Mitochondrial ribosomal proteins developed unconventional mitochondrial targeting signals due to structural constraints

Yury S. Bykov[1,*], Tamara Flohr[2,*], Felix Boos[2,3], Johannes M. Herrmann[2,§], Maya Schuldiner[1,§]

1. Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel
2. Division of Cell Biology, University of Kaiserslautern, Germany
3. Current address: Department of Genetics, Stanford University, Stanford, CA

* Equal contribution

§ Correspondence to: maya.schuldiner@weizmann.ac.il, hannes.herrmann@biologie.uni-kl.de

Keywords: mitochondria, targeting, translocation, mitochondrial ribosome, mitochondrial targeting signal, ribosome structure

Abstract

Mitochondrial ribosomes are complex molecular machines indispensable for respiration. Their assembly involves the import of several dozens of mitochondrial ribosomal proteins (MRPs), encoded in the nuclear genome, into the mitochondrial matrix. Available proteomic and structural data as well as computational predictions indicate that up to 25% of MRPs do not have a conventional N-terminal mitochondrial targeting signal (MTS). We characterized a set of 15 yeast MRPs *in vivo* and showed that 30% of them use internal mitochondrial targeting signals. We isolated a novel internal targeting signal from the conserved MRP Mrp17 (bS6). The Mrp17 targeting signal shares some properties as well as import components with conventional MTS-containing proteins but is not reliably predicted indicating that mitochondrial protein targeting is more versatile than expected. We hypothesize that internal targeting signals arose in MRPs when the N-terminus extension was constrained by ribosome assembly interfaces.

Introduction

Mitochondria are descendants of ancient bacteria that formed eukaryotic cells together with their archaeal host (Martijn et al., 2018; Sagan, 1967; Zaremba-Niedzwiedzka et al., 2017). Since then, mitochondria have lost their autonomy and their reproduction depends entirely on the nuclear genome, which encodes the majority of mitochondrial proteins. However all mitochondria capable of respiration have retained small vestigial genomes of their own and fully functional gene expression machineries of bacterial origin (Roger et al., 2017). Mitochondrial ribosomes (mitoribosomes) are the most complex components of the mitochondrial gene expression system and consist of several RNA molecules and 60 to 80 different proteins (Greber and Ban, 2016; Ott et al., 2016). Mitoribosome dysfunction has adverse consequences leading to a broad spectrum of diseases (Boczonadi and Horvath, 2014).

While it took many years to solve the first ribosome structures (Ban et al., 2000; Carter et al., 2000; Schluenzen et al., 2000), rapid progress in cryo-electron microscopy is now rapidly revealing the structural details of mitoribosomes of many different organismic groups (Amunts et al., 2015; Desai et al., 2017; Itoh et al., 2020; Kummer et al., 2018; Ramrath et al., 2018; Tobiasson and Amunts, 2020; Waltz et al., 2020). The availability of so many structures highlighted an interesting feature of mitoribosomes - their incredible evolutionary diversity (Kummer and Ban, 2021; Waltz and Giegé, 2019). The composition of mitochondrial ribosomes in different eukaryotic lineages underwent dramatic changes caused by multiple losses of RNA segments and mitoribosomal proteins (MRPs) as well as acquisition of new, lineage-specific RNA segments and MRPs (Desmond et al., 2011; Petrov et al., 2019; Sluis et al., 2015; Smits et al., 2007). As a result, mitoribosomes contain a core set of MRPs homologous to the bacterial ribosomal proteins (BRPs) and a variable set of MRPs that can be common for all mitochondrial ribosomes or specific only to certain eukaryotic lineages. In addition, during their evolution many MRPs acquired significant expansions of their C- and N-termini while retaining structurally conserved domains of their BRP ancestors (Melnikov et al., 2018; Sluis et al., 2015; Vishwanath et al., 2004).

Mitochondrial genomes in many eukaryotic cells (with the exception of animals) still contain genes for a number of ribosomal proteins, indicating that their successful transfer to the nuclear genome might be less easily feasible than that of many other matrix proteins (Bertgen et al., 2020). However, most, in animals even all, MRPs are nuclear encoded. Thus, similarly to the majority of mitochondrial proteins (numbering from around 800 in yeast to around 1500 in mammals), they must be imported from the cytosol (Morgenstern et al., 2017; Pagliarini et al., 2008; Vögtle et al., 2017).

