckman). Gradients were centrifuged for 1 hour at 24000 rpm, 4°C. The middle-diffused phase containing mitochondrial vesicles (15-20 ml per tube) was collected, washed twice in the STE buffer, snap-frozen in liquid nitrogen and stored at -80°C.

Purification of mitoribosomes

Mitochondria were further purified further using a stepped sucrose gradient (60 %, 32 %, 23 %, 15%) in a low ionic strength buffer (50 mM HEPES/KOH pH 7.5, 5 mM MgOAc, 2 mM EDTA). A thick pellet at the 60-32% interface was collected and lysed by mixing with 5 volumes of detergent containing lysis buffer (25 mM HEPES/KOH pH 7.5, 100 mM KCl, 15 mM MgOAc, 1.7 % Triton X-100, 2 mM DTT, Complete-EDTA Free Protease Inhibitor). The lysate was centrifuged at 30,000 xg twice, saving the supernatant after each spin. The supernatant was then subjected to differential PEG precipitation; PEG 10,000 was added to reach a concentration of 1.5 % (w/v) and incubated on ice for 10 mins, followed by a spin at 30,000 xg. The supernatant was transferred to a fresh tube, and PEG 10,000 was added to reach a concentration of 8 % (w/v) then incubated on ice for 10 mins, followed by a spin at 30,000 xg.

The pellet was then resuspended in 800 µl of lysis buffer and then layered onto a 34% sucrose cushion (25 mM HEPES/KOH pH 7.5, 100 mM KCl, 15 mM MgOAc, 1.0 % Triton X-100, 2 mM DTT, Complete-EDTA Free Protease Inhibitor) in a TLA120.2 centrifuge tube (0.4 ml of cushion per tube). Mitoribosomes were pelleted through the cushion by centrifugation at 231,550 xg for 45 min. Pelleted mitoribosomes were resuspended using a total of 100 µl of resuspension buffer (25 mM HEPES/KOH pH 7.5, 100 mM KCl, 15 mM MgOAc, 0.01 % β-DDM, 2 mM DTT). The resuspended mitoribosomes were then layered onto a continuous 15-30 % sucrose gradient and centrifuged in a TLS55 rotor for 120 min at 213,626 xg. The gradient was fractionated manually, and fractions containing mitoribosome as judged by the 260 nm absorbance were pooled and buffer exchanged in a centrifugal concentrator.

Cryo-EM and model building

For cryo-EM analysis, 3 µL of the sample at a concentration of OD260 3.5, was applied onto a glow-discharged (20 mA for 30 seconds) holey-carbon grid (Quantifoil R2/2, copper, mesh 300) coated with continuous carbon (of ~3 nm thickness) and incubated for 30 seconds in a controlled environment of 100% humidity and 4 ºC temperature. The grids were blotted for 3 seconds, followed by plunge-freezing in liquid ethane, using a Vitrobot MKIV (FEI/Thermofischer). The data was collected on a FEI Titan Krios (FEI/Thermofischer; Scilifelab, Stockholm, Sweden, and ESRF, Grenoble, Frane) transmission electron microscope operated at 300 keV, using C2 aperture of 70 µm; slit width of 20 eV on a GIF quantum energy filter (Gatan). A K2 Summit detector (Gatan) was used to collect images at a pixel size of 1.05 Å (magnification of 130,000X) with a dose of ~35 electrons/Å2 fractionated over 20 frames. A defocus range of 0.8 to 3.5 µm was applied.

19,158 micrographs (after bad images were removed based on real and reciprocal space features) were collected across 5 non-consecutive data acquisition sessions and processed together using RELION. 896,263 particles were picked using Warp and coordinates were imported into RELION for particle extraction at an initial binning factor of two. The particles were subjected to supervised 3D classification using references generated previously in a screening dataset, which was started based on the *T. brucei* cytosolic ribosome as an initial model. This crude separation classified the 207,788 particles as mtLSU-like, and the remaining as mature mtLSU-like, SSU-like or monosomes. This subset was subjected to auto-refinement separately to improve the angular assignments and then classified further using fine-angular searches with a solvent mask applied. From the mtLSU-like particles, 32,339 particles were retained as pre-mtLSU of good quality and the rest were discarded as non-particles. The retained pre-mtLSUs were then subjected to auto-refinement once more to improve the angles further, this time applying a solvent mask during the refinement procedure, and then the 3D reconstructions obtained were used as a reference for CTF refinement to improve the reconstruction. The final map was then estimated for local resolution using RELION and sharpened with a B-factor appropriate for the reconstruction as estimated automatically using the postprocessing procedure.

Model building was done using *Coot* 0.9 (Emsley et al 2010). First the model of the mature mtLSU (PDB ID:6HIX) was fitted to the density. Chains present in the pre-mtLSU were then individually fitted and locally refined. Additional chains were first identified using information from sidechain densities. First the map density, chemical environment and sidechain interactions were used to create probable sequences. Those sequences were then queried against *T. brucei* specific databases; potential hits were evaluated individually and finally assigned. Models were modeled de-novo. All models were refined iteratively using PHENIX (Liebschner et al 2019) realspace refinement and validated using MolProbity (Williams et al 2018). The data collection, model refinement and validation statistics are presented in Supplementary data file 1. All figures were prepared either in Chimera (Pettersen et al 2004) or ChimeraX (Goddard et al 2018) with additional graphical elements created using Inkscape.

Search for homologs of assembly factors and sequence alignments

Homologs of assembly factors found in our pre-mtLSU and identified by cryo-EM were searched in the NCBI protein database with Position-Specific Iterated BLAST (Altschul et al 1997) using sequences of individual factors from *T. brucei* as queries. The searches were targeted against selected genera. Sequence alignments were generated with the MUSCLE (Larkin et al 2007) algorithm in Geneious (Biomatters Ltd., New Zealand) and corrected manually.

Data availability

The electron density map has been deposited in EMDB under accession code EMD-XXX. The model has been deposited in PDB under accession code XXXX. All data is available in the paper or Supplementary Information.

