https://doi.org/10.1093/nar/gkad422

Miniature RNAs are embedded in an exceptionally protein-rich mitoribosome via an elaborate assembly pathway

Matus Valach [1,*,†], Corinna Benz [2,*,†], Lisbeth C. Aguilar [3], Ondřej Gahura [2], Drahomíra Faktorová [2,4], Alena Zíková [2,4], Marlene Oeffinger [1,3,5], Gertraud Burger [1], Michael W. Gray [6] and Julius Lukeš [2,4]

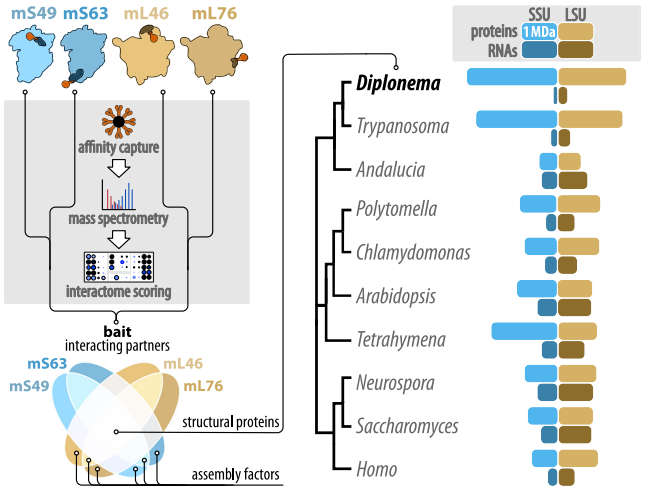
[1]Department of Biochemistry and Robert-Cedergren Centre for Bioinformatics and Genomics, Université de Montréal, Montréal, Quebec, Canada, [2]Institute of Parasitology, Biology Centre, Czech Academy of Sciences, České Budějovice (Budweis), Czech Republic, [3]Center for Genetic and Neurological Diseases, Institut de recherches cliniques de Montréal (IRCM), Montréal, Quebec, Canada, [4]Faculty of Sciences, University of South Bohemia, České Budějovice (Budweis), Czech Republic, [5]Division of Experimental Medicine, McGill University, Montréal, Quebec, Canada and [6]Department of Biochemistry and Molecular Biology and Institute of Comparative Genomics, Dalhousie University, Halifax, Nova Scotia, Canada

Received March 01, 2023; Revised April 20, 2023; Editorial Decision May 04, 2023; Accepted May 08, 2023

ABSTRACT

The mitochondrial ribosome (mitoribosome) has diverged drastically from its evolutionary progenitor, the bacterial ribosome. Structural and compositional diversity is particularly striking in the phylum Euglenozoa, with an extraordinary protein gain in the mitoribosome of kinetoplastid protists. Here we report an even more complex mitoribosome in diplonemids, the sister-group of kinetoplastids. Affinity pull-down of mitoribosomal complexes from *Diplonema papillatum*, the diplonemid type species, demonstrates that they have a mass of > 5 MDa, contain as many as 130 integral proteins, and exhibit a protein-to-RNA ratio of 11:1. This unusual composition reflects unprecedented structural reduction of ribosomal RNAs, increased size of canonical mitoribosomal proteins, and accretion of three dozen lineage-specific components. In addition, we identified >50 candidate assembly factors, around half of which contribute to early mitoribosome maturation steps. Because little is known about early assembly stages even in model organisms, our investigation of the diplonemid mitoribosome illuminates this process. Together, our results provide a foundation for understanding how runaway evolutionary divergence shapes both biogenesis and function of a complex molecular machine.

GRAPHICAL ABSTRACT



INTRODUCTION

Mitochondria play pivotal roles in energy conversion and ATP production. They are also important for the synthesis of iron sulphur clusters and heme, and are involved in many other cellular functions such as ion homeostasis and apoptosis (1–3). This organelle has diverged dramatically from its early origin as an endosymbiotic alphaproteobacterium to its present-day state as exemplified by the progressive loss of coding capacity of the mitochondrial genome, with some genes transferred to the nucleus and others lost entirely (4).

*To whom correspondence should be addressed. Tel: +1 514 343 6111 (Ext 5172); Fax: +1 514 343 2210; Email: matus.a.valach@gmail.com
Correspondence may also be addressed to Corinna Benz. Tel: +420 387 77 5445; Fax: +420 38 531 03882; Email: corinna.benz@paru.cas.cz
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

In extant eukaryotes, the number of protein-coding genes in mitochondrial DNA (mtDNA) varies by a factor of 30, ranging from 66 in the jakobid *Andalucia godoyi* (5) to a mere two in the alveolate *Chromera velia* (6), albeit more commonly leaning towards the lower end of these extremes. For translation of mRNAs specifying proteins encoded by mtDNA, a dedicated mitochondrial ribosome (mitoribosome) is employed, the essential RNA components of which have been retained in the mitochondrial genome in all organisms studied to date.

Like their bacterial counterparts, mitoribosomes consist of two subunits, the small (SSU) and large (LSU) subunits, which in turn are composed of a catalytic core, the mitochondrial ribosomal RNAs (mt-rRNAs), wrapped in proteins (7). While the SSU binds mRNAs and decodes genetic information, the LSU catalyzes polypeptide chain synthesis. The mitoribosome of the last eukaryotic common ancestor (LECA) is thought to have consisted of a core set of mitoribosomal proteins (mtRPs) of bacterial origin but also a number of subsequently acquired components (8). Interestingly, most mtRPs are encoded in the nucleus even in *A. godoyi*, the organism with the highest number of genes retained in mitochondria (9).

In recent years, mitoribosome structure has been resolved by cryo-electron microscopy (cryo-EM) at near-atomic level resolution from representatives of several eukaryotic supergroups, namely animals (*Homo sapiens*, *Sus scrofa*), fungi (*Saccharomyces cerevisiae*, *Neurospora crassa*), land plants (*Arabidopsis thaliana*, *Brassica oleracea*), green algae (*Chlamydomonas reinhardtii*, *Polytomella magna*), ciliates (*Tetrahymena pyriformis*), and kinetoplastids (*Trypanosoma brucei*, *Leishmania tarentolae*) (10–18). These data illustrate that eukaryotic evolution has resulted in an extraordinary diversity of present-day mitoribosomes, which by far exceeds that of cytosolic ribosomes (19). Due to conspicuous extensions of the mtRPs of bacterial origin, emergence of mitochondria-specific mtRPs, and lineage-specific acquisitions of new proteins, mitoribosomes are overall much larger than their bacterial counterparts (7). The most protein-rich mitoribosome reported to date is that of *T. brucei*, with as many as 122 proteins (compared to 54 and 80 in ribosomes of *Escherichia coli* and yeast mitochondria, respectively), a molecular weight of >4.5 MDa (twice as large as is typical in bacteria), and a protein-to-RNA ratio of ∼8:1 (compared to ∼1:2 in bacteria) (13).

Because the protein gain is often accompanied by the truncation of the rRNA components, it was postulated earlier that the loss of rRNA structural elements has been secondarily compensated for by the acquisition of additional proteins (20). This view, however, has recently been refuted because the mitoribosome of jakobid protists has the most bacteria-like mt-rRNAs among eukaryotes, but at the same time contains almost all of the proteins that were assumed to fulfill stabilizing roles in mitoribosomes with truncated mt-rRNAs (21). Another hypothesis proposes that considerable truncation of SSU and LSU mt-rRNAs and complete loss of the 5S mt-rRNA were facilitated by an increase of taxon-specific mtRPs (e.g. in trypanosomes) (13). (Note that a mtDNA-encoded 5S rRNA has otherwise been retained in plants, red algae, and numerous protist lineages (22)). Interestingly, 5S mt-rRNA has also been lost from mammals and fungi but functionally replaced by a tRNA in the former lineage (23) and by LSU rRNA expansion segments in fungi (24).

In some organisms, mt-rRNA sequences have diverged to extreme degrees. For example, in apicomplexans (25), ciliates (16,26) and green algae (18,27), mt-rRNA genes have undergone fragmentation, with sub-genic rDNA modules being expressed as multiple stand-alone RNA products that assemble and interact through intermolecular base-pairing. In other instances, reduced mt-rRNAs only contain a few short stretches of otherwise conserved sequence motifs and secondary structures. Prime examples of such an evolutionary reduction are the marine flagellates diplonemids, the sister clade of kinetoplastids. Diplonemids emerged only recently as one of the most abundant groups of protists in the ocean (28). Their mt-rRNAs are the shortest described to date, tallying as little as 363 and 908 nt for SSU and LSU rRNAs, respectively, which represents ∼24% and ∼31% of the average size of their bacterial (as well as 'average' mitochondrial) counterparts. In addition, the SSU mt-rRNA of diplonemids contains numerous inosine (I) residues, generated by RNA editing, not observed in any other rRNA to our knowledge (29–31). These traits raise numerous questions about the composition of diplonemid mitoribosomes. For example, are the diplonemid-specific deviations of mt-rRNA accompanied by accretion of new mtRPs and/or sequence extension of those mtRPs that have homologs in other organisms? Does the assembly of RNA and protein components into the mature form of this deviant mitoribosome require unique assembly and maturation factors?

In the diplonemid type species, *Diplonema papillatum*, all mtRPs are encoded by the nuclear genome (32). As this species has recently become tractable to genetic manipulation (33,34), we are now able to express tagged mtRPs and identify their interaction partners via affinity capture coupled to mass spectrometry. Here, we determined the mitoribosomal protein interaction networks from four *D. papillatum* cell lines, each genetically altered to express a distinct mtRP carrying a Protein-A (PrA) tag. Our approach has allowed us to purify not only mature mitoribosomes but also assembly intermediates, providing insights into both conserved and unique features of the diplonemid mitochondrial translation apparatus and its biogenesis.

MATERIALS AND METHODS

Cell cultivation

Diplonema papillatum (ATCC 50162) was originally obtained from the American Type Culture Collection (ATCC) and cultivated axenically without shaking at 15–22°C in ocean salt medium (OS) containing 33 g/l Instant Ocean Sea Salt (Instant Ocean) and supplemented with 1% (v/v) horse serum as described earlier (35). For extended cultivations, chloramphenicol (40 mg/l) was added to prevent bacterial contamination. For extraction of mitochondria and whole-cell preparations for rapid affinity purification, yeast extract was added to the medium to 0.04% (w/v). Alternatively, for tagged strain construction and immunoprecipitation-mass spectrometry (IP-MS) experiments, *D. papillatum* was cultured axenically at 27°C in sea water supplemented with 1% (v/v) fetal bovine serum and

0.1% (w/v) tryptone. For the selection and preservation of the tagged cell lines, the antibiotic G418 was added at 75–150 mg/l as necessary.

DNA and RNA extraction, PCR, and strain generation

DNA was isolated using a kit (GeneAll) or a phenol-chloroform extraction procedure. RNA was extracted with a 'home-made' Trizol substitute (36). RNA extraction from BN-PAGE gels was performed using the crush and soak method, essentially as described in (37). To generate cell lines expressing C-terminally PrA-tagged mitoribosomal subunits, we created a fusion PCR product of the PrA tag and neomycin resistance marker from plasmid pDP002 (33) flanked by ∼1.6 kb homologous regions from the respective mtRP ORF and 3′ UTR using Phusion polymerase (NEB). Primers are listed in Supplementary Table S1. All tagging cassettes were transformed as described earlier (33,34) with G418 selection ranging from 45 to 83 μg/ml. Following ∼2 weeks of selection, potential clones were transferred into fresh medium with 75 μg/ml G418, and expression of the tagged protein was confirmed by western blotting. Generation of the cell line expressing mitochondria-targeted PrA (mtPrA) was described previously (33).

Immunofluorescence microscopy

To localize PrA-tagged mtRPs by immunofluorescence, a 10 ml culture of logarithmically growing cells was pelleted and fixed in 2% PFA in PBS-sea water (1:1) for 15 min at room temperature. Following a wash in 1× PBS, cells were permeabilized in ice-cold methanol on ice for 10 min. Fixed cells were then rehydrated in PBS and blocked in 1% BSA in PBS for 1 h followed by incubation with primary antibody for 1 h on ice (anti-PrA, Sigma P3775, RRID: AB_261038, 1:2000). Following three washes with 1× PBS, samples were incubated in AlexaFluor 555-conjugated secondary antibody (goat anti-rabbit, Invitrogen A21428, RRID: AB_2535849, used at 1:1000). After three further washes in 1 × PBS, a drop of ProLong Gold Antifade reagent with DAPI (ThermoScientific) was added to the cell suspension and mixed. This mixture was then spread on a glass slide (ThermoScientific), a coverslip applied, and the slide sealed with nail polish. Slides were imaged with a Zeiss Axioscope and images overlayed using Gimp 2.10.8 software.

Mitochondrial sample preparation and sucrose gradient velocity separation

Mitochondria were extracted as described earlier (35), with minor variations (see Supplementary Information). A detailed protocol is accessible at https://doi.org/10.17504/protocols.io.pkydkxw. Mitochondrial enrichment was gauged by examining the content of mitochondrial rRNAs and activity of Complexes I and V by in-gel staining (35). Proteins from thawed mitochondria were quantified by the Bradford method. Mitochondrial pellets were mixed with the appropriate amount of buffer to obtain a homogenate containing at the final concentration ∼10 mg/ml total proteins, 20 or 100 mM KCl, 20 mM $MgCl_2$, 2.5 mM DTT, 30 mM Tris–HCl pH7.6 and 1× cOmplete EDTA-free (representing 1 × sucrose gradient buffer). The homogenate was immediately frozen in liquid nitrogen and an equivalent of ∼1.5 g was pulverized using TissueLyserII (Qiagen) in a 10-ml grinding jar with a 10-mm steel ball, twice for 30 s at 30 Hz with intermittent cooling in liquid nitrogen. The fine grindate was stored at −70°C until further use. Pulverized mitochondria were lysed by the addition of 1% Triton X-100. After homogenization and a 10-min incubation on ice, the mixture was centrifuged (25 000 × *g*, 10 min, 4°C) and the resulting supernatant was concentrated to a volume of ∼750 μl on a 30-kDa Vivaspin 500 (Sartorius) ultrafiltration device (12 000 × *g*, 10–15 min, 4°C) following the manufacturer's instructions. Around 1/20 (∼500 μg proteins) was set aside to be later used as the 'input' reference, while the remaining 19/20 were loaded on top of a 5-ml 10–30% sucrose gradient (containing 1 × sucrose gradient buffer and 0.02% Triton X-100) and centrifuged in an AH-650 swinging-bucket rotor (247 500 × *g*, 130 min, 4°C). After centrifugation, gradient fractions of 210 μl were collected from the top, snap-frozen in liquid nitrogen, and stored at −70°C. Mitoribosome migration was monitored by examining the profile of extracted RNA by agarose gel electrophoresis, as well as of tagged proteins by western blotting. Ultracentrifugation conditions were optimized in preliminary experiments with wild-type mitoribosomes before being applied to cell lines expressing PrA-tagged mtRPs.