The import of mitochondrial proteins can be conceptually subdivided in two steps: (1) targeting of the newly synthesized mitochondrial protein precursors to the mitochondrial membrane. This can occur either post-translationally or co-translationally involving ribosome-nascent chain complexes. (2) Translocation of the unfolded precursors through the mitochondrial membrane(s) to deliver them to their final destination within mitochondria (Bykov et al., 2020). Effective targeting and translocation are mediated by specialized protein complexes that recognize targeting and translocation signals within precursor protein sequences.

Most matrix and inner membrane proteins are synthesized with N-terminal matrix targeting sequences (MTSs), also called presequences, which are both necessary and sufficient for mitochondrial targeting. MTSs have a characteristic structure that can be predicted by prediction algorithms (Armenteros et al., 2019; Claros and Vincens, 1996; Emanuelsson et al., 2000; Fukasawa et al., 2015). MTSs are typically between 10 and 60 residues in length and can form an amphipathic α-helix with one positively charged surface and one hydrophobic surface. In most cases MTSs are proteolytically removed during protein import, giving rise to mature forms of mitochondrial matrix or inner-membrane proteins (Bedwell et al., 1989; Vögtle et al., 2009; von Heijne, 1986).

In contrast to basically all other proteins of the mitochondrial matrix, many MRPs lack N-terminal MTSs (Woellhaf et al., 2014). In some cases, MRPs use N-terminal regions that mimic the properties of MTSs but are not cleaved (un-cleaved MTSs). Such un-cleaved MTSs are also found in some matrix proteins that are not associated with the ribosome, such as Hsp10 (Poveda-Huertes et al., 2020). Surprisingly, a number of MRPs do not contain any regions that show MTS-like features and it is unknown how mitochondria recognize and import these proteins. For now, there are only two well characterized examples of MRPs with unconventional MTSs – Mrpl32 (bL32, by new nomenclature (Ban et al., 2014)) and Mrp10 (mS37) whose import path deviates from the canonical matrix-targeting route (Bonn et al., 2011; Longen et al., 2014; Nolden et al., 2005).

In this study, we studied the mechanisms by which MRPs are imported and assembled into the mitoribosome. We systematically examined N-termini of unconventional MRPs and analyzed them *in silico* and experimentally. We further focused on the biogenesis of Mrp17 (bS6) as a representative of the unconventional group of MTS-less MRPs. We discovered a novel mitochondrial matrix targeting region that is displayed in the internal sequence of the protein. This stretch shares properties with mitochondrial targeting sequences such as positive charges for receptor binding and membrane potential-dependent translocation, but differs in its structural features and position in the protein. The efficient import of Mrp17 shows that the mitochondrial import machinery is much more versatile in its substrate

spectrum than expected. More generally, our work shows how structural restrictions forced the generation of unconventional targeting motifs.

Results

Mapping unconventional MRP targeting signals

To systematically investigate MRP targeting signals in detail, we compiled all existing data on the maturation of their N-termini in yeast (Table S1). We used direct N-terminal sequencing data (Boguta et al., 1992; Dang and Ellis, 1990; Davis et al., 1992; Graack et al., 1988; Graack et al., 1991; Grohmann et al., 1989; Grohmann et al., 1991; Kitakawa et al., 1990; Kitakawa et al., 1997; Matsushita and Isono, 1993; Matsushita et al., 1989), N-terminal proteomics (Vögtle et al., 2009) and predictions performed by UniProt annotators, as well as by ourselves using MitoFates for cleavage site prediction (Fukasawa et al., 2015). Importantly, we also used available structural information (Desai et al., 2017). In particular, mitoribosome structures were helpful to identify proteins that do not have a cleavable MTS – such proteins had their N-termini contained within the structure and hence could not have been cleaved after import into the mitochondrial matrix. We reanalyzed ribosome profiling data on translation initiation in yeast (see Methods for details) to ascertain that none of these proteins has mis-annotated translation start sites that might produce an N-terminal extension accounting for a missing cleavable MTS (Fig. S1). In the yeast mitochondrial ribosome structure (PDB:5MRC), the structures of six proteins started with amino acid number 1 (Met), structures of 12 started with amino acid number 2, five – with amino acids 3 to 9, and the rest, 50, with amino acid number 10 and more. The number of the first amino acid present in the structure was moderately conserved among the determined mitoribosome structures (Fig. S2) and was not restricted to any particular group of MRPs classified by origin (bacterial, mitochondria-specific, or yeast-specific) or position in the structure (Fig. S3).