Acknowledgements

The authors thank the SciLifeLab cryo-EM and mass spectrometry facilities for data collection, Nikhil Jain for comments. This work was supported by the Swedish Foundation for Strategic Research (FFL15:0325), Ragnar Söderberg Foundation (M44/16), Swedish Research Council (NT_2015-04107), Cancerfonden (2017/1041), European Research Council (ERC-2018-StG-805230), Knut and Alice Wallenberg Foundation (2018.0080), EMBO Young Investigator Program to A.A., and by and by Czech Science Foundation (18-17529S) and ERD fund (CZ.02.1.01/0.0/0.0/16_019/0000759) to A.Z., and by Czech Science Foundation (20-04150Y) to O.G. The cryo-EM facility is funded by the Knut and Alice Wallenberg, Family Erling Persson, and Kempe foundations.

Author contributions

Project conceptualization: OG, AZ, AA; Sample preparation for cryo-EM: OG, SA, AA; Data acquisition and processing: SA; Model building and validation: VT, OG, SA, RB; Structural data interpretation: VT, OG, AA; Manuscript writing and figure preparation: VT, OG, SA, RB, AZ, AA.

References

Achila, D., Gulati, M., Jain, N. & Britton, R. A. (2012) Biochemical characterization of ribosome assembly GTPase RbgA in Bacillus subtilis, *J Biol Chem.* 287, 8417-23

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*, *25*(17), 3389-3402

Amunts, A., Brown, A., Bai, X. C., Llacer, J. L., Hussain, T., Emsley, P., Long, F., Murshudov, G., Scheres, S. H. & Ramakrishnan, V. (2014) Structure of the yeast mitochondrial large ribosomal subunit, *Science.* 343, 1485-9

Amunts, A., Brown, A., Toots, J., Scheres, S. H. W. & Ramakrishnan, V. (2015) Ribosome. The structure of the human mitochondrial ribosome, *Science.* 348, 95-98

Antonicka, H. and Shoubridge, E.A., 2015. Mitochondrial RNA granules are centers for posttranscriptional RNA processing and ribosome biogenesis. *Cell reports*, *10*(6), pp.920-932

Barrientos, A., Korr, D., Barwell, K. J., Sjulsen, C., Gajewski, C. D., Manfredi, G., Ackerman, S. & Tzagoloff, A. (2003) MTG1 codes for a conserved protein required for mitochondrial translation, *Mol Biol Cell.* 14, 2292-302

Bogenhagen, D.F., Martin, D.W. and Koller, A., 2014. Initial steps in RNA processing and ribosome assembly occur at mitochondrial DNA nucleoids. *Cell metabolism*, *19*(4), pp.618-629

Brown, A., Rathore, S., Kimanius, D., Aibara, S., Bai, X. C., Rorbach, J., Amunts, A. & Ramakrishnan, V. (2017) Structures of the human mitochondrial ribosome in native states of assembly, *Nat Struct Mol Biol*. 24, 866-9

Couvillion MT, Soto IC, Shipkovenska G, Churchman LS., 2016. Synchronized mitochondrial and cytosolic translation programs. *Nature*, 533(7604):499–503

Davis, J. H., Tan, Y. Z., Carragher, B., Potter, C. S., Lyumkis, D. & Williamson, J. R. (2016) Modular Assembly of the Bacterial Large Ribosomal Subunit, *Cell.* 167, 1610-1622 e15

Davis, J. H. & Williamson, J. R. (2017) Structure and dynamics of bacterial ribosome biogenesis, *Philos Trans R Soc Lond B Biol Sci.* 372

De Castro, E., Sigrist, C. J., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P. S., Gasteiger, E., ... & Hulo, N. (2006). ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic acids research*, *34*(suppl_2), W362-W365

De Silva, D., Tu, Y.T., Amunts, A., Fontanesi, F. and Barrientos, A., 2015. Mitochondrial ribosome assembly in health and disease. *Cell Cycle*, *14*(14), pp.2226-2250

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr*. 66, 486–501

Fiedorczuk, K., Letts, J. A., Degliesposti, G., Kaszuba, K., Skehel, M., & Sazanov, L. A. (2016). Atomic structure of the entire mammalian mitochondrial complex I. *Nature*, *538*(7625), 406-410

Greber, B. J., Boehringer, D., Leibundgut, M., Bieri, P., Leitner, A., Schmitz, N., Aebersold, R. & Ban, N. (2014) The complete structure of the large subunit of the mammalian mitochondrial ribosome, *Nature.* 515, 283-6

Greber, B. J., Bieri, P., Leibundgut, M., Leitner, A., Aebersold, R., Boehringer, D. & Ban, N. (2015) Ribosome. The complete structure of the 55S mammalian mitochondrial ribosome, *Science.* 348, 303-8

Greber, B. J. & Ban, N. (2016) Structure and Function of the Mitochondrial Ribosome, *Annu Rev Biochem.* 85, 103-32.

Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H. and Ferrin, T.E. (2018) UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Science* 27, 14–25.

599 Gutgsell, N. S., Deutscher, M. P. & Ofengand, J. (2005) The pseudouridine synthase RluD is
600 required for normal ribosome assembly and function in Escherichia coli, *RNA.* 11, 1141-52.

601 Gutgsell, N. S., Del Campo, M., Raychaudhuri, S. & Ofengand, J. (2001) A second function for
602 pseudouridine synthases: A point mutant of RluD unable to form pseudouridines 1911, 1915, and
603 1917 in Escherichia coli 23S ribosomal RNA restores normal growth to an RluD-minus strain,
604 *RNA.* 7, 990-8.