Western blotting

About 10^6 cells were resuspended in 20 μl of 2× Laemmli buffer and boiled at 95°C for 5 min. Whole-cell protein samples were run on 12% Tris–glycine SDS-PAGE gels. For sucrose gradient analyses, equal-volume aliquots of gradient fractions were run on 10% Tris–tricine SDS-PAGE gels essentially as described (38). Proteins were then blotted onto PVDF membranes (Amersham Hybond PDVF 0.45 μm, GE Healthcare). Following a blocking step in 5% fat-free milk in 1× PBS (or 1× TBST) for 1 h, blots were incubated in primary antibody overnight at 4°C (anti-PrA, Sigma P3775, RRID: AB_261038, used at 1:20 000; anti-α-tubulin, Sigma T9026, RRID: AB_477593, used at 1:5 000). After several washes in 1× PBS (or 1× TBST), blots were incubated in secondary antibody (anti-rabbit-HRP, Sigma 12-348, RRID: AB_390191; anti-rabbit-HRP, Cell Signaling Technology 7074, RRID: AB_2099233; anti-mouse-HRP, Sigma 12-349, RRID: AB_390192; used at 1:2 000) for 1 h at room temperature. The signal was then visualized using Clarity Western ECL substrate (Bio-Rad) or SignalFire ECL reagent (Cell Signaling Technology) on a ChemiDoc MP instrument (Bio-Rad).

Protein A-mediated pulldowns

Single-step affinity purification (AP) was performed in quadruplicate as previously described for humans (39), with slight modifications. The relative abundance of the PrA-tagged bait proteins in the cleared lysates was determined by western blot. The intensity signals were used to determine the volumes of cleared lysate to use for the affinity purifications, such as to saturate 3.75 mg of Dynabeads M-270

Epoxy conjugated in advance with rabbit IgG (Sigma I5006, RRID: AB_1163659). *Diplonema* cell pellets were frozen in liquid nitrogen and ground by cryo-milling (40). Cryogenically preserved cell grindates were weighed out in 100 mg aliquots, which were resuspended in 500 μl of the AP buffer (30 mM Tris–HCl pH7.6, 20 mM KCl, 25 mM $MgCl_2$) containing 1% Triton X-100 and supplemented with protease inhibitor cocktail (Mini-complete EDTA-free, Roche). After vortexing, the cell powder was further solubilized by a brief sonication (20 W, 2 s) and the lysate was cleared by centrifugation (16 100 × *g*, 4°C, 10 min). The pre-determined volume of cleared lysate was topped to 900 μl with the AP buffer and the detergent concentration adjusted to the final 0.1%. The lysates were incubated with 3.75 mg IgG-conjugated Dynabeads M-270 (pre-washed in the AP buffer with 0.1% Triton X-100) for 30 min at 4°C on a rotating wheel. After six washes with the AP buffer with 0.1% detergent, the beads were washed once in the AP buffer with 0.01% detergent for 5 min with agitation; and then four times (three times quickly, once for 5 min) in 30 mM Tris–HCl pH7.6 and 5 mM $MgCl_2$. This was followed by on-bead tryptic digestion (see below). Immunoprecipitation (IP) experiments were performed in triplicate essentially as described (41) (for details, see Supplementary Information).

Mass spectrometry

Overnight on-bead digestion of affinity-purified (AP) complexes was performed using 500 ng trypsin in 20 mM Tris–HCl pH8.0 at 37°C. The digestion was stopped by adding formic acid to a final concentration of 2%. Tryptic peptides were dried (speedvac) and stored at −80°C. Prior to LC–MS/MS, protein digests were re-solubilized under agitation for 15 min in 10 μl of 0.2% formic acid. Desalting/cleanup of the digests was performed by using C18 ZipTip pipette tips (Millipore). Eluates were dried down in vacuum centrifuge and then re-solubilized under agitation for 15 min in 10 μl of 1% ACN and 1% formic acid. Samples were loaded into a 75 μm i.d. × 150 mm Self-Pack C18 column installed in the Easy-nLC 1200 system (Proxeon Biosystems). The buffers used for chromatography were 0.2% formic acid (buffer A) and 100% acetonitrile with 0.2% formic acid (buffer B). Peptides were eluted with a two-slope gradient at a flow-rate of 250 nl/min. Solvent B first increased from 1 to 34% in 74 min and then from 34 to 93% in 8 min. The HPLC system was coupled to Orbitrap Fusion mass spectrometer (Thermo Scientific) through a Nanospray Flex Ion Source. Nanospray and S-lens voltages were set to 1.4 kV and 60 V, respectively. Capillary temperature was set to 250°C. Full scan MS survey spectra (m/z 360–1550) in profile mode were acquired in the Orbitrap with a resolution of 120 000 and a target value of 4e5. The 20 most intense peptide ions were fragmented in the HCD collision cell and analyzed in the linear ion trap with a target value of 1e4 and a normalized collision energy at 30 V. MS3 scanning was performed upon detection of a neutral loss of phosphoric acid (48.99, 32.66 or 24.5 Th) in MS2 scans. Target ions selected for fragmentation were dynamically excluded for 25 s after two MS2 events. Data acquisition relied on Xcalibur v4.0 and Tune v2.0. Mass spectrometry of immunoprecipitated (IP) samples and wild-type mitoribosomes was outsourced (for details, see Supplementary Information). The list of all analyzed proteomics samples and datasets is provided in Supplementary Table S2.

Analyses of mass spectrometry and interactome data

Prior to data analysis, the Thermo RAW format was converted to mzML using ThermoRawFileParser v1.3.2 (42). Peptide searches in the raw MS/MS datasets were performed using MSFragger v3.3 (43), followed by filtering and scoring by *Philosopher* v4.0.0 (44) and quantification by *IonQuant* v1.7.5 (45). For details, including the *Philosopher* pipeline configuration file, see Supplementary Information. Precursor intensity-based protein quantification was done essentially as described in earlier publications (21,35) by calculating iBAQ (intensity-Based Absolute Quantification) values (46). To determine which proteins were enriched in the samples, the *Philosopher* pipeline was instructed to produce Reprint-type reports, which were then used to score protein-protein interactions (PPIs) using *SAINTexpress* (47) as implemented at the *CRAPome* database and analysis website (48). PPIs were then inspected using the *ProHits-viz* tool suite (49). Lastly, the iBAQ values and *SAINTexpress* scores were combined for the purposes of visualization and categorization of identified PPI candidates. Thresholds for new mtRP and mtAF candidates were based on the distribution of assigned mtRPs and other factors, respectively, in individual samples of the AP pulldowns, because overall this approach displayed a much better replicate correlation (∼0.66 for MRPS49 and MRPS63, ∼0.62 for MRPL46, and ∼0.54 for MRPL76 and mtPrA) than IP pulldowns. For the protein categorization, any candidate had to satisfy the thresholds in both the MRPS49 and MRPS63 pulldowns to be classified as an mtSSU interacting partner. Because of the reduced resolution of the MRPL76 dataset, we applied thresholds only to MRPL46; however, for a candidate to be considered authentic, it had to be at least detected in the MRPL76 dataset. More specifically, for a protein to classify as an mtRP candidate, it was required to have a SAINT probability score (SP) >0.98, primary fold-change score (FC_A) > 4, and $\log_2$_iBAQ > 24.5 (MRPS49, MRPS63) or > 23 (MRPL46). To classify a prey as a subunit-specific mtAF candidate, the thresholds were set at SP > 0.98, FC_A > 1.2 and $\log_2$_iBAQ > 19, under the condition that the prey protein was detected in pulldowns of both baits from the same subunit. Proteins observed only in one of the bait pulldowns were excluded from the latter category and were classified as 'bait-specific' mtAF candidates (with FC_A >2 and >1.5 for MRPS49/MRPS63 and MRPL46/MRPL76, respectively, and with SP > 0.98 and $\log_2$_iBAQ > 20 for all baits). The cut-off values were determined based on the values of assigned mtRPs and mtAFs. Additional filtering of candidate mtRPs and mtAFs was done based on a protein's levels in whole cell lysates, cytosol, and enriched mitochondria of *D. papillatum* using the dataset PXD035104 available in the PRIDE database. If detected as enriched in the cytosol compared to whole cells, the protein was considered an evident contaminant. Levels of assigned mtRPs fluctuated relatively widely in the mitochondrial fraction (Supplementary Table S3), but 95% values were below $\log_2$_iBAQ = 19.2. Therefore, any prey

protein detected in mitochondria above this threshold was considered to have been observed in the pulldowns due to spurious interactions.

Analyses of protein and RNA sequences and structure modeling

Sequence analyses (including local sequence similarity searches, domain identification, multiple sequence alignment, phylogenetic reconstruction, RNA secondary structure prediction, protein structure visualization and analysis) followed the procedures established previously (21), using tools and databases listed below. Most assignments of *Diplonema*'s proteins were confidently achieved via sequence comparison, except for a few instances of particularly divergent sequences (for details, see Supplementary Information; Supplementary Figure S1). In these cases, we compared structural predictions created by AlphaFold (50) to the experimentally determined structures of mtRPs from various organisms. To improve the prediction reliability of AlphaFold as described in (51), we expanded the underlying Uniprot database with the EukProt database (52) containing protein sequences from > 1 000 eukaryotes underrepresented or missing from public databases. Tools used: HMMER (53), ClustalOmega (54), MAFFT (55), WebLogo3 (56), FastTree2 (57), IQ-TREE v2.2.0 (58), Vienna RNA package 2.0 (59), RNAstructure v6.4 (60), pdb-tools v2.4.8 (61), ChimeraX v1.4 (62), AlphaFold v2.1.2 & v2.2.4 (50). Databases used: SMART (63), Pfam (64), NCBI CDD (65), Uniprot (66), PDBe (67).

RESULTS

Isolation of mitoribosomal subunits from *Diplonema*

To allow immunoprecipitation and affinity purification of mitoribosomes, we transformed *Diplonema* with plasmid constructs carrying the 3′-terminal moiety of an mtRP gene joined to a Protein-A (PrA) tag. Plasmid design took the following criteria into account. For tagging, we chose mtRPs whose homologs in the *Trypanosoma* mitoribosome are located at the surface and have solute-exposed C-termini (13). To minimize non-homologous integration of the construct, mtRP genes were chosen that were not bounded by repetitive sequences within 2 kb of the genomic regions upstream and downstream of the stop codon. Four cell lines were generated each expressing a distinct PrA-tagged mtRP: mS49 (gene name *MRPS49*; gene identifier *DIPPA_31280*), mS63 (*MRPS63*; *DIPPA_25715*), mL46 (*MRPL46*; *DIPPA_18631*) or mL76 (*MRPL76*; *DIPPA_24761*). Extrapolating from the *Trypanosoma* mitoribosome structure (13), the mS49 and mS63 homologs in *Diplonema* are located in the mtSSU subunit head and tip of the body, respectively, mL46 is positioned at the top of the mtLSU central protuberance (CP), and mL76 bulges out on the mtLSU intersubunit side, below the L1 stalk (Figure 1A, Supplementary Figure S2). As a control, we also generated a cell line with an ectopically expressed mitochondrially targeted PrA (mtPrA, a mitochondrial presequence joined to PrA), which integrated in a non-coding region between the genes *DIPPA_21441* and *DIPPA_21439* (for details, see Materials and Methods).

Genomic PCR and western blot analysis with anti-PrA antibodies together showed that each *Diplonema* cell line carrying a tagged mtRP expressed the fusion protein but lacked the original, unmodified gene (Supplementary Figure S3A,B). The expression levels of the tagged mS49, mS63 and mL46 proteins were similar, whereas mL76-PrA was ∼90% less abundant. Mitochondrial localization of the fusion proteins was verified by immunofluorescence microscopy using anti-PrA antibodies. Concurrent staining with the DNA-binding dye DAPI showed that the tagged proteins coincided with the location of mtDNA (Figure 1B), displaying a pattern consistent with a reticulated, subcortical mitochondrion typical for diplonemids (32,68,69).

The molecular mass of tagged mitoribosomes was examined by their sedimentation behaviour during sucrose gradient ultracentrifugation using mtSSU and mtLSU rRNAs (374 and 930 nt, respectively, in *D. papillatum*) as mitoribosomal markers. Both markers peaked in the fractions corresponding to ∼40–50S (∼2–3 MDa) followed by a gradual decrease (Figure 1C, Supplementary Figure S4A–C, E, F). Particles of that size are characteristic of individual subunits in kinetoplastids (13,70), which indicates that the *Diplonema* mitoribosome dissociates into subunits of similar sizes.

To test whether the tagged *Diplonema* mtRPs were incorporated into mitoribosomal particles, mitochondrial extracts from the four cell lines were sedimented by sucrose gradient centrifugation under different salt conditions. Western blotting with anti-PrA revealed that at low salt, the protein complexes including tagged mS49, mS63, mL46, and mL76 had the same sedimentation pattern as the complexes monitored by mt-rRNAs (Figure 1C, Supplementary Figure S4A, E, F). Three of the tagged mtRPs behaved the same way under all conditions, whereas most of the mL76-PrA shifted to fractions corresponding to smaller particles (∼15–20S) at high salt. This indicated that the tagged mL76 was relatively loosely attached to the mitoribosome (Figure 1C, Supplementary Figure S4B,C).

The tagged mtRPs together with their associated proteins were purified by two techniques: conventional immunoprecipitation (IP) (41) and rapid single-step affinity purification (AP) (71) (for more details, see Materials and Methods). Successful pulldown of the tagged proteins was confirmed by western blotting (Supplementary Figure S4G), as well as by testing enrichment of mt-rRNA in the eluates (Supplementary Figure S4H).