Interestingly, a simple distinction by the first amino acid appearing in the structure separates MRPs into two classes. In the first group are those MRPs that are derived from cleaved precursors (which consistently have high MTS prediction scores). In addition, this group may contain proteins with an uncleavable N-terminus of a flexible nature which would then be unresolved in the available structures. Some of the latter may have poor mitochondrial targeting scores in prediction algorithms. In the second group are those whose structure starts with amino acid number less than 10. Most of these proteins score very poorly with different software predicting N-terminal MTS (Fig. 1A, Fig. S2). Many MRPs of this group lack conventional, N-terminal import signals, and their targeting signals are not predictable by available software. Thus, the available structures of mitochondrial ribosomes confirm the

previous conclusion that many MRPs are made without N-terminal MTSs (Woellhaf et al., 2014).

Next, we experimentally analyzed the targeting information in the sequences of different MRPs by GFP fusion proteins. To this end, we selected 15 MRPs with different properties (Fig. S3, Table S2). Then we tested whether the N-terminal 30 residues of these proteins were necessary and/or sufficient for mitochondrial targeting. The length of 30 residues was chosen as it corresponds to the most common size of a cleavable yeast MTS (Vögtle et al., 2009). To test this, we expressed each MRP in diploid yeast fused to GFP. To assay if the N-terminus is necessary we expressed a truncated version with the first 30 amino acids deleted ($MRP_{\Delta 30}$-GFP). To test if the N-terminus is sufficient we expressed a version with only the first 30 amino acids (MRP_{1-30}-GFP). As a control we used the full-length version (MRP_{Full}-GFP) (Fig. 1B). The distribution of GFP signals was imaged in cells in which mitochondria were stained with MitoTracker Orange (Fig. 1C-E, Fig S4, Table S2). Six proteins (Mrpl15, Rsm26, Rsm18, Rsm25, Mrpl40, and Rsm27) contained targeting information within their N-termini (Fig. 1C); of them, only Mrpl15 (mL57) had high MTS prediction scores consistent with highly confident annotation of a cleavable 29-amino acid long MTS (Table S1). Other proteins whose N-termini were able to target GFP to mitochondria had low MTS prediction scores (Table S2) indicating that their N-terminal signals have distinct properties, not similar to conventional MTSs. For four proteins (Mrps16, Mrp51, Mrpl38, and Mrpl28) neither the N-terminal 30 residues nor the internal segment on its own were sufficient for targeting, indicating that the necessary targeting information is contained in an N-terminal segment longer than 30 amino acids or distributed over the whole length of these proteins (Fig. 1D). Finally, five proteins (Mrp17, Pet123, Mrpl23, Mrp35, and Mrp20) were targeted to mitochondria independently of their N-terminal regions indicating that the targeting signals in these proteins are internal (Fig. 1E).

Interestingly, many of the N-terminally truncated MRP versions accumulated outside of the mitochondria in the cytosol or, in many cases, in the nucleus (Fig. 1C-E, Fig. S4B). These observations agree with the recent discovery that mistargeted mitochondrial proteins can accumulate in the nucleus and get degraded in perinuclear puncta (Shakya et al., 2021). Despite the mislocalization of several of these forms, none of them resulted in obvious growth defects (Fig. S5).

To summarize, we selected a subset of MRPs with diverse structural and sequence features and characterized the mitochondrial targeting capacity of their N termini. We observed that many of these MRPs contain multiple unconventional targeting signals, often outside of the 30 N-terminal residues, thus apparently scattered over their sequence. One particularly intriguing MRP was Mrp17 (bS6), a protein of the small subunit of the yeast mitoribosome.

Mrp17 lacks any identifiable targeting signal while being one of the most structurally conserved proteins across all mitoribosomal structures studied to date. Hence, we chose Mrp17 for further investigation.