605 Jomaa, A., Jain, N., Davis, J. H., Williamson, J. R., Britton, R. A. & Ortega, J. (2014) Functional
606 domains of the 50S subunit mature late in the assembly process, *Nucleic Acids Res.* 42, 3419-35

607 Kimanius, D., Forsberg, B. O., Scheres, S. H. & Lindahl, E. (2016) Accelerated cryo-EM
608 structure determination with parallelisation using GPUs in RELION-2, *Elife.* 5

609 Itoh, Y., Naschberger, A., Mortezaei, N., Herrmann, J. & Amunts, A. (2020) Analysis of
610 translating mitoribosome reveals functional characteristics of protein synthesis in mitochondria
611 of fungi, *bioRxiv* 2020.01.31.929331

612 Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., ...
613 & Thompson, J. D. (2007). Clustal W and Clustal X version 2.0. *bioinformatics*, *23*(21), 2947-
614 2948

615 Li, N., Chen, Y., Guo, Q., Zhang, Y., Yuan, Y., Ma, C., Deng, H., Lei, J. & Gao, N. (2013)
616 Cryo-EM structures of the late-stage assembly intermediates of the bacterial 50S ribosomal
617 subunit, *Nucleic Acids Res.* 41, 7073-83

618 Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I., Hintze, B.,
619 Hung, L.W., Jain, S., McCoy, A.J. and Moriarty, N.W. (2019) Macromolecular structure
620 determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta*
621 *Crystallogr D Struct Biol* 75, 861–877

622 Masud, A.J., Kastaniotis, A.J., Rahman, M.T., Autio, K.J. and Hiltunen, J.K. (2019).
623 Mitochondrial acyl carrier protein (ACP) at the interface of metabolic state sensing and
624 mitochondrial function. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1866,
625 118540

626 Ni, X., Davis, J. H., Jain, N., Razi, A., Benlekbir, S., McArthur, A. G., Rubinstein, J. L., Britton,
627 R. A., Williamson, J. R. & Ortega, J. (2016) YphC and YsxC GTPases assist the maturation of
628 the central protuberance, GTPase associated region and functional core of the 50S ribosomal
629 subunit, *Nucleic Acids Res.* 44, 8442-55

630 Nikolay, R., Hilal, T., Qin, B., Mielke, T., Burger, J., Loerke, J., Textoris-Taube, K., Nierhaus,
631 K. H. & Spahn, C. M. T. (2018) Structural Visualization of the Formation and Activation of the
632 50S Ribosomal Subunit during In Vitro Reconstitution, *Mol Cell.* 70, 881-893 e3

633 Ott, M., Amunts, A., & Brown, A., 2016. Organization and regulation of mitochondrial protein
634 synthesis. *Annual Review of Biochemistry*, *85*, 77-101.

635 Pausch, P., Steinchen, W., Wieland, M., Klaus, T., Freibert, S. A., Altegoer, F., Wilson, D. N. &
636 Bange, G. (2018) Structural basis for (p)ppGpp-mediated inhibition of the GTPase RbgA, *J Biol*
637 *Chem.* 293, 19699-19709

638 Pearce, S. F., Rebelo-Guiomar, P., D'Souza, A. R., Powell, C. A., Van Haute, L., & Minczuk, M.
639 (2017). Regulation of mammalian mitochondrial gene expression: recent advances. *Trends in*
640 *biochemical sciences*, *42*(8), 625-639.

641 Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and
642 Ferrin, T.E. (2004) UCSF Chimera – a visualization system for exploratory research and
643 analysis. *Journal of Computational Chemistry* 25, 1605–1612

644 Petrov, A. S., Wood, E. C., Bernier, C. R., Norris, A. M., Brown, A. & Amunts, A. (2019)
645 Structural Patching Fosters Divergence of Mitochondrial Ribosomes, *Mol Biol Evol.* 36, 207-219

646 Ramrath, D. J. F., Niemann, M., Leibundgut, M., Bieri, P., Prange, C., Horn, E. K., Leitner, A.,
647 Boehringer, D., Schneider, A. & Ban, N. (2018) Evolutionary shift toward protein-based
648 architecture in trypanosomal mitochondrial ribosomes, *Science.* 362

649 Saurer, M., Ramrath, D. J. F., Niemann, M., Calderaro, S., Prange, C., Mattei, S., Scaiola, A.,
650 Leitner, A., Bieri, P., Horn, E. K., Leibundgut, M., Boehringer, D., Schneider, A. & Ban, N.
651 (2019) Mitoribosomal small subunit biogenesis in trypanosomes involves an extensive assembly
652 machinery, *Science.* 365, 1144-1149.

653 Seffouh, A., Jain, N., Jahagirdar, D., Basu, K., Razi, A., Ni, X., Guarne, A., Britton, R. A. &
654 Ortega, J. (2019) Structural consequences of the interaction of RbgA with a 50S ribosomal
655 subunit assembly intermediate, *Nucleic Acids Res.* 47, 10414-10425

656 Schneider, A., Charriere, F., Pusnik, M. & Horn, E. K. (2007) Isolation of mitochondria from
657 procyclic Trypanosoma brucei, *Methods Mol Biol.* 372, 67-80

658 Tegunov, D. & Cramer, P. (2019) Real-time cryo-electron microscopy data preprocessing with
659 Warp, *Nat Methods.* 16, 1146-1152

660 Tobiasson, V. & Amunts, A. (2020) Ciliate mitoribosome illuminates evolutionary steps of
661 mitochondrial translation, *Elife.* 9, e59264

662 Van Vranken, J.G., Jeong, M.Y., Wei, P., Chen, Y.C., Gygi, S.P., Winge, D.R. and Rutter, J.
663 (2016). The mitochondrial acyl carrier protein (ACP) coordinates mitochondrial fatty acid
664 synthesis with iron sulfur cluster biogenesis. *Elife*, 5, e17828.

665 32

666 Waltz, F., Soufari, H., Bochler, A., Giege, P. & Hashem, Y. (2020) Cryo-EM structure of the
667 RNA-rich plant mitochondrial ribosome, *Nat Plants.* 6, 377-383