Liquid chromatography–tandem mass spectrometry (LC–MS/MS) was performed on three (IP) and four (AP) biological replicates per tagged protein together with the same number of replicates from a wild-type (IP) or mtPrA (AP) control. As an independent validation, we also performed LC–MS/MS on wild-type mitoribosomes purified by Blue-Native polyacrylamide gel electrophoresis (BN-PAGE; Supplementary Figure S4D), which produced similar results (see below), albeit with lower yield, coverage, and confidence values. Mass spectrometry data allowed us to calculate the abundance of proteins in the purified mitoribosomal fractions and also the level of enrichment by comparing the abundance before and after purification. Overall, the AP datasets contained fewer contaminants

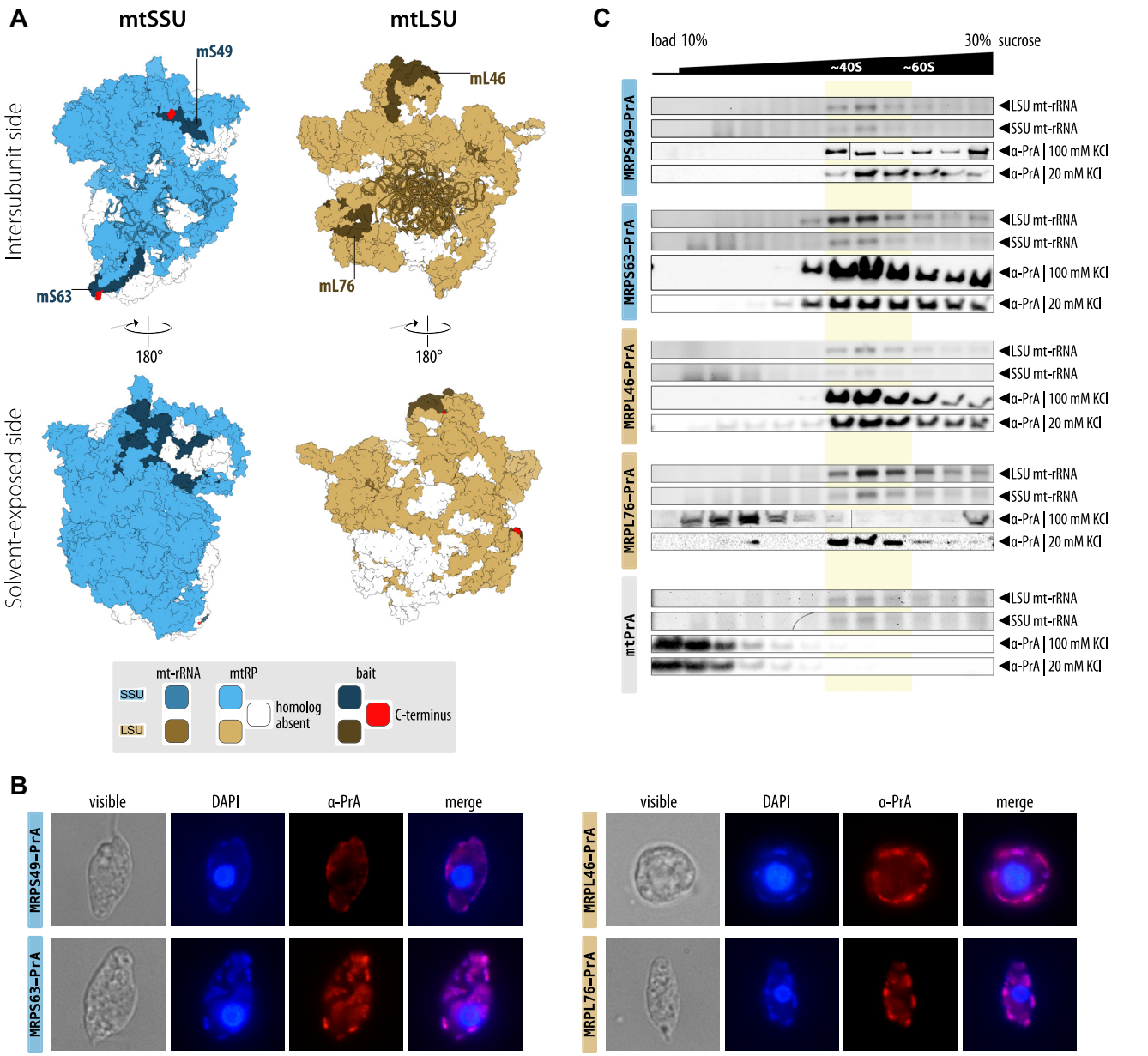


Figure 1. PrA-tagged mitoribosomal proteins are properly integrated into *Diplonema* mitoribosomes. (**A**) Views of the mitoribosomal small and large subunit of the kinetoplastid *Trypanosoma brucei* from the intersubunit and solvent-exposed sides. mtRPs with homologs present or absent in *Diplonema papillatum* are shown in different shades as indicated in the key. Locations of the four PrA-tagged baits and their C-termini are highlighted. RNA is shown as ribbon. (**B**) Sub-cellular localization of the PrA-tagged mtRPs using immunofluorescence microscopy (visible light, DNA staining with DAPI, anti-PrA antibodies, and merged fluorescence channels). Note that *Diplonema* contains a single reticulated mitochondrion, which is subcortically located; the dotted pattern of stained DNA with diffuse surrounding signal is due to image focusing onto a select plane. (**C**) Distribution of *Diplonema*'s mitoribosomal subunits in 10–30% sucrose velocity gradients in two different ionic conditions. Twelve fractions were collected from the top of the gradients for each cell line expressing the fusion mtRP or the mtPrA control. The peak fraction containing separated subunits is highlighted in yellow; the trail likely corresponds to a minor fraction of intact monosomes and certain assembly intermediates. The RNA fractions shown are from separations in 100 mM KCl. For full gels and blots, see Supplementary Figure S4.

and provided the strongest quantitative signals (see below; Supplementary Table S3).

Composition of the *Diplonema* mitoribosome and its assembly intermediates

The sequences of proteins that co-purified with the four tagged *Diplonema* mtPRs were compared to a collection of experimentally confirmed mtRPs and mitochondrial assembly (or maturation) factors (mtAFs) from a variety of eukaryotes (Supplementary Table S3). Further, the three-dimensional (3D) structures of the detected *Diplonema* proteins were predicted *in silico* using AlphaFold (see Materials and Methods). By integrating the sequence-, structure-, and MS/MS-based information, we established four categories: (i) function-assigned mtRPs with

counterparts in other organisms (referred to as 'assigned mtRPs'), (ii) candidate mtRPs with homologs exclusively in diplonemids (i.e., candidate mtRPs), (iii) function-assigned mtAFs, and (iv) diplonemid-specific candidate mtAFs (for details, see Materials and Methods; Figure 2, Supplementary Figures S5 and S6). Owing to the high reproducibility and efficacy of the AP pulldowns, it was possible to establish clearcut abundance and enrichment thresholds for the assigned mtRPs and mtAFs. As expected, the mtRPs were typically much more abundant than the mtAFs, reflecting the predominance of mature particles in the captured mitoribosomal population, as well as the transient nature of assembly-factor interactions compared to the integral proteins. The MS/MS-derived thresholds served to identify with confidence diplonemid-specific mitoribosomal components (Supplementary Table S3).

Diplonema's mitoribosome accumulated more proteins than its counterparts from other organisms

The *Diplonema* mitoribosome contains 94 assigned mtRPs, 42 in the SSU and 52 in the LSU. These proteins were consistently detected in pulldowns of a given subunit via different tagged mtRPs and a majority (41 SSU and 46 LSU) was also found in the wild-type mitoribosomes (Supplementary Table S3). Sixteen of the corresponding SSU and 16 of the LSU proteins were known from kinetoplastids (13) and therefore must have already been present in the common ancestor of diplonemids and kinetoplastids. Conversely, 11 SSU and 12 LSU kinetoplastid-specific mtRPs are either absent from diplonemids or have diverged beyond recognition by sequence comparison. Further, among the *Diplonema* proteins enriched in the pulldowns, we did not detect any homologs of the non-core mtRPs that are considered specific to opisthokonts, ciliates or plants (Supplementary Table S3).

Absence of numerous otherwise widely conserved mtRPs (11 from SSU and 19 from LSU) is another common feature of kinetoplastid and diplonemid mitoribosomes (Figure 3). Five further mtRPs (namely, uS3m, mS38, uL2m, uL10m, uL14m) appear to be missing or extremely divergent specifically in the diplonemid lineage (for more details, see Supplementary Information; Supplementary Figure S7). As in the cases of the contentious uS3m in kinetoplastids (13) and mS38 in yeast (10), this issue could be resolved by determining the *Diplonema* mitoribosome structure.

In addition, the SSU and LSU pulldowns revealed 20 and 16 candidate mitoribosomal proteins, respectively, with abundance and enrichment profiles convincingly similar to those of the assigned mtRPs and with counterparts limited to diplonemids (Supplementary Table S3). A mere quarter of the postulated diplonemid-specific mtRPs carry a conserved protein domain or have significant sequence similarity to a protein in the Uniprot database, with two themes emerging from their analyses. First, eight polypeptides either represent or resemble metabolic enzymes or nucleic acid-interacting factors (see Supplementary Information), a phenomenon that has also been observed in mitoribosomes of other species (19). Second, four candidate proteins may represent divergent paralogs of assigned diplonemid mtRPs (namely, uS11m, uS15m, mS23, uL24m) because they display remote sequence and/or predicted structural similarity (see Supplementary Information).

Diplonema's mitoribosome biogenesis involves at least 80 factors

Mitoribosomes undergo numerous maturation steps including base modification of the mt-rRNAs and folding of both mt-rRNAs and mtRPs directed by mtAFs. From predicted diplonemid proteomes we retrieved homologous sequences for 17 out of 44 SSU and 15 out of 29 LSU mtAFs reported in literature (from model organisms including animals, fungi, and kinetoplastids). Half of these mtAFs have homologs exclusively in kinetoplastids and diplonemids (Supplementary Table S3). The majority of such shared mtAFs were shown in *Trypanosoma* to bind close to the central plane of the mitoribosome and in proximity to the functionally most relevant segments of the mt-rRNAs (Figure 4A, B).

Homologs of the mtRPs that we tagged in *Diplonema* are incorporated into the kinetoplastid mitoribosome at early (mS49, mS63, mL76) or late (mL46) assembly stages (14,72–74) (Figure 4A, C, Supplementary Figure S2). Co-purification of certain proteins involved in the maturation of the mitoribosomal subunits—in addition to mtRPs—is therefore possible. Assuming sufficient shotgun-proteomics coverage and assembly intermediate stability, if a given mtAF is not co-captured with an mtRP bait (despite its presence in the proteome), it can be reasoned that the mtAF has dissociated prior to the recruitment of that particular mtRP. Conversely, the presence of a co-captured mtAF means that either it was already present in the intermediate, to which the mtRP is bound, or it was recruited after that mtRP. As we explore in the following sections, to a certain extent, this allowed us to infer the order in which the factors act.

Out of the 31 assigned mtAFs found in the *Diplonema* proteome, we detected by MS/MS 14 associated with the SSU and four with the LSU (Figure 2, Supplementary Figures S5, S6, S8, Supplementary Table S3). In addition to these function-assigned mtAFs, 50 proteins were pulled down with one or several baits and displayed similar abundance and enrichment levels as the assigned mtAFs (see Materials and Methods). These diplonemid-specific candidate factors were virtually absent from the preparation of the mature wild-type mitoribosomes resolved by native PAGE, had homologs in the inferred proteomes across diplonemids, and nearly all were predicted to be imported into mitochondria (Supplementary Figure S6, Supplementary Table S3). About half of these candidates carry functional motifs or domains reported in mtAFs from other organisms, such as pentatricopeptide repeat (PPR), RNA-binding motif, pseudouridine synthase, RNA methyltransferase, or RNA helicase, or display similarity to already assigned mtRPs or mtAFs (for details, see Supplementary Information). Taken together, these features support the notion that these proteins represent functionally relevant assembly factors.