Defining Mrp17 targeting and translocation signals

To investigate the unconventional mitochondrial targeting signals of Mrp17 in more detail we created a systematic set of Mrp17 truncations fused to GFP and expressed them in diploid yeast (Fig. S6). We observed that the internal fragment of Mrp17 between amino acids 20 and 100 was the minimal fragment able to target GFP to mitochondria similarly to full-length Mrp17 (131 amino acids) without producing cytosolic background signal (Fig. 2A). This indicates that similarly to the N-terminus, the C-terminus is dispensable for targeting. Splitting this fragment in two halves showed that the N-terminal part ($Mrp17_{21-60}$) was still able to target GFP to mitochondria although with significant cytosolic background while the C-terminal part ($Mrp17_{61-100}$) was cytosolic (Fig. 2A). We conclude that, *in vivo,* Mrp17 region 21-60 is necessary for mitochondrial targeting but is not sufficient for efficient targeting, which is promoted by additional signals distributed over the whole length of the protein (Fig. S7A,B).

The microscopic analysis does not allow us to discriminate between targeting to the mitochondrial surface from complete translocation into the matrix. To elucidate the import efficiency of different Mrp17 regions we used *in vitro* translocation assays into isolated yeast mitochondria. Since Mrp17 is very small and many fragments lacked methionine residues that are necessary for radiolabeling, we fused Mrp17 to an unfolded mutant of the mouse dihydrofolate reductase - $DHFR_{mut}$ (Vestweber and Schatz, 1988). The full-length Mrp17-$DHFR_{mut}$ fusion was effectively imported into isolated yeast mitochondria at the same rate as untagged Mrp17 but gave much stronger signals in autoradiography (Fig. 2B, Fig. S7C). In agreement with the targeting experiments performed *in vivo*, the short N-terminal region of Mrp17 was neither necessary nor sufficient for efficient translocation ($Mrp17_{21-131}$, $Mrp17_{1-20}$ in Fig. 2B). The first 60 amino acids of Mrp17 were sufficient for translocation narrowing down the import signal to the N-terminal half of the protein ($Mrp17_{1-60}$ in Fig. 2B). Leaving only the first 40 amino acids or removing them from the N-terminus reduced the translocation speed indicating that regions 20-40 and 40-60 are equally important parts of the signal ($Mrp17_{1-40}$ and $Mrp17_{41-131}$ in Fig. 2B). Finally, we narrowed down the Mrp17 region containing the translocation signal to amino acids 30-60 (Fig. 2B, bottom; Fig. S7D). However, similarly to the results of *in vivo* experiments, even short fragments of Mrp17 outside this region retained some translocation capacity (Fig. 2B, Fig. S7C).

To summarize, we determined that the main mitochondrial targeting and translocation signal of Mrp17 is positioned between amino acids 30 to 60. However, there exist additional signals that improve mitochondrial targeting efficiency or stability *in vivo*. These additional signals reside in the C-terminal half of Mrp17. This again indicates, that the mitochondrial targeting regions are scattered over the Mrp17 sequence, and for this protein, the N-terminal region is irrelevant for efficient mitochondrial import.

Characterizing the features of the Mrp17 targeting and translocation signal

Next, we analyzed the unconventional internal targeting region of Mrp17 located between residues 30-60 in more detail. Standard prediction algorithms do not find an MTS-like sequence in this region (Fig. S3). The Mrp17 structure mostly contains beta-strands in this region and only a part of a helical stretch (Fig. 3A). Mrp17 is generally rich in positive charges (its pI is 10.5) and a high content of positive charges are a general feature of ribosomal proteins that interact with negatively charged mRNA. Interestingly, during evolution, the positive charges in MRPs (and particularly their lysine content) was further increased suggesting that positive charges might play a role beyond their relevance for neutralizing the negative charges of ribosomal RNA (Fig. S8A, B).