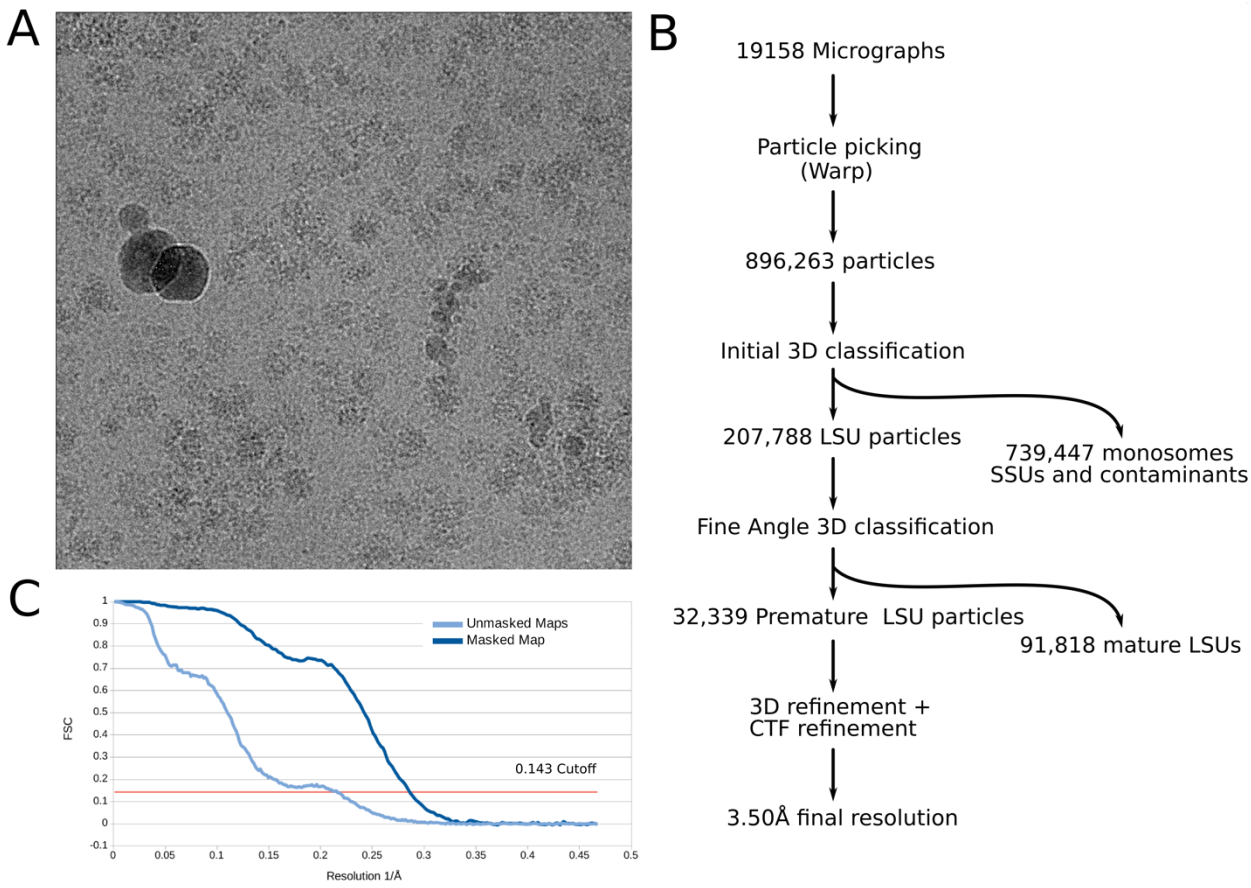
668 Williams, JS. et al. (2018) MolProbity: More and better reference data for improved all-atom
669 structure validation. *Protein Science* 27, 293-315

670 Zikova, A., Panigrahi, A. K., Dalley, R. A., Acestor, N., Anupama, A., Ogata, Y., Myler, P. J. &
671 Stuart, K. (2008) Trypanosoma brucei mitochondrial ribosomes: affinity purification and
672 component identification by mass spectrometry, *Mol Cell Proteomics.* 7, 1286-96.

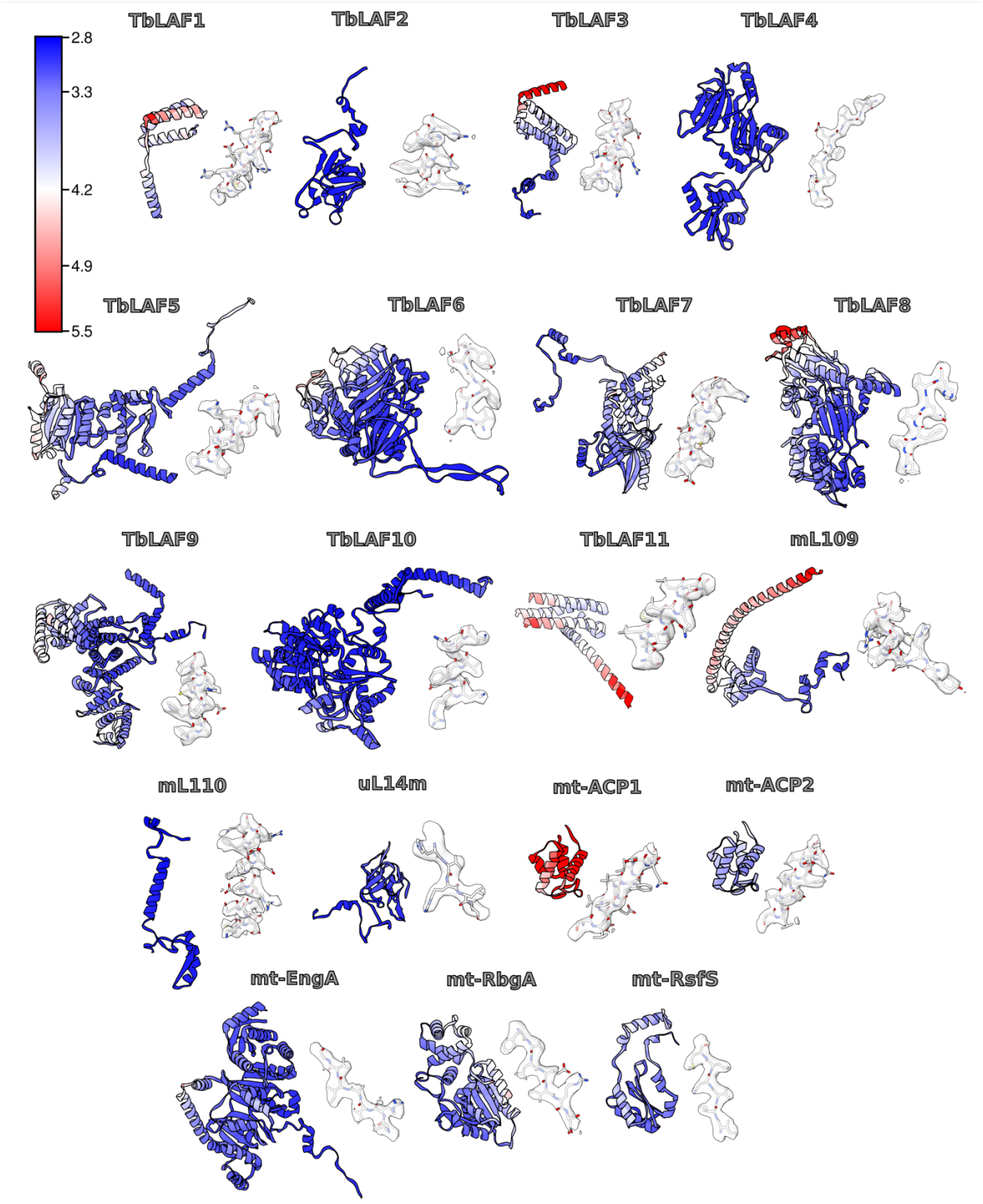
673 Zhang, X., Yan, K., Zhang, Y., Li, N., Ma, C., Li, Z., Zhang, Y., Feng, B., Liu, J., Sun, Y., Xu,
674 Y., Lei, J. & Gao, N. (2014) Structural insights into the function of a unique tandem GTPase
675 EngA in bacterial ribosome assembly, *Nucleic Acids Res.* 42, 13430-9