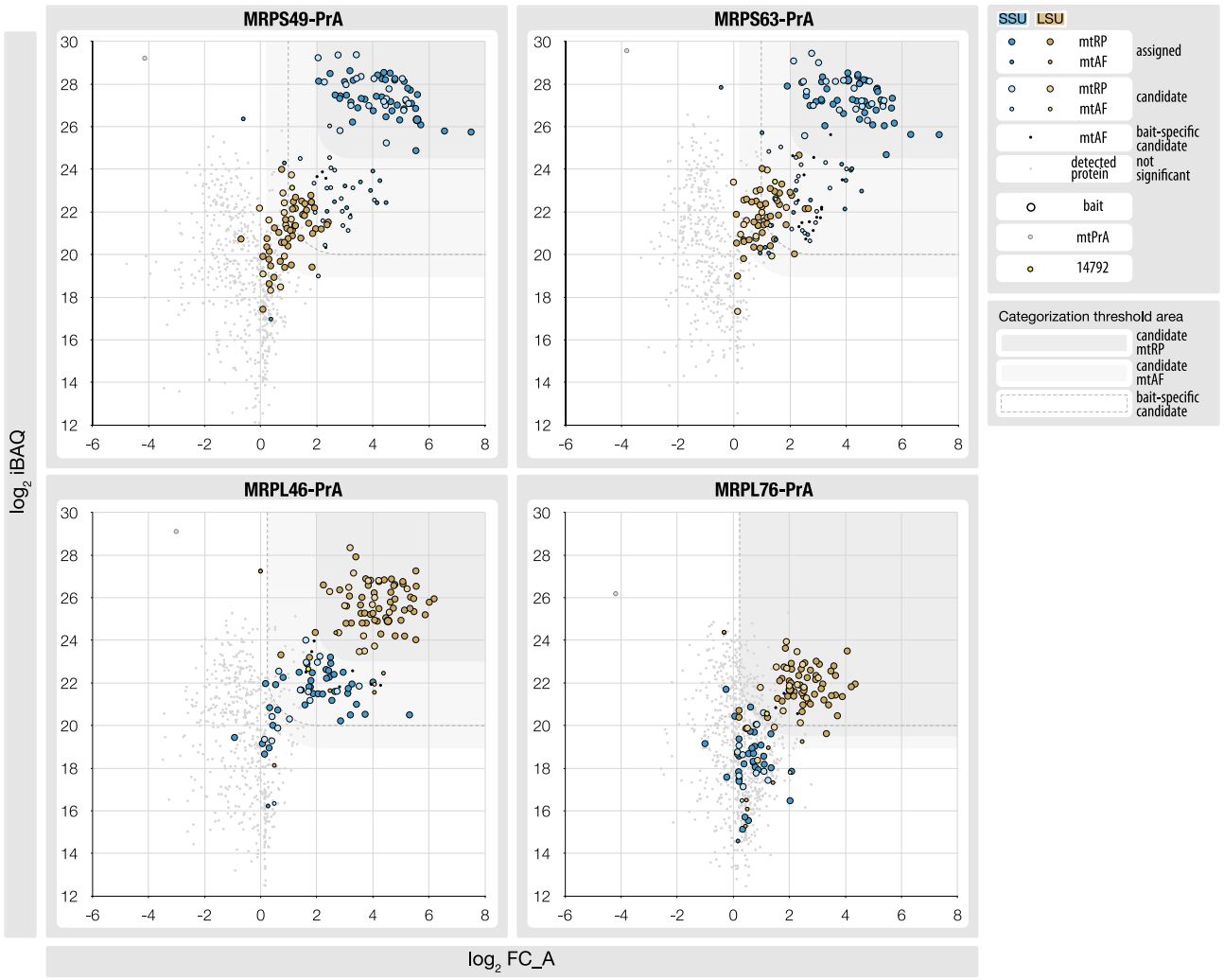


Figure 2. Cooccurrence with the assigned mitoribosomal components allows identification of *Diplonema*-specific mtRPs and mtAFs. Enrichment-abundance plots of proteins pulled down together with PrA-tagged mtRPs using the rapid affinity purification (AP) approach. Proteins were classified based on their enrichment (average fold change; FC_A), abundance (iBAQ), probability score (SP), and sequence similarity to known proteins (mitoribosomal proteins, mtRPs; assembly factors, mtAFs). Plot areas within enrichment and abundance thresholds used for prey categorization (see Materials and Methods) are highlighted in grey shades (mtRP and mtAF) and by dotted lines (bait-specific preys). Individual proteins are highlighted based on their categorization: (i) mtSSU- and mtLSU-associated in blue and orange hues, respectively; (ii) assigned (recognized homologs) and new candidates (diplonemid-specific) in darker and lighter shades, respectively and (iii) mtRP and mtAF by larger and smaller circles, respectively. Black dots indicate proteins that were above thresholds only in a specific AP pulldown sample (e.g. only in the four replicates of MRPS49-PrA). Two additional proteins, mtPrA (i.e. the control bait) and DIPPA_14792 (see Results), are shown as white and yellow-filled circles, respectively. Note that for a protein to be highlighted, it had to pass the SP $\geq$ 0.98 threshold in at least one of the pulldown conditions. The only exception is the mitochondrial acyl-carrier protein (mtACP), which is a multifunctional mitochondrial protein known to also act as an mtAF for both SSU and LSU in kinetoplastids (mt-SAF32 and mt-LAF11, respectively); mtACP is represented by the only high-abundance and low-enrichment point across the four plots (top-left quadrants). For MS/MS data from the traditional immunoprecipitation approach and analysis of wild-type mitoribosomes, see Supplementary Figure S5.

mtSSU assembly pathway: insights into early biogenesis

A total of 56 *Diplonema* proteins are predicted to be involved in the assembly of the mtSSU, notably 17 known from model systems and 39 diplonemid-specific mtAF candidates (Supplementary Figure S6, Supplementary Table S3). This number is comparable to what has been reported for kinetoplastids (reviewed in (75)).

In kinetoplastids, mS49 and mS63 are already present in the earliest characterized assembly intermediate called 'SSU assemblosome' (73), thus these proteins integrate in the mtSSU relatively early. The same appears to be the case in *Diplonema* because the interaction partners of these two mtRPs include a set of seven mtAFs exclusive to the intermediate (Figure 4A, C, Supplementary Figure S6). Among the conserved factors found in *Diplonema*, two assemblosome-bound mtAFs that remain associated with the maturing mitoribosome until later stages are also present (Figure 4A, C). Lastly, as in kinetoplastids, the maturing mtSSU of *Diplonema* contains homologs of mt-IF2 and an mt-IF3-like protein mt-SAF39, which contribute to late assembly steps (73,76). The presence of all these mtAF homologs in the mtSSU pulldowns shows that in *Diplonema*

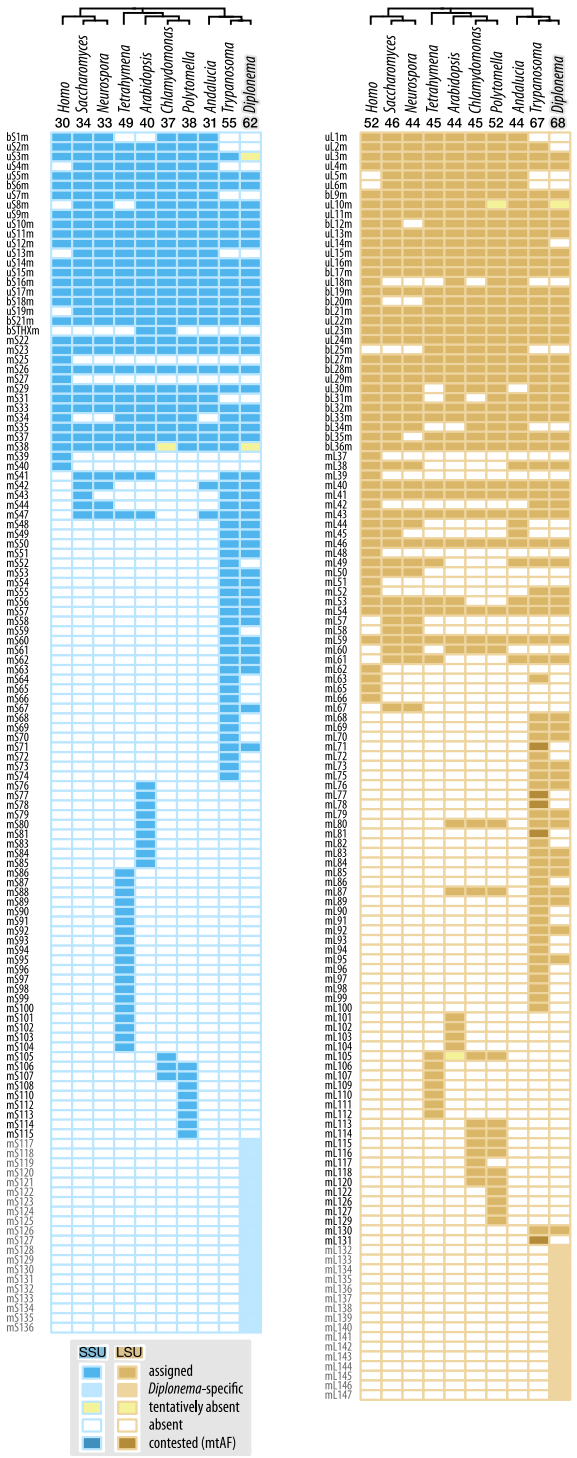


Figure 3. *Diplonema* contains the most protein-rich mitoribosomes. Distribution of currently known SSU and LSU mitoribosomal proteins in *Diplonema* and other eukaryotes. Presence of assigned or *Diplonema*-specific SSU and LSU mtRP homologs is indicated in different shades of blue and orange, respectively. Homolog absence is shown in white, while the yellow tint highlights tentative absence (no homolog retrieved by sequence and structure-based searches). Darker shades indicate proteins originally classified as mtRPs but contested as actually representing mitochondrial assembly factors (mtAF). For more details, see Supplementary Information and Supplementary Table S3. Numbers below species names specify their total number of SSU and LSU mtRPs (contested and tentatively absent mtRPs are excluded).

they are also involved in maturation steps after the incorporation of the two tagged mtRPs.

The complement of assigned and candidate mtAFs that was co-captured with *Diplonema* mS49 represented a subset of those retrieved in experiments with mS63 as a bait, albeit several were at relatively low abundance in the latter (Figure 4C, Supplementary Figure S6). Given that mS49 is part of the head while mS63 resides in the body of the mtSSU (Figure 4A), the asymmetric distribution of the interacting mtAFs indicates that the assembly of the mS63-containing pre-mitoribosomal particles requires a set of distinct mtAFs, which are released after completion of this portion, and that only then is mS49 recruited to the head region, accompanied by its associated set of factors.

From among the assigned mtAFs found associated with *Diplonema* mtSSU, this disparity also affects homologs of the GTPases YqeH (human mitochondrial counterpart MTG3/NOA1) and Era (ERAL1), as well as the RNA methyltransferase KsgA/RsmA (TFB1M) (Figure 4C, Supplementary Figure S6, Supplementary Table S3). In bacteria and human mitochondria, the two latter factors together with the RNA-binding protein RbfA (RBFA in human mitochondria) interact with and methylate the rRNA helix h45 (77,78). In *Diplonema,* these four mtAFs appear to be involved in the early SSU assembly, because they are detected in large quantities in the pulldowns with mS63, which integrates prior to mS49, as we documented above. Interestingly, the mS49 pulldowns (although less efficient than mS63) yielded mtAF levels that are positively correlated with the order in which the h45-interacting mtAF trio acts: ERAL1 < TFB1M < RBFA (Figure 4C, Supplementary Table S3). This observation suggests that by the assembly stage during which mS49 integrates, ERAL1 and TFB1M (operating as the first and second interactor in the cascade, respectively) have effectively left the maturing mtSSU. This conclusion is in line with a recent demonstration that all four mtAFs are part of early assembly intermediates in human mitochondria (79,80). Lastly, while analyzing *Diplonema*'s homologs, we determined that the kinetoplastid assemblosome protein designated mt-SAF18 (73) actually represents a divergent homolog of human RBFA (see Supplementary Information). This means that in kinetoplastid mitochondria, the h45 methylation is also a relatively early process, since the last contributor in the cascade is present in the assemblosome particles.

The exact role of the h45-interacting mtAF trio is to sequentially bind to the rRNA regions in proximity to the helix, during which two widely conserved consecutive adenosines in the corresponding loop are dimethylated by the methyltransferase (77,78). These functions are conserved in human mitochondria (79,80), whereas the role of the trio is unlikely to be conventional in *Diplonema*. While the diplonemid ERAL1 and RBFA homologs carry all functionally important segments and residues, the catalytic pocket of their TFB1M, which normally binds *S*-adenosylmethionine (SAM) and the adenosines of h45 of mtSSU rRNA, differs notably from its well-studied counterparts (81,82), having substituted most of its function-critical residues (Figure 4D, Supplementary Figure S9). This suggests that diplonemid TFB1M has become purely an rRNA folding factor instead of an enzyme with a

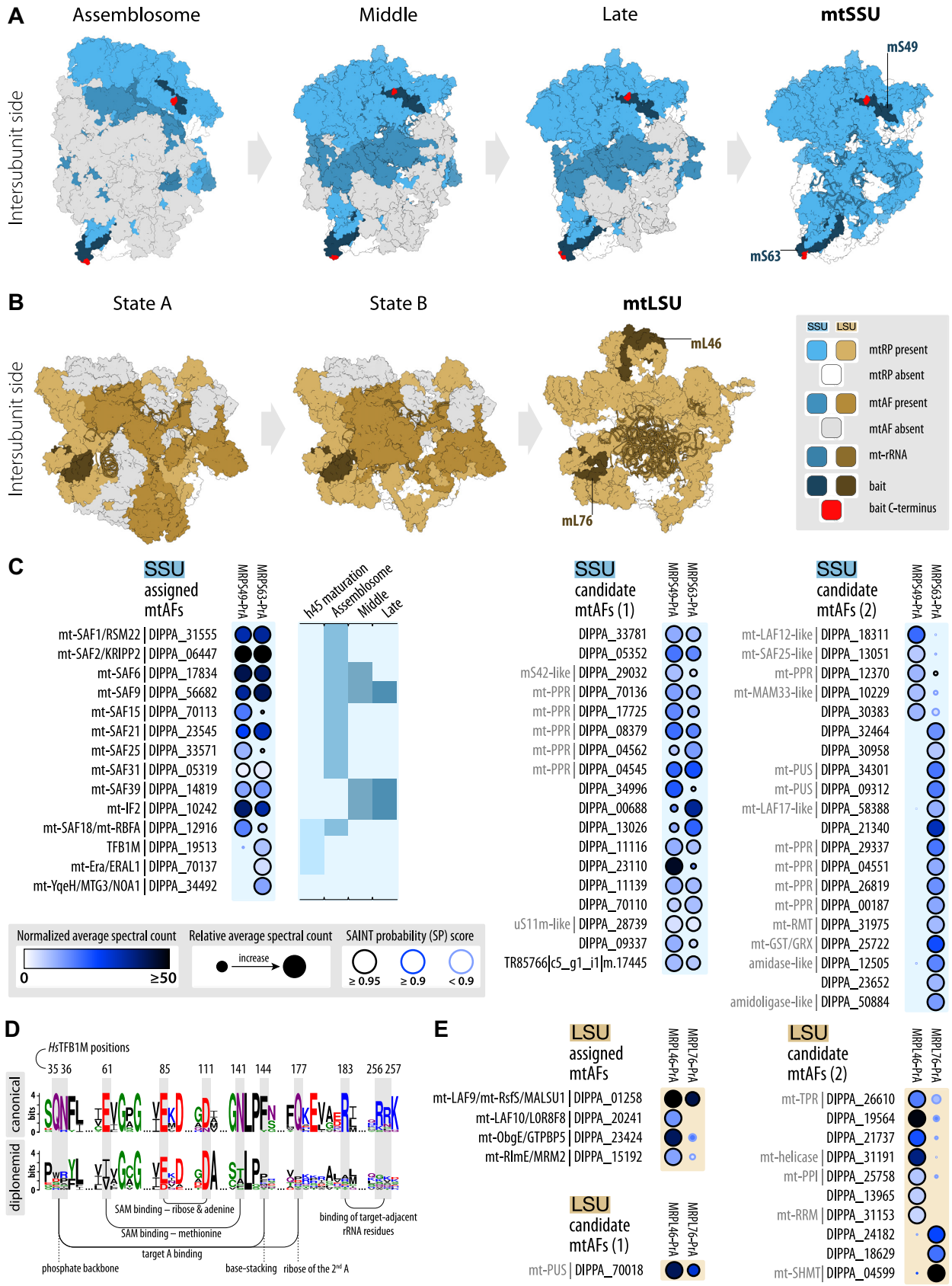


Figure 4. Spatiotemporal characteristics of bait incorporation facilitate dissection of mitoribosome biogenesis stages. (A, B) Intersubunit side views of mitoribosomal small (**A**) and large (**B**) subunit from *Trypanosoma brucei* at successive assembly stages. Proteins (mtRPs or mtAFs) with homologs present

secondary folding capacity (for details, see Supplementary Information).