To find out whether mitochondrial targeting and import require the positive charge and if so, whether there is a specific dependence on lysines, we constructed mutants of Mrp17 with all lysines substituted with arginines thus maintaining the charge ($Mrp17_{K\text{-}R}$) or all lysines substituted for alanines thus decreasing the charge ($Mrp17_{K\text{-}A}$). The $Mrp17_{K\text{-}R}$ mutant fused to $DHFR_{mut}$ was efficiently imported into isolated mitochondria, while the $Mrp17_{K\text{-}A}$ mutant was not (Fig. 3B). Similarly to the results obtained *in vitro*, $Mrp17_{K\text{-}R}$ fused to GFP and expressed in yeast colocalized with mitochondria as the wild type Mrp17 while $Mrp17_{K\text{-}A}$-GFP remained cytosolic (Fig. 3C, Fig. S8C). Thus, the positive charge is an important feature of the Mrp17 translocation signal. We also constructed Mrp17 mutants with lysines substituted with alanines only in the 1-60 or 30-60 regions. These mutants fused to $DHFR_{mut}$ were also not imported indicating that the positive charges indeed must be positioned in the targeting signal region (Fig. 3D). Interestingly, $Mrp17_{K\text{-}R}$ was able to rescue the growth defect of Δ*mrp17* cells on respiratory media suggesting that lysines per se are not essential for import and assembly of Mrp17, but the positive charges are (Fig. 3E).

Using mutagenesis, we determined that positive charge, most critically between amino-acids 30-60, is important for the function of Mrp17 targeting and translocation signal.

Determining the import pathway of Mrp17

Since Mrp17 lacks the typical features and targeting signals of matrix proteins, we wondered how this protein was first recognized by mitochondria and later imported. To test the receptor requirement for the mitochondrial import of Mrp17, we treated isolated yeast mitochondria with trypsin that removes all the import receptors of the Translocate of the Outer Membrane (TOM) complex from the mitochondrial surface but spares membrane-embedded TOM subunits (Ohba and Schatz, 1987). Trypsin treatment strongly inhibits the import of MTS containing matrix proteins such as Atp1 (Fig. 4A) but does not affect the import of most intra-membrane space (IMS) proteins (Gornicka et al., 2014; Lutz et al., 2003). The import of Mrp17-DHFR$_{mut}$ was even more sensitive to trypsin treatment than the import of Atp1 (Fig. 4A). Thus, Mrp17 import into mitochondrial matrix strongly depends on the presence of the mitochondrial outer membrane receptors.

To determine the dependence of Mrp17 import on individual receptors we purified mitochondria from yeast lacking Tom20 (Δ*tom20*), the main receptor for recognition of canonical MTSs. Mrp17-DHFR$_{mut}$ import into Δ*tom20* mitochondria was strongly reduced (Fig. 4B,C) indicating that Tom20 is involved in the import process. Deletion of Tom70 and its lowly expressed paralog Tom71 had no effect on the import of Mrp17-DHFR$_{mut}$ (Fig. S9), consistent with the absence of internal MTLS-like sequences (iMTS-Ls, (Backes et al., 2018)) in the protein (Fig. S3).

Since Tom20 deletion can affect the import of many mitochondrial proteins, its effect on Mrp17 can be indirect. However, we could find two potential motifs for binding Tom20 (ΦXXΦΦ, where Φ is hydrophobic amino acid and X – any amino acid; Fig. 3A)(Muto et al., 2001; Obita et al., 2003). To test whether the requirement for Tom20 binding is direct we mutated the Tom20-binding motifs in the Mrp17 translocation signal. For this, we constructed a mutant where each amino acid of both Tom20 binding motifs was substituted with alanine – Mrp17$_{\Delta\Phi XX\Phi\Phi}$. Indeed, the mutant fused to DHFR$_{mut}$ was not imported into isolated mitochondria supporting an essential role of these Tom20-binding motifs for import (Fig. 4D). However, the residual translocation of Mrp17-DHFR$_{mut}$ in Δ*tom20* compared to trypsinized mitochondria suggests that Mrp17 can use additional receptors, albeit at lesser efficiency. Such a bypass receptor could potentially be Tom22 which was shown to work in concert with Tom20 in MTS recognition (Yamano et al., 2008).

Next, we used the temperature-sensitive mutants *tim17-ts* and *tim22-ts* that reduce the protein import along either of the two Translocase of the Inner Membrane (TIM) pathways – the TIM23 and TIM22 pathways, respectively. Mrp17 import into *tim17-ts* mitochondria was

reduced whereas its import into *tim22-ts* mitochondria was not affected showing that Mrp17 uses the TIM23 pathway, similar to canonical MTS-containing proteins (Fig. 4E).