676 Zivanov, J., Nakane, T., Forsberg, B. O., Kimanius, D., Hagen, W. J., Lindahl, E., Scheres, S. H.
677 (2018) New tools for automated high-resolution cryo-EM structure determination in RELION-
678 3, *Elife* 7

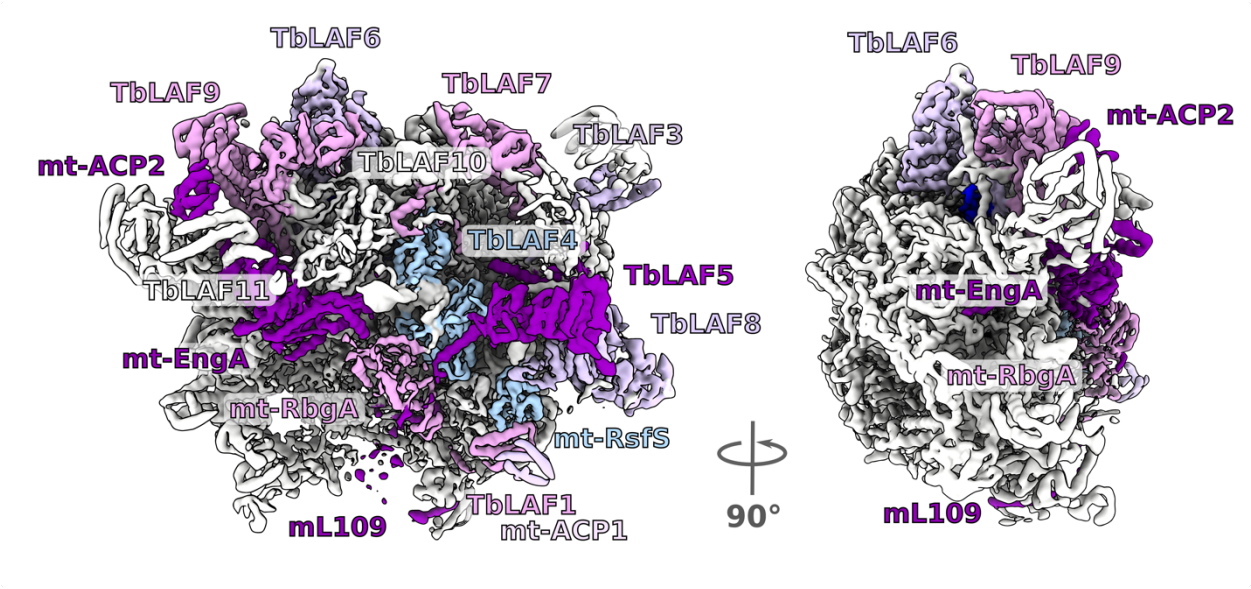
679 Zhu, J., King, M.S., Yu, M., Klipcan, L., Leslie, A.G. and Hirst, J. (2015) Structure of
680 subcomplex Iβ of mammalian respiratory complex I leads to new supernumerary subunit
681 assignments. *Proceedings of the National Academy of Sciences*, 112, 12087-12092