mtLSU assembly pathway displays predominantly conserved features in its late stages

As mentioned above, 19 (assigned or candidate) protein factors involved in the assembly of the *Diplonema* mtLSU were identified by mass spectrometry of RNPs pulled down with either mL46 or mL76 or both mtRPs. This contrasted with 12 and 27 mtAFs involved in the mtLSU maturation in human and kinetoplastids, respectively (Supplementary Table S3). The low number of LSU mtAFs detected in *Diplonema*, especially compared to kinetoplastids, can be attributed to a combination of technical difficulties of working with the cell line expressing mL76-PrA and mL46 being among the last mtRPs to integrate into the mitoribosome, according to studies in *Trypanosoma* (72,74) (Figure 4B). In contrast to the SSU baits, we observed only two proteins that co-purified together with both LSU baits above our strict thresholds: *Diplonema*'s homolog of mt-LAF9 (also known as mt-RsfS or MALSU1; see also below) and a mitochondrial TruD-superfamily pseudouridine synthase (mt-PUS, a candidate diplonemid-specific AF; Figure 4E, Supplementary Figure S6, Supplementary Table S3). In the case of the SSU mtAFs, the asymmetric distribution of several candidate mtAFs associated more specifically with mS49 than with mS63 (Figure 4C) allowed us to infer the order of mtRP recruitment during the assembly. Similarly, five candidate mtAFs displayed more pronounced enrichment in the mL46 pulldowns, but still associated, to a lesser degree, with mL76 (Figure 4E), which is in line with mL76 being incorporated before mL46 as in kinetoplastids (Figure 4B).

In addition to the aforementioned mt-LAF9, four of the assigned *Diplonema* mtAFs were substantially enriched exclusively in the mL46 pulldowns. Two are close binding partners of mt-LAF9, namely mt-LAF10 (i.e. L0R8F8) and mt-LAF11 (mitochondrial acyl-carrier protein; mt-ACP). In both human (83) and kinetoplastids (14,72,74), the three LAF proteins form a protrusion on the intersubunit side of mtLSU precursors, which prevents premature association with the mtSSU. It was experimentally demonstrated for human mitoribosomes (reviewed in (84)) that the dissociation of these LAFs from the assembly intermediate represents the last maturation step.

The two remaining mtAFs detected at high levels in the pulldowns of *Diplonema*'s mL46, and to a lesser extent in those of mL76 (Figure 4E), are orthologs of mt-RlmE/MRM2 and mt-ObgE/GTPBP5. The former protein methylates the ribose 2′-O of the first U in the H92 loop (aka the 'A-loop'), while the latter is a GTPase that probes the peptidyl-transferase center (PTC) helices H89–H93 and the A-loop, presumably as a quality-control mechanism (85–87). *Diplonema*'s MRM2 carries all residues important for its catalytic activity, an observation that is in line with the conserved features of its LSU mt-rRNA in the PTC region, including the presumed target of the methyltransferase (30). In the human mitoribosome, MRM2 and GTPBP5 are present in the very late assembly intermediates when bL36m is being incorporated (88–90). Based on their significant enrichment in the mL46 pulldowns (and presence to a lesser extent in the mL76 samples), we conclude that diplonemids and most likely kinetoplastids as well share this late assembly intermediate of the mitoribosome maturation pathway with humans.

Both subunits interact with a component of the tRNA-import machinery

To identify components that interact with both the SSU and LSU of the *Diplonema* mitoribosome, we examined, which proteins were significantly enriched (fold-change score > 1) across all experiments. Besides 27 assigned and five candidate mtRPs, a single protein was ubiquitously found: DIPPA_14792 (Figure 2, Supplementary Figures S5, S6). Highly conserved across diplonemids (61% pairwise identity), this protein with a single predicted C-terminal transmembrane helix is a homolog of Tb927.9.7830 (previously Tb09.v1.0420), a conserved kinetoplastid protein that is a critical component of *Trypanosoma*'s mitochondrial tRNA-import machinery (91). We detected DIPPA_14792 in earlier MS/MS analyses of purified mitochondria and mitochondrial high-molecular weight complexes (35) and now find it associated more specifically with the enriched wild-type mitoribosome (Supplementary Figure S5). Overall, these findings indicate that the protein interacts, in a relatively stable manner, with the mature mitoribosome. We discuss possible implications in the Supplementary Information.

Diplonemid SSU mt-rRNA carries only two canonical helices, h18 and h44

We previously determined by deep RNA-Seq of sucrose gradient-enriched *Diplonema* mitoribosomes that these particles contain two RNA species of 374 and 930 nt, representing the SSU and LSU mt-rRNA, respectively (30), a

←

or absent in *Diplonema papillatum* are shown in different shades as indicated by the key (for more details, see Supplementary Figure S2). (**C**) Dot plot visualization of relative protein abundances measured in average spectral counts (see the key for details). *Diplonema*'s proteins are grouped together into assigned mtAFs, candidate mtAFs present in pulldowns of both baits from a given subunit (1), and candidate mtAFs significantly prevalent in pulldowns of a single bait (2). The bar plot shows to which of the four identified mitoribosome biogenesis stages the assigned mtAFs contribute. Assignments of mtAFs are shown next to their gene identifiers; for the candidate mtAFs, similarity to known proteins or functional categorization is indicated (PPR/TPR, pentatricopeptide/tetratricopeptide repeat protein; PUS, pseudouridine synthase; RMT, RNA methyltransferase; RRM, RNA recognition motif; PPI, peptidyl prolyl isomerase; GST/GRX, glutathione-S-transferase/glutaredoxin). (**D**) Diplonemid homologs of the methyltransferase TFB1M lack most catalytic residues. Excerpts from the sequence logo based on the alignment of TFB1M homologs from diplonemids and other eukaryotes. Shown are only sequence segments around catalytic residues involved in the binding of the co-factor *S*-adenosylmethionine (SAM), the target As of the SSU mt-rRNA, and the target-adjacent rRNA residues. Protein residue numbering is provided using human TFB1M as a reference. (For more details, see Supplementary Figure S9.) (**E**) Dot plot visualization of relative protein abundances for the most prominent mtLSU-interacting proteins. Abbreviations of functional categorization: TPR, tetratricopeptide repeat protein; PUS, pseudouridine synthase; RRM, RNA recognition motif; PPI, peptidyl prolyl isomerase; SHMT, serine hydroxymethyltransferase.

finding now confirmed *via* protein pulldown experiments (Supplementary Figure S4H). Using the earlier published secondary structure model of the LSU mt-rRNA from *Diplonema* (30), here we compared the rRNA sequences across 11 diplonemids, which pinpointed the sparse conserved sequence motifs. Essentially all conserved signature positions and structural elements had been correctly assigned in *D. papillatum*, except that the domain VI appears more conventional than originally proposed (see also Supplementary Information).

We also examined the diplonemid SSU mt-rRNA, which has not been analyzed in detail so far. This structural RNA varies in size across eukaryotes, yet maintains a degree of base-pair formation comparable with that of the bacterial SSU rRNA, which contains 45 helices with ∼60% of its residues engaged in base-pairing (Supplementary Figure S10) (10–11,16–18,21). In contrast, kinetoplastid SSU mt-rRNAs are not only substantially shorter (by 60% compared to a typical bacterial counterpart), but their residues pair to a much lesser degree: for example, *Trypanosoma* SSU mt-rRNA contains only 53 canonical and 17 non-canonical base pairs (13), amounting to 22% of its 620-nt size (Figure 5A). Out of the 45 typical helices, only 13 are preserved at least to some extent, while 12 others appear to have been substituted by single-stranded regions effectively mimicking the spatial arrangement of the canonical helices but without engaging in proper base pairing (13,14); we refer to these as 'helix-mimics'. Only three canonical helices are well preserved at both the sequence and structural level: h18 in the 5′ domain, as well as h44 and h45 in the 3′-minor domain (Figure 5A). The helices h18 and h44 contain some of the few function-critical, universally conserved residues (92); they are juxtaposed in the ribosome and cooperate in the decoding process (93,94).

SSU mt-rRNAs in diplonemids have contracted even further: at a median length of 380 nt, they represent a mere ∼25% of a typical bacterial SSU rRNA. However, downsizing is not the only unorthodox feature of these rRNAs. They are extremely variable across diplonemids (26% and 4% sequence identity across closely related *Diplonema* spp. and all diplonemids, respectively); identifying any recurrent secondary structure elements is therefore difficult. We could only assign two short conserved motifs diagnostic of SSU rRNA, whose surrounding regions have the potential to adopt canonical structures of the helices h18 and h44 (Figure 5B, C). Importantly, in all diplonemids, the two elements contain at the expected positions the counterparts of the residues G530 (conventional *E. coli* numbering; in h18), as well as A1492 and A1493 (in h44) that play an essential role in selecting cognate tRNAs in the ribosomal A-site during translation (Figure 5B, C). In contrast to kinetoplastids, it is not clear if a genuine helix h45 exists across diplonemids: its size is ≥ 8 bp across eukaryotes, but in diplonemids, the corresponding region downstream of h44 is only ∼23 nt-long on average and has the potential to fold into a 2–5 bp-long helix at most, if at all. Therefore, it seems equally plausible that this region could merely spatially approximate the h45 shape. Overall, the paucity of conserved secondary structure elements in diplonemid SSU mt-rRNAs is highly reminiscent of their kinetoplastid counterparts and even more pronounced.

The most conspicuous portion of the diplonemid SSU mt-rRNAs is located just upstream of h18, covering most of the 5′ domain. The ∼100 nt-long region is rich in U and non-conventional inosine (I) residues, with only a few As and Cs. This situation resembles the prominent A + U enrichment and C depletion seen in the same region of the kinetoplastid molecule (Figure 5D). Note that in *Diplonema*, 41% of nucleotides in this region originate from deamination RNA editing and the rate is even higher in certain other diplonemids (29). The evolutionary scenarios leading to the massive editing of mt-rRNAs in diplonemids have been discussed earlier (35,95–96).

The diplonemid SSU mt-rRNAs are ∼40% shorter than the kinetoplastid homologs. To elucidate whether in *Diplonema* the 'missing' mt-rRNA regions could have been replaced by proteins, we examined the extent to which mtRPs that contact rRNA elements in the *Trypanosoma* mtSSU 3D structure differ from their *Diplonema* counterparts. Among such 43 *Trypanosoma* proteins, seven homologs in *Diplonema* carry extensions of at least 20 amino acid (aa) residues precisely in those polypeptide-chain segments that contact the rRNA in *Trypanosoma* (Figure 5E). For example, the N-terminus of *Trypanosoma* uS11m is wedged in between the helix h29 and the helix-mimics of h38 and h43; the *Diplonema* homolog contains an ∼60 aa-long N-terminal extension (excluding the mitochondrial targeting peptide, the removal of which from the mature mtPR was confirmed via MS/MS). Three-dimensional structure prediction of these seven mtPRs showed that the extensions in the *Diplonema* proteins do not form well-defined shapes, but rather exist as largely disordered segments with occasional short α-helical stretches. mtRP extensions in other organisms are also predominantly unstructured (16,97) and when positioned near RNA, such disordered extensions tend to meander through the spaces freed by the missing rRNA elements (11,13). Thus, we posit that diplonemid-specific mtRP extensions occupy spaces corresponding to h6–h13, h20–h23, h32–h33 and h42 of the *Trypanosoma* mt-SSU rRNA, which represent peripheral portions of the 5′, central, and 3′-major domains of the RNA. It would be interesting to investigate, by combining structure determination with expression of truncated mtRPs, whether the replacement of RNA by proteins plays a role in stabilization or decoding function of the diplonemid mtSSU.

Lastly, diplonemids appear to lack homologs of seven other *Trypanosoma* mtRPs that interact with the SSU mt-rRNA, contacting its 5′ and 3′-major domains (Figure 5E). We postulate that in *Diplonema*, these proteins are substituted by some of the 20 diplonemid-specific mtRPs identified here.