To summarize, Mrp17 has an internal targeting signal that shares some properties such as positive charge with a regular MTS. Indeed, its import pathway is similar to MTS containing proteins suggesting that mitochondrial components recognize it as a *bona fide* MTS yet even state-of-the-art prediction algorithms do not, suggesting that we lack information on certain MTS characteristics (Fig. 4F).

Comparison of Mrp17 to its bacterial homolog

Unlike many other core MRPs that significantly extended their structures with insertions and N/C-terminal expansions, Mrp17 maintained the overall structure of its bacterial ancestors (Fig. 5A, Fig. S10). This means that in the course of evolution, the acquisition of an N-terminal MTS was either not possible so that the Mrp17 mitochondrial targeting signal had to be accommodated within the existing “bacterial” structure, or the ribosomal protein was already predisposed for mitochondrial import. Since RNA-binding regions show similarity to mitochondrial targeting signals with both requiring positively charged and hydrophobic amino acids – this was an appealing hypothesis.

To investigate both possibilities, we compared average MTS prediction scores of Mrp17 (and other structurally conserved MRPs) to their closest bacterial homologs with known structure, i. e. *Escherichia coli* ribosomal proteins (Fig. 5B, Fig. S10). Two out of the 10 structurally conserved proteins that lacked cleavable MTS, indeed had much higher MTS prediction scores in mitoribosomes compared to bacterial ribosome meaning that uncleavable MTS-like signals were developed at their N-termini. For other proteins, both homologs had equally low scores (less than 0.5) indicating that some sequence properties of bacterial and mitoribosomal proteins are similar.

Next, we tested if these bacterial proteins already have properties that allow them to be imported into mitochondria *in vivo* by expressing them as GFP fusions in yeast (Fig. 5C, Fig. S11A). None of the proteins localized to mitochondria showing that these BRPs do not have an intrinsic mitochondrial targeting capacity and that targeting signals were incorporated into conserved MRP structures in the course of evolution. Interestingly, two of the proteins (EcL13 and EcL14) had an intrinsic nuclear targeting capacity which was tolerated by the cells (Fig. S11B).

We aimed to check if any particular sequence or structural features present in Mrp17 compared to its *E. coli* homolog, protein S6 (EcS6) can be responsible for mitochondrial

targeting. Most prominent of these features are two loop extensions (the first one is not in the range of amino acids 30-60 but the second is at the end of it and contains additional positive charges) and the two Tom20-binding motifs (Fig. 5A,D). We therefore introduced these four Mrp17 features into the homologous regions of EcS6 altogether or in different combinations, but none supported mitochondrial targeting of the chimeric constructs fused to GFP (Fig 5D, Fig. S11C). *In vitro*, adding to EcS6 the whole targeting sequence of Mrp17 (from amino acids 30 to 60, sufficient for translocation of $DHFR_{mut}$) was still not sufficient for translocation of EcS6-$DHFR_{mut}$ when fused to it at the N-terminus (Fig. 5E) suggesting that the overall negative charge of EcS6 (-6.5, without C-terminal glutamines) strongly inhibits mitochondrial import of chimeric constructs. When the same domain, $Mrp17_{30-60}$, was introduced into its homologous position in EcS6 and fused to $DHFR_{mut}$, such chimeric construct showed some import capacity (Fig. 5E). This indicates that in addition to the charge effects, the import signal contained in $Mrp17_{30-60}$ is also dependent on its spatial context to be functional.

Based on the comparison with the bacterial homolog, we suggest that the main features ensuring mitochondrial import of Mrp17 are the Tom20-binding motifs and its overall positive charge both of which were acquired without significant structural rearrangements.

Discussion

In this work we collected the available information on the MTSs of yeast MRPs from proteomic and structural studies as well as MTS prediction algorithms. Using *in vivo* structure-function assays we found that some MRPs possessed internal targeting signals that can be poorly predicted. We characterized the internal targeting signal of Mrp17 in more detail and found that it shares some features and import pathway with regular MTS-containing proteins.