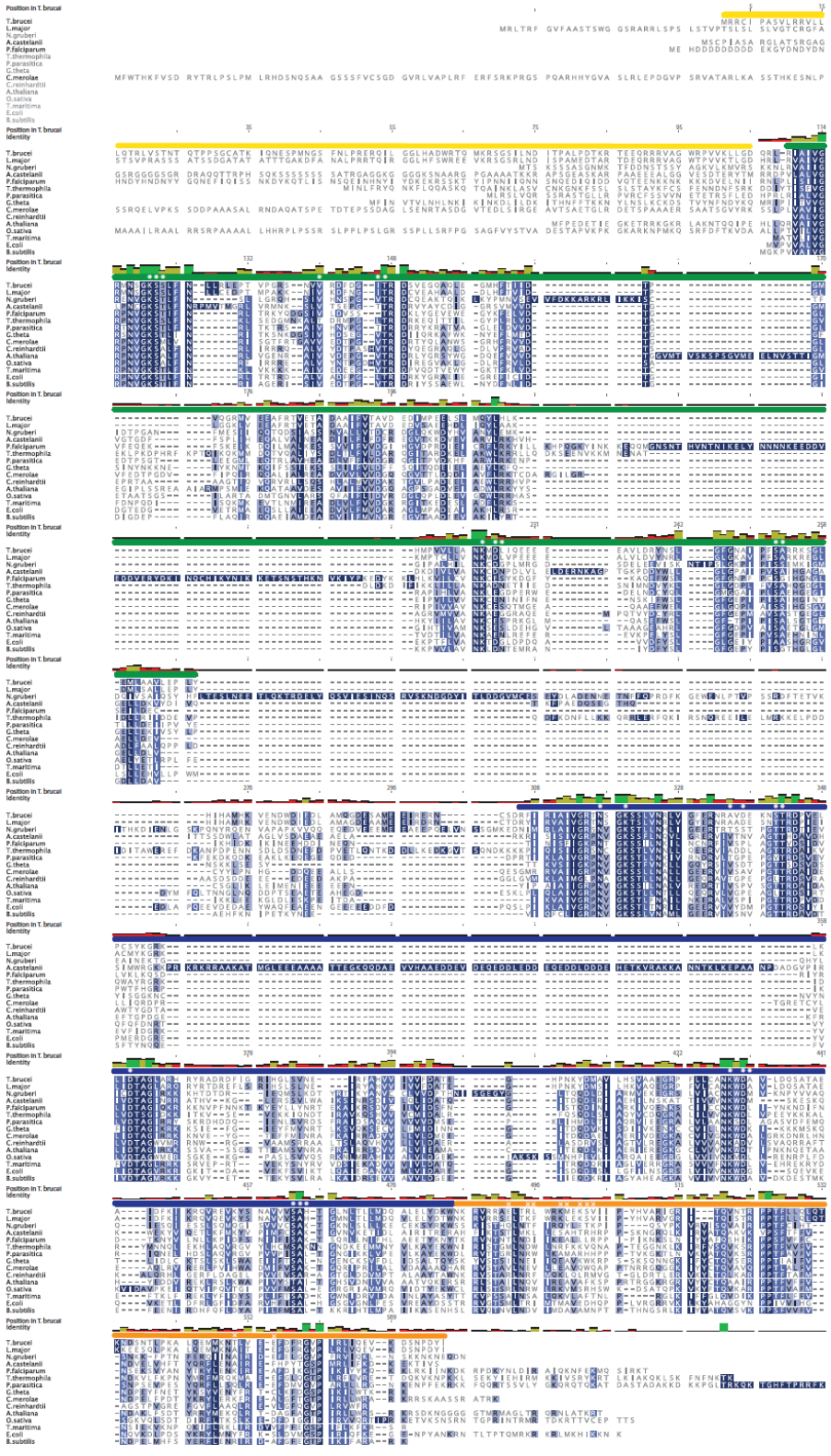
Appendix Figure S1. Cryo-EM data processing. (**A**) Representative micrograph. (**B**) Processing workflow. (**C**) Fourier shell correlation (FSC) curves. Resolution is estimated based on the 0.143 FSC cut-off criterion (red line).



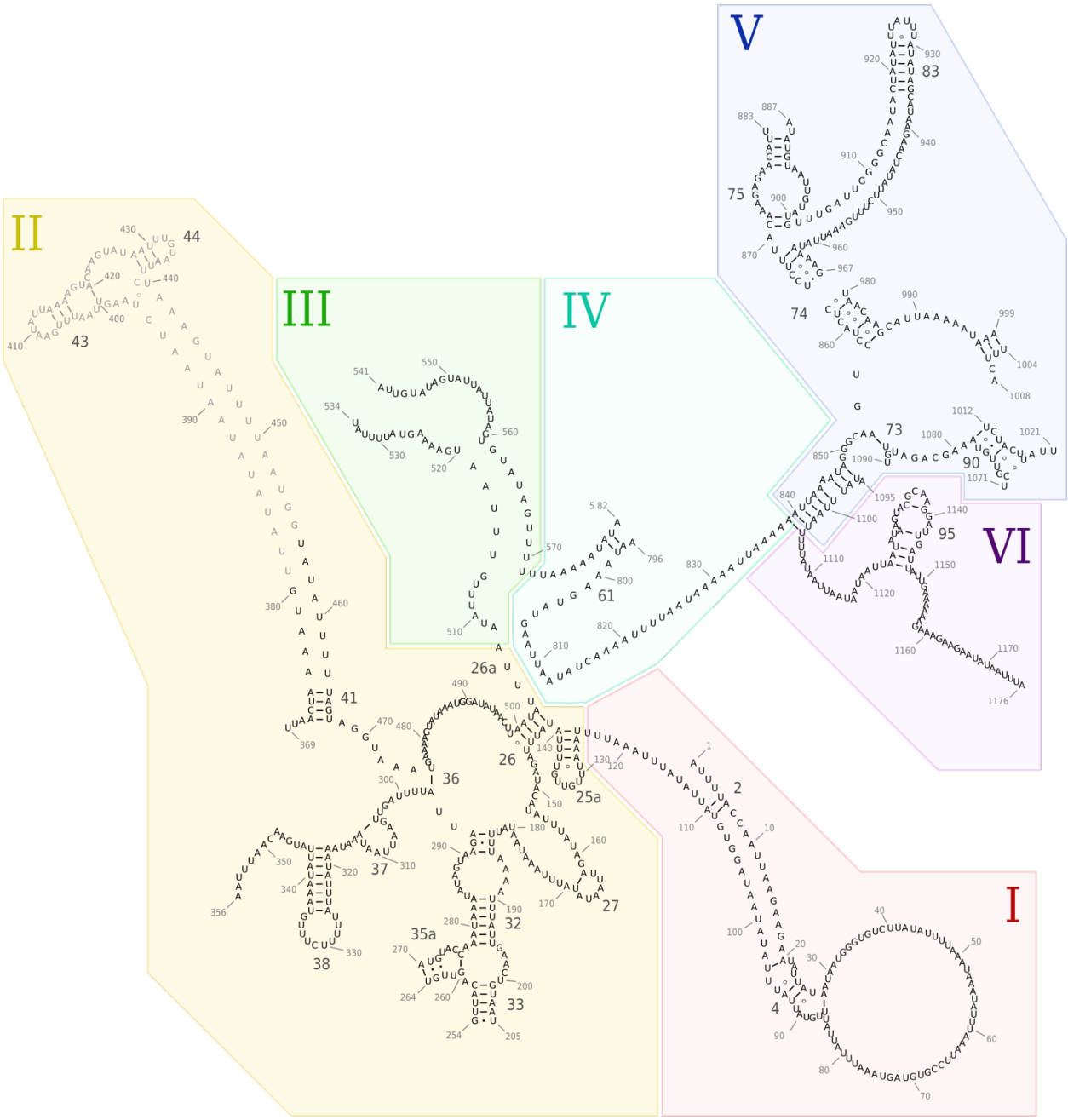
Appendix Figure S2. Examples of densities and models for individual assembly factors and newly identified proteins colored by local resolution of the density in the corresponding regions.