DISCUSSION

The diplonemid mitoribosome displays the fewest conserved and numerous unorthodox features

Here we elucidated the composition of the mitoribosome from *Diplonema papillatum*, which not only differs significantly from its bacterial ancestor, but also from any other studied mitoribosome. The dramatically reduced LSU and SSU mt-rRNAs of this protist were first identified based

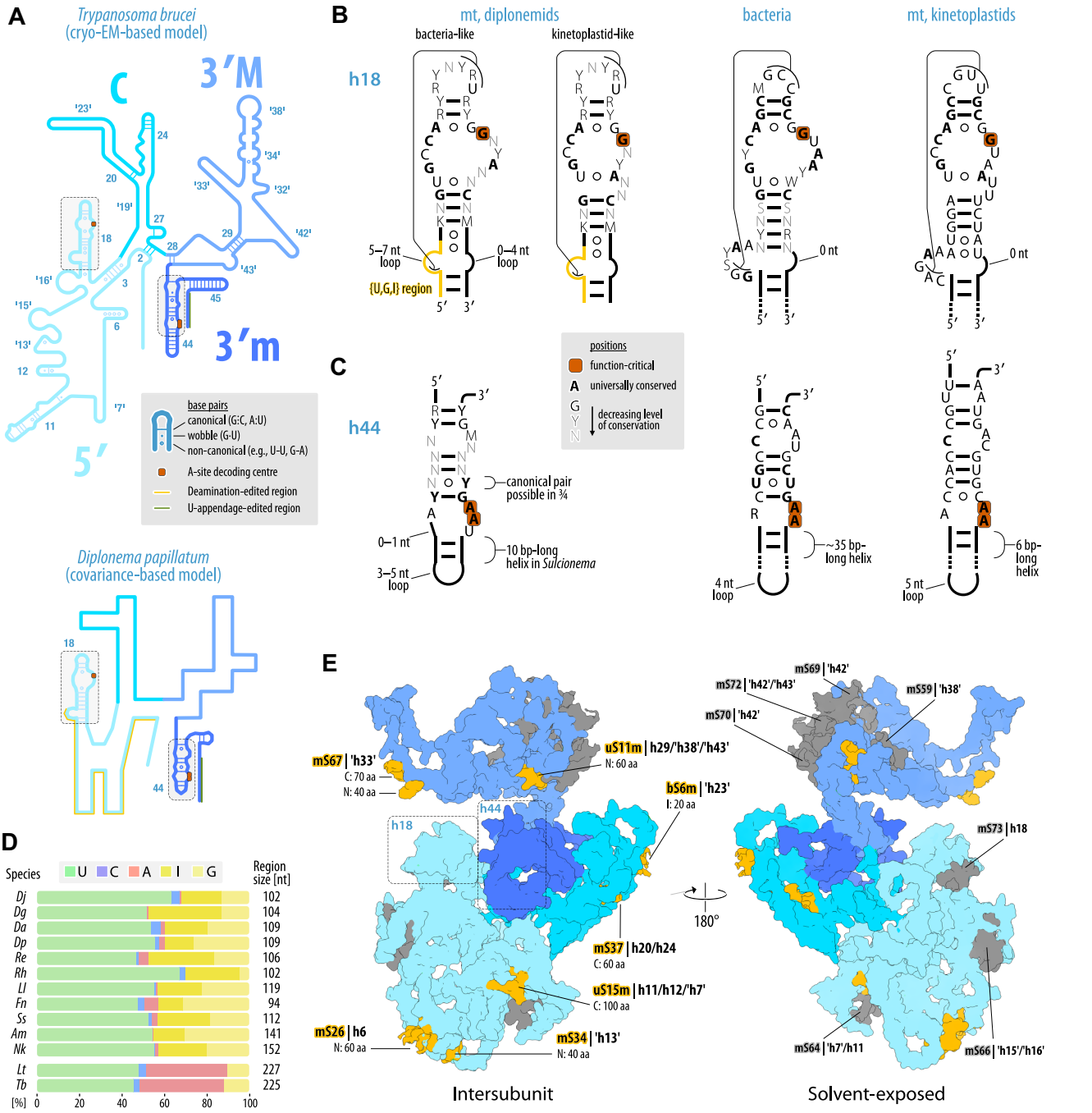


Figure 5. Derived and diminutive diplonemid SSU mt-rRNA contains little canonical structure beyond the A-site decoding center. (**A**) Secondary structure models of the divergent SSU mt-rRNAs from *Trypanosoma* and *Diplonema*. The four structural domains (5′, central [C], 3′ major [3′M], and 3′ minor [3′m]) of the SSU rRNA are shown in different shades of blue. Regions of the helices h18 and h44, which constitute the decoding center, and the positions of the function-critical rRNA residues are highlighted in gray and red, respectively. In the case of the *Diplonema* covariance-based model, secondary structure can only be predicted for h18, h44 and reduced h45; the rest of the model is drawn to scale and approximately indicates which domain segments the remaining sequence most likely covers (see also (E)). (**B**) Alternative secondary structure models of the helix h18 from diplonemid mitochondria compared to the structures in bacterial and kinetoplastid mitochondrial SSU rRNA. Canonical (bacteria-like) and a kinetoplastid-like base-pair arrangements are shown for diplonemids. The 5′ proximal segment of the diplonemid h18 overlaps the deamination-edited U + G + I-rich region (in yellow). (**C**) Secondary structure model of the helix h44 from diplonemid SSU mt-rRNA compared to the corresponding structures in bacteria and kinetoplastid mitochondria. Note that the helix is extremely truncated except in *Sulcionema*. (**D**) Size and base-composition comparison of the 5′ domain upstream of h18 across diplonemids and kinetoplastids. In diplonemids, instead of adenosines, the region is prominently enriched in the non-conventional nucleoside inosine. (**E**) Locations where RNA is likely substituted by proteins in the *Diplonema* mtSSU. A cryo-EM-based structure of *Trypanosoma*'s SSU mt-rRNA, viewed from the intersubunit and solvent-exposed sides, is coloured in the same shades of blue as in (A) to specify the four SSU rRNA structural domains. Predicted positions of long diplonemid-specific mtRP extensions are highlighted in orange, indicating which mtRP is extended and which RNA secondary structure elements are in the vicinity. Many if not all of the kinetoplastid-specific mtRPs that are apparently missing in diplonemids (shown in gray) are most likely functionally substituted by diplonemid-specific mtRPs (see the text for details).

on their sheer abundance in the mitochondrial transcriptome together with a handful of sequence motifs generally well conserved across eukaryotes but rather degenerate in diplonemids (30,31). We have confirmed here that *Diplonema* and, by extrapolation, diplonemids as a whole have only two mt-RNA species. The 930-nt molecule associated with the LSU and the 374-nt molecule associated with the SSU represent the most diminutive, experimentally confirmed contiguous mature rRNAs identified thus far. Comparison across ten diplonemids showed that outside the critical regions spatially proximal to the PTC in the LSU and the A-site decoding centre in the SSU, the mt-rRNAs vary considerably in sequence and secondary structure. This suggests that the rest of the rRNA molecules provides a minimalist, lenient scaffold for arranging mtRPs.

The radical downsizing of the *Diplonema* mt-rRNAs has been accompanied by the loss of only a few core components but a remarkable gain of three dozen diplonemid-specific mtRPs. As a consequence, the total number of the organism's mtRPs has grown to a total of 130, exceeding by far the 75 in the most bacteria-like mitoribosome of *Andalucia* (21) and even the 122 of *Trypanosoma*, the hitherto most protein-rich mitoribosome known (13). Furthermore, the majority of canonical mtRPs have experienced N- or C-terminal extensions, and this to a larger extent than seen in most other organisms (see Supplementary Information; Supplementary Figure S11). Thus, despite the reduction of the RNA content, the protein size gain translates into an overall exceptionally massive mitoribosome with a calculated mass of 5.1 MDa, compared to ∼2.7 and ∼4.5 MDa in human and *Trypanosoma*, respectively.

Consequently, the protein-to-RNA ratio in *Diplonema*'s mitoribosomes increased to ∼11:1, which is—compared to the ratios 8:1 in *Trypanosoma*, 2:1 in mammals, and 1:1 in *Andalucia*—the highest value determined to date. It is mostly the mtSSU that skews the ratio: while *Diplonema*'s 62 SSU mtRPs total ∼2.7 MDa (excluding MTPs), the tiny rRNA contributes a mere ∼120 kDa, resulting in an unparalleled 22:1 ratio of this subunit. As we documented in detail for the SSU, a number of *Diplonema*-specific extensions of canonical mtRPs, and potentially some of the new mtRPs, most likely occupy the space where RNA structural elements would be positioned in a mitoribosome with a more conventional rRNA.

Evolutionary and functional aspects of unconventional mitoribosomal proteins

The accretion of novel proteins to mitoribosomes is commonly observed across eukaryotes (19), which is indicative of processes framed under the theory of constructive neutral evolution (98) being rampant in the evolutionary path leading to extant mitoribosomes. In the investigated members of the Archaeplastida lineage—*Arabidopsis*, *Chlamydomonas* and the non-photosynthetic green alga *Polytomella magna*—mtRPs have been primarily recruited from two families of helical repeat proteins, the pentatricopeptide repeat (PPR) and octatricopeptide repeat (OPR) proteins. Most members of these protein families interact with RNA (15). Among the mtRPs recruited in kinetoplastids and ciliates are only a few helical repeat proteins (reviewed in (99)), whereas the large majority of the lineage-specific mtRPs lack well-defined domains (13,16). A similar pattern emerges in diplonemids, with only one among 36 *Diplonema*-specific mtRPs being a helical repeat protein (PPR protein DIPPA_32001 in mtSSU), while nearly 80% of *Diplonema*-specific mtRPs have no discernable domains.

The third category of new recruits found in mitoribosomes from yeasts (10) to ciliates (16) to kinetoplastids (13) is represented by proteins derived from metabolic enzymes that now play a purely architectural role in mitoribosomes. In most cases, these repurposed proteins lack residues necessary for their original catalytic activities. For example, DIPPA_19549 and DIPPA_04650, two *Diplonema* LSU mtRPs, bear a non-functional hydrolase and *O*-acetyl-ADP-ribose deacetylase domain, respectively (see Supplementary Information).

In a curious twist on this recurrent trend, two unorthodox mtSSU proteins of *Diplonema* have apparently retained their enzymatic activity. DIPPA_03129 is a mitochondrial monothiol glutaredoxin and DIPPA_25989 encodes a BolA-type protein, both critical components of the mitochondrial iron-sulfur (Fe-S) cluster assembly and trafficking. Studies in human and yeast show that the two proteins form a transient complex and participate in the generation of [4Fe–4S] and [2Fe–2S] clusters and the incorporation of these prosthetic groups into target enzymes (reviewed in (100)). The proteins' prominence and specificity in both mS49 and mS63 pulldowns argues against merely spurious interaction with the *Diplonema* mitoribosome. We speculate that in diplonemids, the association of Fe-S cluster synthesis with the mitoribosome facilitates the maturation of the mtDNA-encoded and mitochondrially translated Nad8 (aka NDUFS8). This protein is a subunit of the mitochondrial respiratory chain Complex I and binds two [4Fe–4S] clusters. Another possible target is the cluster-containing assembly factor mt-SAF1 (METTL7/RSM22). Alternatively, the function of the protein pair in *Diplonema*'s mitoribosome could be entirely separate from its canonical roles in the Fe-S cluster assembly, as in the case of documented multi-functional, 'moonlighting' proteins (see Supplementary Information). Determining the role(s) of these two proteins is bound to provide new insights into the interplay of metabolism and translation.

Variations of the mitoribosome assembly process

Our approach allowed us to sample not only the presumably predominant population of mature subunits and full mitoribosomes in *Diplonema*, but also certain assembly intermediates. Overall, mitoribosome biogenesis appears to be similarly complex in diplonemids and kinetoplastids, as the sheer number of the high-confidence non-mtRP interactors indicates; for instance, the number of those associated with *Diplonema*'s mtSSU almost matches that of mtRPs (56 versus 62). Importantly, our study also provides a view into processes that have not been previously investigated in kinetoplastids.

In all organisms examined, the mtRP mL46 is among the last to be incorporated into the mitoribosome. In

Diplonema, mL46 co-purified with, in particular, two PTC-interacting maturation factors mt-RlmE/MRM2 and mt-ObgE/GTPBP5. This finding corroborates the view that in diplonemids these two mtAFs participate only in the very final phases of the mitoribosome biogenesis. The same situation was observed in human mitoribosomes (reviewed in (84)). That the interaction of these two mtAFs is conserved in highly derived diplonemid mitoribosomes suggests that their action is ubiquitous, including in kinetoplastids, as well as evolutionarily ancient.

Immunoprecipitations with *Diplonema* mS49 and mS63 provided insight into steps preceding the SSU assemblosome, i.e., prior to the stages that have been investigated before (73). For instance, we inferred that 20 diplonemid-specific mtAFs, representing 50% of all SSU-associated mtAFs, act in such early steps. Currently, their roles can be extrapolated only in about half the cases (e.g. rRNA modification by pseudouridine synthases or folding of rRNA assisted by PPR proteins), justifying their further biochemical and structural investigation. Several assigned *Diplonema* SSU mtAFs that have not been analyzed in kinetoplastids and only sparingly in other organisms were mapped to early assembly stages, namely mt-YqeH/MTG3/NOA1, mt-Era/ERAL1, and TFB1M. The first factor is a GTPase, which was only very recently shown to attach to the earliest examined human mtSSU assembly intermediate (79). Its presence at a similar stage in diplonemids suggests a conserved role (but see also Supplementary Information). The other two proteins commonly cooperate with RBFA/mt-SAF18 in a sequential manner to remodel and methylate the SSU mt-rRNA helix h45. We advance the idea that in diplonemids, as well as in kinetoplastids, this coordinated mtAF trio commences its activity relatively early in the assembly process, similarly to its human counterpart (79,80). However, in contrast to the human system, diplonemid and kinetoplastid proteins complete their tasks at a comparatively earlier stage as more numerous maturation steps are required to assemble the protists' much more complex mtSSU RNP (see also Supplementary Information).

Finally, we posit that specifically in diplonemids, the TFB1M homolog has only retained its secondary function, h45 remodelling (101,102), but lost its RNA methyltransferase activity. This is because the protein's sequence has undergone catalysis-compromising substitutions, possibly because its target rRNA positions have been lost and the target helix has been substantially reduced. At least one analogous case has recently been documented experimentally: in human cells, catalytically inactive mutants of the LSU mt-rRNA-targeting methyltransferase MRM2 support normal mitochondrial translation, which indicates that in humans, this protein is primarily involved in mitoribosome biogenesis as a chaperone, not as an RNA modifying enzyme (103).

Concluding remarks

The diplonemid mitoribosome pushes the limits of how far a translation machinery is allowed to diverge from its proto-mitochondrial bacterial ancestor while still remaining functional. The bulk of the diminutive mt-rRNAs appears to have adapted to imitating helical structures without necessarily forming base pairs, while losses of otherwise widely distributed canonical mitoribosomal proteins coupled with new acquisitions highlight the gradual encroachment by proteins not only on the RNA structure, but apparently also on its function, such as decoding. The singularity of diplonemid mitoribosomes makes a case for future cryo-EM structure determination, but also for the development of platforms for RNAi- or CRISPR-Cas-mediated gene manipulation in these protists for eventual functional studies of integral components and maturation factors. For now, our catalog of mS63-associated mtAFs provides a window into the very early steps of the mitoribosome assembly. The continued exploration of divergent systems such as the diplonemid mitoribosome will be instrumental in understanding how evolution can impact both biogenesis and function of complex molecular machines.