Why would Mrp17 and other MRPs rely on internal signals instead of evolving a canonical N-terminal MTS? For Mrp17 and other MRPs that originated from bacterial ancestors and have not acquired any additional sequence extensions – they may have been simply because they did not need to evolve new signals as they already possessed mitochondrial import capacity due to their positive charge and hydrophobicity that are needed to bind ribosomal RNA. Such intrinsic import capacity was indeed shown for some *E. coli* proteins (Lucattini et al., 2004). However, in our work even strongly positively charged *E. coli* ribosomal proteins S12, S16, L13, and L14, although structurally conserved with their MRP homologs, were not targeted to mitochondria when expressed in yeast suggesting no initial targeting predisposition (Fig. 5C). Instead, we propose that the N-termini of these proteins were too

conserved to be extended with a cleavable MTS. So proteins, such as Mrp17, had to accommodate a targeting signal elsewhere in their sequences.

The N-terminus of Mrp17 is tightly positioned on the contact with neighboring proteins (Fig. 6A, Fig. S12A). If this interface arose before the final maturation of the mitochondrial protein import system, it would create challenges for the evolution of an N-terminal targeting signal. In principle, such and engagement of the N-terminus in a structural interface can still be kept with a cleavable MTS if proteases cleave the protein precisely before the first structurally conserved residue. Indeed, some MRPs have evolved in that way (Fig. S10). We hypothesize that for Mrp17, however, such precise cleavage could not evolve. First reason for that is that the two first amino acids were too conserved to be adjusted for a cleavage site. Second, an intermediate stage of cleavable MTS evolution, which could be a sub-optimally cleaved N-terminus with few extra amino acids, could not be accommodated in the structure, so it never got to evolve a precisely cleaved MTS. Instead, Mrp17 remodeled the charge for the whole protein and developed internal targeting signals. Interestingly, the main signal identified in our work was not positioned in structurally distinct areas that are different between Mrp17 and EcS6 (Fig. 5A,D; loop expansions, C-terminal region) but mostly distributed between amino acids 30-60 which form a lot of new protein-protein contacts within the mitoribosome compared to the bacterial ribosome (Fig. S12A,B). So, the Mrp17 targeting signal might have co-evolved with protein-protein interactions which is similar to the evolution of nucleolar-targeting signals in cytosolic ribosomal proteins (Melnikov et al., 2015).

Can structural constraints also explain why other MRPs did not develop cleavable MTSs? To assay that, we divided all MRPs into three groups according to the relative position of their MTS cleavage site (if present) and the first structured residue (Fig. 6B). We hypothesized that despite being involved in structure formation, the proteins that developed an MTS precisely cleaved off before the first structured residue might have N-termini less buried in the structure that could provide some flexibility for MTS cleavage evolution compared to the proteins that failed to develop an MTS and have a deeply buried N-terminus, like Mrp17. However, this was not the case and these two groups had their N-termini equally deeply buried in the structure (Fig. 6C, “No MTS” and “Precise MTS” groups). The validity of such measurement was confirmed by the fact that the MRP N-termini with few more unstructured residues before the structure start are indeed more often positioned on the ribosome surface (Fig. 6C, “Other” group). Thus, there is no strict limitation for a buried N-terminus to develop an MTS.

Interestingly, we found that the N-termini without a cleavable MTS had more protein and less RNA around them in the ribosome structure with Mrp17 being an extreme example (Fig.

382 6B,C). This suggests that unlike protein-protein contacts, protein-RNA contacts can be more
383 permissible to the change in the protein component that will allow the adjustment of the N-
384 terminus for protease cleavage. Our observation agrees with the study of sequence
385 conservation of ribosomal proteins that revealed that amino acids on the protein-RNA
386 interfaces are indeed less conserved than the amino acids on the protein-protein interfaces
387 (Pilla and Bahadur, 2019). Supposedly, RNA-protein interfaces are the most ancient and
388 important for ribosome assembly, so their smaller conservation compared to protein-protein
389 interfaces might seem counterintuitive (Fox, 2010). However, this could be explained by the
390 nature of protein-RNA binding that often relies on positively charged and hydrophobic
391 amino-acids while the fine shape of the protein surface is not so important. On the other
392 hand, protein-protein interfaces very much rely on the exact match of surface shapes which
393 makes amino-acid substitutions much more detrimental for the interface integrity.