Appendix Figure S3. Cryo-EM density map of the pre-mtLSU showing distribution of the assembly factors.



Appendix Figure S4. Sequence alignment of EngA homologs from representative bacterial and eukaryotic species. The yellow, green, blue and orange horizontal bars mark the N-terminal extension, GTPase domain (GD) 1, GD2, and the KH domain, respectively. The white asterisks and crosses mark side chains in *T. brucei* mt-RbgA that coordinate GTP and interact with mt-RbgA, respectively. The green, yellow, and red vertical bars above the alignment correspond to 100%, <100% and ≥30%, and <30% identities at the respective position.



Appendix Figure S5. Secondary structure rRNA diagram derived from the model and colored by domain. Unmodeled sections that appear in the mature mtLSU are shown in gray. Domains in Roman numerals.

Supplementary File 1. Cryo-EM data collection, refinement and validation statistics

	Consensus map
Data collection and processing	
Magnification	130000x
Voltage (kV)	300

Electron exposure (e–/Å^2)	35
Defocus range (μm)	-0.8 ~ -3.5
Pixel size (Å)	1.05
Symmetry imposed	C1
Initial particle images (no.)	896,263
Final particle images (no.)	32,339
Map resolution (Å) FSC threshold 0.143	3.50
Map resolution range (Å)	3.0~10

Refinement	
Map sharpening *B* factor (Å^2)	-70
Model composition	
Non-hydrogen atoms	146831
Protein residues	17415
Ligands	10
B factors min/max/avg (Å^2)	
Protein	17/172/68
Nucleotide	22/281/57
Ligand	35/194/61
R.m.s. deviations	
Bond lengths (Å)	0.002
Bond angles (°)	0.46
Validation	
MolProbity score	1.65
Clashscore	6.7
Poor rotamers (%)	0.35
Ramachandran plot	
Favored (%)	95.6
Allowed (%)	4.10
Disallowed (%)	0.03

Supplementary File 2. Summary of pre-mtLSU components

Alias	Chain ID	TriTrypDB Gene ID (Lister strain 427)	TriTrypDB Gene ID (reference strain TREU927)	Uniprot ID (reference strain TREU927)	Full size	Modeled residues	Comment
12S rRNA	AA	rRNA	rRNA	N/A	1176	1-205, 254-264, 270-356, 369-380, 404-413, 445-450, 456-534, 541-582, 591-594, 796-883, 887-967, 980-999, 1004-1008, 1012-1021, 1071-1090, 1095-1176	
uL3m	AE	Tb427.03.5610	Tb927.3.5610	Q580R4	473	38-265, 272-404	
uL4m	AF	Tb427tmp.02.3810	Tb927.11.6000	Q385G8	351	18-459	
bL9m	AI	Tb427.05.3410	Tb927.5.3410	Q57UC5	263	9-220	
uL11m	AK	Tb427.02.4740	Tb927.2.4740	N/A	342	26-200, 207-235, 239-306	239-306 built as UNK
uL13m	AN	Tb427.04.1070	Tb927.4.1070	Q580D5	202	10-180	
uL14m	XG	Tb427.04.930	Tb927.4.930	Q580C1	217	20-107, 114-189	
uL15m	AP	Tb427.05.3980	Tb927.5.3980	Q57U68	374	10-136, 150-322, 354-363	
bL17m	AR	Tb427.08.5860	Tb927.8.5860	Q57YI7	301	11-266	
bL19m	AT	Tb427.01.1210	Tb927.1.1210	Q4GZ98	144	2-139	
bL20m	AU	Tb427tmp.01.1930	Tb927.11.10170	Q383R2	213	10-140, 162-205	
bL21m	AV	Tb427.07.4140	Tb927.7.4140	Q57UP4	188	6-185	
uL22m	AW	Tb427.07.2760	Tb927.7.2760	Q57Y86	278	2-278	
uL23m	AX	Tb427tmp.03.0260	Tb927.11.870	Q387G3	246	64-228	
uL24m	AY	Tb427.03.1710	Tb927.3.1710	Q57ZE0	378	1-311, 318-340	
bL28m	A1	Tb427.06.4040	Tb927.6.4040	Q586A2	241	10-226	
uL29m	A2	Tb427tmp.160.5240	Tb927.9.7170	Q38EM7	471	9-233, 248-471	
uL30m	A3	Tb427tmp.211.0230	Tb927.9.8290	Q38ED8	218	51-200	
bL32m	A5	Tb427.04.2330	Tb927.4.2330	Q584F4	80	26-80	2Fe-2S cluster binding
bL35m	A8	Tb427.10.1870	Tb927.10.1870	Q38C55	181	40-181	
mL41	Ae	Tb427tmp.01.1600	Tb927.11.9830	Q383U6	197	47-161	

mL42	Af	Tb427tmp.01.1840	Tb927.11.10080	Q383S1	189	41-173	
mL43	Ag	Tb427.04.4600	Tb927.4.4600	Q583E5	260	2-186	
mL49	Al	Tb427.05.3110	Tb927.5.3110	Q57Z82	218	37-101, 114-218	
mL52	Ao	Tb427tmp.02.2250	Tb927.11.4650	Q385V2	152	19-151	
mL53	Ap	Tb427.07.2990	Tb927.7.2990	N/A	309	16-303	
mL63	At	Tb427.07.7010	Tb927.7.7010	Q57XS1	154	10-154	
mL64	Av	Tb427tmp.01.3500	Tb927.11.11630	Q383B7	242	27-222	NAD binding
mL67	BA	Tb427tmp.55.0016	Tb927.11.1630	Q386Z1	831	27-83, 130-328, 335-542, 562-824	
mL68	BB	Tb427.10.600	Tb927.10.600	Q38CI0	541	62-258, 264-294, 304-341, 346-450	264-294, 304-341 built as UNK
mL70	BD	Tb427.06.4200	Tb927.6.4200	Q586Y7	547	105-521	
mL71	BE	Tb427.07.3460	Tb927.7.3460	Q57WG1	449	11-190, 228-448	
mL72	BF	Tb427.06.3930	Tb927.6.3930	Q585Z1	426	26–62, 118-421	
mL74	BH	Tb427.10.7380	Tb927.10.7380	Q38AM5	349	89-314	
mL75	BI	Tb427.10.380	Tb927.10.380	Q38CK0	342	20-342	
mL76	BJ	Tb427tmp.01.2340	Tb927.11.10570	Q383M2	333	173-333	
mL77	BK	Tb427.06.2480	Tb927.6.2480	Q584Q8	386	84-156, 188-233, 254-269, 280-386	254-259 built as UNK
mL78	BL	Tb427.10.11050	Tb927.10.11050	Q389N4	312	31-130, 141-197, 216-265, 281-306	
mL80	BN	Tb427.06.1440	Tb927.6.1440	Q585A3	302	53-266	
mL81	BO	Tb427tmp.02.3230	Tb927.11.5530	Q385L5	262	36-193, 210-262	
mL83	BQ	Tb427.07.3430	Tb927.7.3430	Q57WF8	231	16-200	
mL84	BR	Tb427.06.4080	Tb927.6.4080	Q586A6	205	11-205	
mL85	BS	Tb427tmp.160.2250	Tb927.9.3640	Q38FG8	198	20-163	
mL86	BT	Tb427.05.4120	Tb927.5.4120	Q57Z37	191	10-176	