DATA AVAILABILITY

New data reported in this article are available in the online supplementary material. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium *via* the PRIDE partner repository (104) with the dataset identifiers PXD040176 (AP), PXD039927 (IP) and PXD039882 (wild-type mitoribosomes).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Victor Tobiasson (Stockholm University, Sweden) for fruitful discussions and help with the choice of mitoribosomal proteins for tagging and Mohamed Aoulad Aissa (Université de Montréal, Canada) for excellent technical assistance in preliminary experiments.

Author contributions: M.V., J.L. conceptualization; M.V., C.B., M.W.G. data curation; C.B., D.F., L.C.A., M.V., M.W.G., O.G. investigation; M.V., C.B. formal analysis; C.B., L.C.A., M.V., O.G. methodology; G.B., J.L. project administration, supervision; A.Z., G.B., J.L., M.O., O.G. funding acquisition; M.V., O.G., C.B. visualization; M.V., C.B. writing – original draft; C.B., G.B., J.L., M.V., M.W.G., O.G. writing – review & editing. All authors read and approved the final manuscript.

FUNDING

Natural Sciences and Engineering Research Council of Canada (NSERC) [RGPIN-2014-05286, RGPIN-2019-04024 to G.B., RGPIN-2020-06924 to M.O.]; Czech Science Foundation (Grantová agentura České republiky; GACR) [23-06479X to J.L., 20-04150Y to O.G.]; European Research Council (ERC) [101044951 (MitoSignal) to A.Z.]; The Gordon and Betty Moore Foundation (GBMF) [9354 to J.L.]; A.Z. and J.L. acknowledge past support of research in this area by OP VVV [16_019/0000759] from the Czech Ministry of Education, Youth and Sports; M.W.G. acknowledges past support of research in this area by CIHR [MOP-4124]. Funding for open access charge: Grantová agentura České republiky.

Conflict of interest statement. None declared.

REFERENCES

1. Kořený,L., Oborník,M., Horáková,E., Waller,R.F. and Lukeš,J. (2021) The convoluted history of haem biosynthesis. *Biol. Rev*, **97**, 141–162.
2. Lill,R., Diekert,K., Kaut,A., Lange,H., Pelzer,W., Prohl,C. and Kispal,G. (1999) The essential role of mitochondria in the biogenesis of cellular iron-sulfur proteins. *Biol. Chem.*, **380**, 1157–1166.
3. Spinelli,J. and Haigis,M. (2018) The multifaceted contributions of mitochondria to cellular metabolism. *Nat. Cell Biol.*, **20**, 745–754.
4. Gray,M. (2015) Mosaic nature of the mitochondrial proteome: implications for the origin and evolution of mitochondria. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 10133–10138.
5. Burger,G., Gray,M., Forget,L. and Lang,B. (2013) Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists. *Genome Biol. Evol.*, **5**, 418–438.
6. Flegontov,P., Michálek,J., Janouškovec,J., Lai,D.-H., Jirků,M., Hajdušková,E., Tomčala,A., Otto,T.D., Keeling,P.J., Pain,A. *et al.* (2015) Divergent mitochondrial respiratory chains in phototrophic relatives of apicomplexan parasites. *Mol. Biol. Evol.*, **32**, 1115–1131.
7. Ott,M., Amunts,A. and Brown,A. (2016) Organization and regulation of mitochondrial protein synthesis. *Annu. Rev. Biochem.*, **85**, 77–101.
8. Desmond,E., Brochier-Armanet,C., Forterre,P. and Gribaldo,S. (2011) On the last common ancestor and early evolution of eukaryotes: reconstructing the history of mitochondrial ribosomes. *Res. Microbiol.*, **162**, 53–70.
9. Gray,M.W., Burger,G., Derelle,R., Klimeš,V., Leger,M.M., Sarrasin,M., Vlček,Č., Roger,A.J., Eliáš,M. and Lang,B.F. (2020) The draft nuclear genome sequence and predicted mitochondrial proteome of *Andalucia godoyi*, a protist with the most gene-rich and bacteria-like mitochondrial genome. *BMC Biol.*, **18**, 22.
10. Desai,N., Brown,A., Amunts,A. and Ramakrishnan,V. (2017) The structure of the yeast mitochondrial ribosome. *Science*, **355**, 528–531.
11. Greber,B.J., Bieri,P., Leibundgut,M., Leitner,A., Aebersold,R., Boehringer,D. and Ban,N. (2015) The complete structure of the 55S mammalian mitochondrial ribosome. *Science*, **348**, 303–308.
12. Itoh,Y., Naschberger,A., Mortezaei,N., Herrmann,J.M. and Amunts,A. (2020) Analysis of translating mitoribosome reveals functional characteristics of translation in mitochondria of fungi. *Nat. Commun.*, **11**, 5187.
13. Ramrath,D., Niemann,M., Leibundgut,M., Bieri,P., Prange,C., Horn,E., Leitner,A., Boehringer,D., Schneider,A. and Ban,N. (2018) Evolutionary shift toward protein-based architecture in trypanosomal mitochondrial ribosomes. *Science*, **362**, eaau7735.
14. Soufari,H., Waltz,F., Parrot,C., Durrieu-Gaillard,S., Bochler,A., Kuhnb,L., Sissler,M. and Hashem,Y. (2020) Structure of the mature kinetoplastids mitoribosome and insights into its large subunit biogenesis. *Proc. Natl. Acad. Sci. U.S.A.*, **117**, 29851–29861.
15. Tobiasson,V., Berzina,I. and Amunts,A. (2022) Structure of a mitochondrial ribosome with fragmented rRNA in complex with membrane-targeting elements. *Nat. Commun.*, **13**, 6132.
16. Tobiasson,V. and Amunts,A. (2020) Ciliate mitoribosome illuminates evolutionary steps of mitochondrial translation. *Elife*, **9**, 59264.
17. Waltz,F., Soufari,H., Bochler,A., Giegé,P. and Hashem,Y. (2020) Cryo-EM structure of the RNA-rich plant mitochondrial ribosome. *Nat. Plants*, **6**, 377–383.
18. Waltz,F., Salinas-Giegé,T., Englmeier,R., Meichel,H., Soufari,H., Kuhn,L., Pfeffer,S., Förster,F., Engel,B.D., Giegé,P. *et al.* (2021) How to build a ribosome from RNA fragments in *Chlamydomonas* mitochondria. *Nat. Commun.*, **12**, 7176.
19. Waltz,F. and Giegé,P. (2020) Striking diversity of mitochondria-specific translation processes across eukaryotes. *Trends Biochem. Sci.*, **45**, 149–162.
20. O'Brien,T. (2002) Evolution of a protein-rich mitochondrial ribosome: implications for human genetic disease. *Gene*, **286**, 73–79.
21. Valach,M., Gonzalez Alcazar,J.A., Sarrasin,M., Lang,B.F., Gray,M.W. and Burger,G. (2021) An unexpectedly complex mitoribosome in *Andalucia godoyi*, a protist with the most bacteria-like mitochondrial genome. *Mol. Biol. Evol.*, **38**, 788–804.
22. Valach,M., Burger,G., Gray,M. and Lang,B.F. (2014) Widespread occurrence of organelle genome-encoded 5S rRNAs including permuted molecules. *Nucleic Acids Res.*, **42**, 13764–13777.
23. Brown,A., Amunts,A., Bai,X.C., Sugimoto,Y., Edwards,P.C., Murshudov,G., Scheres,S.H.W. and Ramakrishnan,V. (2014) Structure of the large ribosomal subunit from human mitochondria. *Science*, **346**, 718–722.
24. Amunts,A., Brown,A., Bai,X., Llácer,J., Hussain,T., Emsley,P., Long,F., Murshudov,G., Scheres,S. and Ramakrishnan,V. (2014) Structure of the yeast mitochondrial large ribosomal subunit. *Science*, **343**, 1485–1489.
25. Feagin,J.E., Harrell,M.I., Lee,J.C., Coe,K.J., Sands,B.H., Cannone,J.J., Tami,G., Schnare,M.N. and Gutell,R.R. (2012) The fragmented mitochondrial ribosomal RNAs of *Plasmodium falciparum*. *PLoS One*, **7**, e38320.
26. Heinonen,T.Y., Schnare,M.N., Young,P.G. and Gray,M.W. (1987) Rearranged coding segments, separated by a transfer RNA gene, specify the two parts of a discontinuous large subunit ribosomal RNA in *Tetrahymena pyriformis* mitochondria. *J. Biol. Chem.*, **262**, 2879–2887.
27. Boer,P.H. and Gray,M.W. (1988) Scrambled ribosomal RNA gene pieces in *Chlamydomonas reinhardtii* mitochondrial DNA. *Cell*, **55**, 399–411.
28. Tashyreva,D., Simpson,A.G.B., Prokopchuk,G., Škodová-Sveráková,I., Butenko,A., Hammond,M., George,E.E., Flegontova,O., Záhonová,K., Faktorová,D. *et al.* (2022) Diplonemids – a review on 'new' flagellates on the oceanic block. *Protist*, **173**, 125868.
29. Kaur,B., Záhonová,K., Valach,M., Faktorová,D., Prokopchuk,G., Burger,G. and Lukeš,J. (2020) Gene fragmentation and RNA editing without borders: eccentric mitochondrial genomes of diplonemids. *Nucleic Acids Res.*, **48**, 2694–2708.
30. Valach,M., Moreira,S., Kiethega,G. and Burger,G. (2014) *Trans*-splicing and RNA editing of LSU rRNA in *Diplonema* mitochondria. *Nucleic Acids Res.*, **42**, 2660–2672.
31. Moreira,S., Valach,M., Aoulad-Aissa,M., Otto,C. and Burger,G. (2016) Novel modes of RNA editing in mitochondria. *Nucleic Acids Res.*, **44**, 4907–4919.
32. Valach,M., Moreira,S., Petitjean,C., Benz,C., Butenko,A., Flegontova,O., Nenarokova,A., Prokopchuk,G., Batstone,T., Lapébie,P. *et al.* (2023) Recent expansion of metabolic versatility in *Diplonema papillatum*, the model species of a highly speciose group of marine eukaryotes. *BMC Biol.*, **21**, 99.
33. Faktorová,D., Kaur,B., Valach,M., Graf,L., Benz,C., Burger,G. and Lukeš,J. (2020) Targeted integration by homologous recombination enables in situ tagging and replacement of genes in the marine microeukaryote *Diplonema papillatum*. *Environ. Microbiol.*, **22**, 3660–3670.
34. Kaur,B., Valach,M., Peña-Diaz,P., Moreira,S., Keeling,P.J., Burger,G., Lukeš,J. and Faktorová,D. (2018) Transformation of *Diplonema papillatum*, the type species of the highly diverse and abundant marine microeukaryotes Diplonemida (Euglenozoa). *Environ. Microbiol.*, **20**, 1030–1040.
35. Valach,M., Léveillé-Kunst,A., Gray,M.W. and Burger,G. (2018) Respiratory chain complex I of unparalleled divergence in diplonemids. *J. Biol. Chem.*, **293**, 16043–16056.
36. Rodríguez-Ezpeleta,N., Teijeiro,S., Forget,L., Burger,G. and Lang,B.F. (2009) Construction of cDNA libraries: focus on protists and fungi. *Methods Mol. Biol.*, **533**, 33–47.
37. Sambrook,J. and Russell,D.W. (2019) Isolation of DNA fragments from polyacrylamide gels by the crush and soak method. *Cold Spring Harb. Protoc.*, **2019**, https://doi.org/10.1101/pdb.prot100479.
38. Schägger,H. (2006) Tricine–SDS-PAGE. *Nat. Protoc.*, **1**, 16–22.
39. Scott,D.D., Trahan,C., Zindy,P.J., Aguilar,L.C., Delubac,M.Y., van Nostrand,E.L., Adivarahan,S., Wei,K.E., Yeo,G.W., Zenklusen,D. *et al.* (2017) Nol12 is a multifunctional RNA binding protein at the nexus of RNA and DNA metabolism. *Nucleic Acids Res.*, **45**, 12509–12528.
40. Oeffinger,M., Wei,K.E., Rogers,R., DeGrasse,J.A., Chait,B.T., Aitchison,J.D. and Rout,M.P. (2007) Comprehensive analysis of diverse ribonucleoprotein complexes. *Nat, Methods*, **4**, 951–956.
41. Puig,O., Caspary,F., Rigaut,G., Rutz,B., Bouveret,E., Bragado-Nilsson,E., Wilm,M. and Séraphin,B. (2001) The tandem