mL87	BU	Tb427tmp.01.0500	Tb927.11.8040	Q384L5	185	104-185	
mL89	BW	Tb427.03.820	Tb927.3.820	Q57WW5	188	2-188	
mL90	BX	Tb427.06.1700	Tb927.6.1700	Q585P1	190	61-100, 108-174	2x Zn binding
mL92	BZ	Tb427tmp.01.1215	Tb927.11.9450	Q383Y4	190	2-190	
mL93	Ba	Tb427.10.11350	Tb927.10.11350	Q389K5	153	19-153	
mL94	Bb	Tb427tmp.160.5050	Tb927.9.6910	Q38EP7	162	38-140	
mL95	Bc	Tb427.10.11370	Tb927.10.11370	Q389K3	146	10-146	
mL98	Bf	Tb427.10.13770	Tb927.10.13770	Q388M2	113	27-68, 75-112	
mL99	Bg	Tb427.02.2590	Tb927.2.2590	Q587H8	105	24-105	
mL100	Bh	N/A	Tb927.9.8905	N/A	92	2-91	Zn binding
mL109	XR	Tb427.01.1390	Tb927.1.1390	Q4GZ80	245	22-203	
mL110	XS	Tb427tmp.01.1810	Tb927.11.10050	Q383S4	102	6-102	
mt-ACP1	XD	Tb427.03.860	Tb927.3.860	Q57WW9	148	64-146	
mt-ACP2	XE	Tb427.03.860	Tb927.3.860	Q57WW9	148	64-146	
mt-EngA	XL	Tb427.07.1640	Tb927.7.1640	Q57TZ4	576	45-504, 514-574	2xGTP binding
mt-RbgA	XQ	Tb427tmp.211.0810	Tb927.9.9150	Q38E75	451	44-202,215-338	
mt-RsfS	XJ	Tb427.06.3420	Tb927.6.3420	Q584Y2	349	163-312	
TbLAF1	XM	Tb427.07.4210	Tb927.7.4210	Q57UQ1	116	25-115	
TbLAF2	XA	Tb427.10.15860	Tb927.10.15860	Q387S8	156	3-156	Zn binding
TbLAF3	XF	Tb427.04.4610	Tb927.4.4610	Q583E6	319	120-170, 194-245, 260-317	
TbLAF4	XP	Tb427tmp.160.2000	Tb927.9.3350	Q38FJ3	406	35-405	
TbLAF5	XO	Tb427tmp.211.3800	Tb927.9.12850	Q38DC9	586	78-344, 363-440, 453-505	
TbLAF6	XC	Tb427tmp.02.3800	Tb927.11.5990	Q385G9	616	1-414, 444-616	

TbLAF7	XH	Tb427.05.2070	Tb927.5.2070	Q57ZS6	634	2-55, 86-113, 184-276, 300-336, 367-412, 471-577, 587-623	
TbLAF8	XN	Tb427.08.3300	Tb927.8.3300	Q57YY3	691	58-101, 123-150, 200-668	
TbLAF9	XI	Tb427.05.3870	Tb927.5.3870	Q57U79	731	25-95, 156-203, 213-289, 319-647, 656-727	
TbLAF10	XB	Tb427tmp.52.0011	Tb927.11.12930	N/A	754	47-391, 443-711	MgADP binding
TbLAF11	XR	Tb427tmp.211.4580	Tb927.9.14050	Q38D50	524	2-165	

Supplementary File 3. Contacts of assembly factors with rRNA. Regions and nucleotides of respective rRNA domains are color-coded as in Fig 4 and Fig EV4.

Assembly factor	Contacts with rRNA	
	Region	Nucleotides
mt-EngA	H74, H75, H88, H90-91	868-874, 891-898, 958-967, 994-999, 1004
mt-RbgA	H90-91	1021, 1071
mt-RsfS	H95	1126-1132, 1138-1141
TbLAF2	H39, H43, H72, H74, H88, H93	132-135 347-351, 404, 821-829, 832-835, 860-862, 904, 965-677, 985-992, 1079-1082, 1119
TbLAF3	H43	411-412, 446
TbLAF4	H39, H90-93	351-353, 857-859, 992-994, 1008, 1012-1021, 1071-1087, 1122-1124
TbLAF5	H43, H90, H95	409-413, 1020-1021, 1136-1137
TbLAF6	H13, H28, H37, H80-81, H88	65-66, 164-166, 313, 328, 908, 917-919, 935-939, 944-953
TbLAF7	H41, H43	407, 448-450
TbLAF8	H95	1131-1140
TbLAF9	H80-81, H88	908-909, 915-916, 919-924, 930-940, 952-957, 961
TbLAF10	H26, H32, H33, H35a, H37, H39, H51, H72, H73-75, H80-81, H93	127-139, 199-200, 270-276, 289-293, 306-316, 344-355, 497, 550, 555, 826-828, 847-848, 850-855, 860-862, 870, 903-922, 930-933, 947-980, 955-965, 983 - 992, 1085-1089