affinity purification (TAP) method: a general procedure of protein complex purification. *Methods*, **24**, 218–229.
42. Hulstaert,N., Shofstahl,J., Sachsenberg,T., Walzer,M., Barsnes,H., Martens,L. and Perez-Riverol,Y. (2019) ThermoRawFileParser: modular, scalable, and cross-platform RAW file conversion. *J. Proteome Res.*, **19**, 537–542.
43. Kong,A.T., Leprevost,F.v., Avtonomov,D.M., Mellacheruvu,D. and Nesvizhskii,A.I. (2017) MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry–based proteomics. *Nat. Methods*, **14**, 513–520.
44. da Veiga Leprevost,F., Haynes,S.E., Avtonomov,D.M., Chang,H.Y., Shanmugam,A.K., Mellacheruvu,D., Kong,A.T. and Nesvizhskii,A.I. (2020) Philosopher: a versatile toolkit for shotgun proteomics data analysis. *Nat. Methods*, **17**, 869.
45. Yu,F., Haynes,S.E., Teo,G.C., Avtonomov,D.M., Polasky,D.A. and Nesvizhskii,A.I. (2020) Fast quantitative analysis of timsTOF PASEF data with MSFragger and IonQuant. *Mol. Cell. Proteomics*, **19**, 1575–1585.
46. Schwanhüusser,B., Busse,D., Li,N., Dittmar,G., Schuchhardt,J., Wolf,J., Chen,W. and Selbach,M. (2011) Global quantification of mammalian gene expression control. *Nature*, **473**, 337–342.
47. Teo,G., Liu,G., Zhang,J., Nesvizhskii,A.I., Gingras,A.C. and Choi,H. (2014) SAINTexpress: improvements and additional features in Significance Analysis of INTeractome software. *J. Proteomics*, **100**, 37–43.
48. Mellacheruvu,D., Wright,Z., Couzens,A.L., Lambert,J.P., St-Denis,N.A., Li,T., Miteva,Y.v., Hauri,S., Sardiu,M.E., Low,T.Y. *et al.* (2013) The CRAPome: a contaminant repository for affinity purification–mass spectrometry data. *Nat. Methods*, **10**, 730–736.
49. Knight,J.D.R., Choi,H., Gupta,G.D., Pelletier,L., Raught,B., Nesvizhskii,A.I. and Gingras,A.C. (2017) ProHits-viz: a suite of web tools for visualizing interaction proteomics data. *Nat. Methods*, **14**, 645–646.
50. Jumper,J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronneberger,O., Tunyasuvunakool,K., Bates,R., Žídek,A., Potapenko,A. *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, **596**, 583–589.
51. Wheeler,R.J. (2021) A resource for improved predictions of *Trypanosoma* and *Leishmania* protein three-dimensional structure. *PLoS One*, **16**, e0259871.
52. Richter,D.J., Berney,C., Strassert,J.F.H., Poh,Y.-P., Herman,E.K., Muñoz-Gómez,S.A., Wideman,J.G., Burki,F. and de Vargas,C. (2022) EukProt: a database of genome-scale predicted proteins across the diversity of eukaryotes. *Peer Community Journal*, **2**, e56.
53. Eddy,S.R. (2011) Accelerated Profile HMM Searches. *PLoS Comput. Biol.*, **7**, e1002195.
54. Sievers,F., Wilm,A., Dineen,D., Gibson,T.J., Karplus,K., Li,W., Lopez,R., McWilliam,H., Remmert,M., Söding,J. *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.*, **7**, 539.
55. Katoh,K. and Standley,D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.*, **30**, 772–780.
56. Crooks,G.E., Hon,G., Chandonia,J.M. and Brenner,S.E. (2004) WebLogo: a sequence logo generator. *Genome Res.*, **14**, 1188–1190.
57. Price,M.N., Dehal,P.S. and Arkin,A.P. (2010) FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS One*, **5**, e9490.
58. Minh,B.Q., Schmidt,H.A., Chernomor,O., Schrempf,D., Woodhams,M.D., Haeseler,A., Lanfear,R. and Teeling,E. (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.*, **37**, 1530–1534.
59. Lorenz,R., Bernhart,S.H., Höner zu Siederdissen,C., Tafer,H., Flamm,C., Stadler,P.F. and Hofacker,I.L. (2011) ViennaRNA Package 2.0. *Algorith. Mol. Biol.*, **6**, 26.
60. Reuter,J.S. and Mathews,D.H. (2010) RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinf.*, **11**, 129.
61. Rodrigues,J.P.G.L.M., Teixeira,J.M.C., Trellet,M. and Bonvin,A.M.J.J. (2018) pdb-tools: a swiss army knife for molecular structures. *F1000Research*, **7**, 1961.
62. Pettersen,E.F., Goddard,T.D., Huang,C.C., Meng,E.C., Couch,G.S., Croll,T.I., Morris,J.H. and Ferrin,T.E. (2021) UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.*, **30**, 70–82.
63. Letunic,I. and Bork,P. (2018) 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res.*, **46**, D493–D496.
64. Finn,R.D., Coggill,P., Eberhardt,R.Y., Eddy,S.R., Mistry,J., Mitchell,A.L., Potter,S.C., Punta,M., Qureshi,M., Sangrador-Vegas,A. *et al.* (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.*, **44**, D279–D285.
65. Marchler-Bauer,A., Derbyshire,M.K., Gonzales,N.R., Lu,S., Chitsaz,F., Geer,L.Y., Geer,R.C., He,J., Gwadz,M., Hurwitz,D.I. *et al.* (2015) CDD: NCBI's conserved domain database. *Nucleic Acids Res.*, **43**, D222–D226.
66. The UniProt Consortium. (2022) UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.*, **51**, D523–D531.
67. Armstrong,D.R., Berrisford,J.M., Conroy,M.J., Gutmanas,A., Anyango,S., Choudhary,P., Clark,A.R., Dana,J.M., Deshpande,M., Dunlop,R. *et al.* (2020) PDBe: improved findability of macromolecular structure data in the PDB. *Nucleic Acids Res.*, **48**, D335–D343.
68. Lukeš,J., Wheeler,R., Jirsová,D., David,V. and Archibald,J.M. (2018) Massive mitochondrial DNA content in diplonemid and kinetoplastid protists. *IUBMB Life*, **70**, 1267–1274.
69. Marande,W., Lukeš,J. and Burger,G. (2005) Unique mitochondrial genome structure in diplonemids, the sister group of kinetoplastids. *Eukaryot. Cell*, **4**, 1137–1146.
70. Maslov,D.A., Spremulli,L.L., Sharma,M.R., Bhargava,K., Grasso,D., Falick,A.M., Agrawal,R.K., Parker,C.E. and Simpson,L. (2007) Proteomics and electron microscopic characterization of the unusual mitochondrial ribosome-related 45S complex in *Leishmania tarentolae*. *Mol. Biochem. Parasitol.*, **152**, 203–212.
71. Trahan,C., Aguilar,L.C. and Oeffinger,M. (2016) Single-step affinity purification (ssAP) and mass spectrometry of macromolecular complexes in the yeast *S. cerevisiae*. *Methods Mol. Biol.*, **1361**, 265–287.
72. Jaskolowski,M., Ramrath,D.J.F., Bieri,P., Niemann,M., Mattei,S., Calderaro,S., Leibundgut,M., Horn,E.K., Boehringer,D., Schneider,A. *et al.* (2020) Structural insights into the mechanism of mitoribosomal large subunit biogenesis. *Mol. Cell*, **79**, 629–644.
73. Saurer,M., Ramrath,D.J.F., Niemann,M., Calderaro,S., Prange,C., Mattei,S., Scaiola,A., Leitner,A., Bieri,P., Horn,E.K. *et al.* (2019) Mitoribosomal small subunit biogenesis in trypanosomes involves an extensive assembly machinery. *Science*, **365**, 1144–1149.
74. Tobiasson,V., Gahura,O., Aibara,S., Baradaran,R., Zíková,A. and Amunts,A. (2021) Interconnected assembly factors regulate the biogenesis of mitoribosomal large subunit. *EMBO J.*, **40**, e106292.
75. Gahura,O., Chauhan,P. and Zíková,A. (2022) Mechanisms and players of mitoribosomal biogenesis revealed in trypanosomatids. *Trends Parasitol.*, **38**, 1053–1067.
76. Lenarčič,T., Niemann,M., Ramrath,D.J.F., Calderaro,S., Flügel,T., Saurer,M., Leibundgut,M., Boehringer,D., Prange,C., Horn,E.K. *et al.* (2022) Mitoribosomal small subunit maturation involves formation of initiation-like complexes. *Proc. Natl. Acad. Sci. U.S.A.*, **119**, e2114710118.
77. Rozanska,A., Richter-Dennerlein,R., Rorbach,J., Gao,F., Lewis,R.J., Chrzanowska-Lightowlers,Z.M. and Lightowlers,R.N. (2017) The human RNA-binding protein RBFA promotes the maturation of the mitochondrial ribosome. *Biochem. Journal*, **474**, 2145–2158.
78. Schedlbauer,A., Iturrioz,I., Ochoa-Lizarralde,B., Diercks,T., Lopez-Alonso,J.P., Lavin,J.L., Kaminishi,T., Capuni,R., Dhimole,N., de Astigarraga,E. *et al.* (2021) A conserved rRNA switch is central to decoding site maturation on the small ribosomal subunit. *Sci. Adv.*, **7**, eabf7547.
79. Harper,N.J., Burnside,C. and Klinge,S. (2022) Principles of mitoribosomal small subunit assembly in eukaryotes. *Nature*, **614**, 175–181.
80. Itoh,Y., Khawaja,A., Laptev,I., Cipullo,M., Atanassov,I., Sergiev,P., Rorbach,J. and Amunts,A. (2022) Mechanism of mitoribosomal small subunit biogenesis and preinitiation. *Nature*, **606**, 603–608.
81. Guja,K.E., Venkataraman,K., Yakubovskaya,E., Shi,H., Mejia,E., Hambardjieva,E., Karzai,A.W. and Garcia-Diaz,M. (2013) Structural basis for S-adenosylmethionine binding and

methyltransferase activity by mitochondrial transcription factor B1. *Nucleic Acids Res.*, **41**, 7947–7959.
82. Liu,X., Shen,S., Wu,P., Li,F., Liu,X., Wang,C., Gong,Q., Wu,J., Yao,X., Zhang,H. *et al.* (2019) Structural insights into dimethylation of 12S rRNA by TFB1M: indispensable role in translation of mitochondrial genes and mitochondrial function. *Nucleic Acids Res.*, **47**, 7648–7665.
83. Brown,A., Rathore,S., Kimanius,D., Aibara,S., Bai,X.C., Rorbach,J., Amunts,A. and Ramakrishnan,V. (2017) Structures of the human mitochondrial ribosome in native states of assembly. *Nat. Struct. Mol. Biol.*, **24**, 866–869.
84. Lavdovskaia,E., Hillen,H.S. and Richter-Dennerlein,R. (2022) Hierarchical folding of the catalytic core during mitochondrial ribosome biogenesis. *Trends Cell. Biol.*, **32**, 182–185.
85. Feng,B., Mandava,C.S., Guo,Q., Wang,J., Cao,W., Li,N., Zhang,Y., Zhang,Y., Wang,Z., Wu,J. *et al.* (2014) Structural and functional insights into the mode of action of a universally conserved Obg GTPase. *PLoS Biol.*, **12**, e1001866.
86. Maiti,P., Lavdovskaia,E., Barrientos,A. and Richter-Dennerlein,R. (2021) Role of GTPases in driving mitoribosome assembly. *Trends Cell. Biol.*, **31**, 284–297.
87. Nikolay,R., Hilal,T., Schmidt,S., Qin,B., Schwefel,D., Vieira-Vieira,C.H., Mielke,T., Bürger,J., Loerke,J., Amikura,K. *et al.* (2021) Snapshots of native pre-50S ribosomes reveal a biogenesis factor network and evolutionary specialization. *Mol. Cell.*, **81**, 1200–1215.
88. Chandrasekaran,V., Desai,N., Burton,N.O., Yang,H., Price,J., Miska,E.A. and Ramakrishnan,V. (2021) Visualizing formation of the active site in the mitochondrial ribosome. *Elife*, **10**, e68806.
89. Cipullo,M., Gesé,G.V., Khawaja,A., Hällberg,B.M. and Rorbach,J. (2021) Structural basis for late maturation steps of the human mitoribosomal large subunit. *Nat. Commun.*, **12**, 3673.
90. Hillen,H.S., Lavdovskaia,E., Nadler,F., Hanitsch,E., Linden,A., Bohnsack,K.E., Urlaub,H. and Richter-Dennerlein,R. (2021) Structural basis of GTPase-mediated mitochondrial ribosome biogenesis and recycling. *Nat. Commun.*, **12**, 3672.
91. Seidman,D., Johnson,D., Gerbasi,V., Golden,D., Orlando,R. and Hajduk,S. (2012) Mitochondrial membrane complex that contains proteins necessary for tRNA import in *Trypanosoma brucei*. *J. Biol. Chem.*, **287**, 8892–8903.
92. Noller,H.F., Donohue,J.P. and Gutell,R.R. (2022) The universally conserved nucleotides of the small subunit ribosomal RNAs. *RNA*, **28**, 623–644.
93. Demeshkina,N., Jenner,L., Westhof,E., Yusupov,M. and Yusupova,G. (2012) A new understanding of the decoding principle on the ribosome. *Nature*, **484**, 256–259.
94. Loveland,A.B., Demo,G., Grigorieff,N. and Korostelev,A.A. (2017) Ensemble cryo-EM elucidates the mechanism of translation fidelity. *Nature*, **546**, 113–117.
95. Burger,G. and Valach,M. (2018) Perfection of eccentricity: mitochondrial genomes of diplonemids. *IUBMB Life*, **70**, 1197–1206.
96. Burger,G., Moreira,S. and Valach,M. (2016) Genes in hiding. *Trends Genet.*, **32**, 553–565.
97. Santos,B., Zeng,R., Jorge,S.F., Ferreira-Junior,J.R., Barrientos,A. and Barros,M.H. (2022) Functional analyses of mitoribosome 54S subunit devoid of mitochondria-specific protein sequences. *Yeast*, **39**, 208–229.
98. Gray,M.W., Lukeš,J., Archibald,J.M., Keeling,P.J. and Doolittle,W.F. (2010) Irremediable complexity?*Science*, **330**, 920–921.
99. Scaltsoyiannes,V., Corre,N., Waltz,F. and Giegé,P. (2022) Types and functions of mitoribosome-specific ribosomal proteins across eukaryotes. *Int. J. Mol. Sci.*, **23**, 3474.
100. Lill,R. and Freibert,S.A. (2020) Mechanisms of mitochondrial iron-sulfur protein biogenesis. *Annu. Rev. Biochem.*, **89**, 471–499.
101. Singh,J., Raina,R., Vinothkumar,K.R. and Anand,R. (2022) Decoding the mechanism of specific RNA targeting by ribosomal methyltransferases. *ACS Chem. Biol.*, **17**, 829–839.
102. Stephan,N.C., Ries,A.B., Boehringer,D. and Ban,N. (2021) Structural basis of successive adenosine modifications by the conserved ribosomal methyltransferase KsgA. *Nucleic Acids Res.*, **49**, 6389–6398.
103. Rebelo-Guiomar,P., Pellegrino,S., Dent,K.C., Sas-Chen,A., Miller-Fleming,L., Garone,C., van Haute,L., Rogan,J.F., Dinan,A., Firth,A.E. *et al.* (2022) A late-stage assembly checkpoint of the human mitochondrial ribosome large subunit. *Nat. Commun.*, **13**, 929.
104. Perez-Riverol,Y., Bai,J., Bandla,C., García-Seisdedos,D., Hewapathirana,S., Kamatchinathan,S., Kundu,D.J., Prakash,A., Frericks-Zipper,A., Eisenacher,M. *et al.* (2022) The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.*, **50**, D543–D552.