394 Still, involvement in protein-protein contacts does not fully explain the absence of a
395 cleavable MTS in some less deeply buried MRPs (Fig. 6C) indicating that other factors might
396 be at play. Some possible explanations can be that a new MRP is recruited to the ribosome
397 with its targeting signal already positioned elsewhere in its sequence. Another option is that
398 a recruited protein has a normal cleavable MTS but after recruitment, the MTS loses its
399 cleavage site and becomes an integral part of the structure still positioned on the ribosome
400 surface. It is also possible that in some MRPs the flexible N-terminal extensions play other
401 important roles incompatible with N-terminal targeting signal properties. In the future, more
402 systematic analysis of mitoribosome structures from different organisms can shed light on
403 the complex interplay between the evolution of mitoribosome assembly on one side and the
404 targeting signals of its components on the other.

405 To summarize, we suggest that at least some of the non-canonical targeting signals can
406 develop in MRPs not only to fulfill certain functions, such as in the case of Mrp10 and Mrpl32
407 (Bonn et al., 2011; Longen et al., 2014), but under the pressure of structural constraints
408 imposed by the mitoribosome assembly. It seems that unlike OXPHOS complexes, which
409 also underwent complex evolution in mitochondria (Sluis et al., 2015), protein components of
410 mitoribosome more often use their N-termini to establish important interactions and thus
411 could not so easily develop a cleavable MTS (Fig. 6D). Instead, they developed multiple
412 internal targeting signals of different strengths such as those found in Mrp17. This might
413 have been relatively easy for ribosomal proteins that either already contain a lot of positive
414 charges, or can easily increase their content due to abundant protein-RNA contacts. An
415 alternative strategy would be to place a targeting signal on the C-terminus. For now, the only
416 known matrix protein with a C-terminal targeting signal is Hmi1 that also cannot tolerate an

N-terminal extension (Lee et al., 1999). Such diversity of mitochondrial targeting signals highlights the need for better targeting signal prediction algorithms.

In this work we have shown that many MRPs have unconventional, non-cleavable targeting signals that can use the same import pathway as a regular MTS. This demonstrates an incredible versatility of the mitochondrial import system that can accommodate such a range of substrates and poses the question what is the mechanistic basic of balancing such versatility with the specificity of protein import process.

Acknowledgements

We are grateful to Prof. Ada Yonath and Dr. Anat Bashan for their valuable comments and discussion, and to Victor Tobiasson and Alexey Amunts for their advice on structure analysis. We thank Amir Fadel for technical help with microscopy and growth assays, and Nikolaus Pfanner for temperature-sensitive yeast mutants. We thank Einat Zalckvar for critical reading of the manuscript.

Collaborative work between Herrmann and Schuldiner labs is supported by a DIP collaborative grant (MitoBalance). Work in the Schuldiner lab is supported by an ERC CoG (OnTarget 864068). We are grateful for funding from the Deutsche Forschungsgemeinschaft (HE2803/9-1 to J.M.H.). Y.B. is supported by EMBO Long-term postdoctoral fellowship. M.S in an incumbent of the Dr. Gilbert Omenn and Martha Darling Professorial Chair in Molecular Genetics. The robotic system of the Schuldiner lab was purchased through the kind support of the Blythe Brenden-Mann Foundation.

Materials and Methods

Yeast strains and plasmids

All yeast strains used for fluorescent protein expression and imaging were constructed on the BY4741 haploid or BY4743 diploid background (Brachmann et al., 1998). For mitochondria purification we used W303 background and for growth rescue with plasmid shuffling – YPH499 background. The strains used in this study are listed in Table S2. All strains were constructed using standard LiAc/ssDNA/PEG based transformation protocol (Gietz and Woods, 2006). For transformation we used the standard plasmids for PCR-based tagging and knock outs (Janke et al., 2004; Longtine et al., 1998) and plasmids generated in this study (Table S3). Mutant versions of Mrp17 genes and genes for bacterial ribosomal proteins optimized for expression in yeast were ordered from GeneWiz or Genescript. When strains were constructed by genomic integration, primers were designed using

Primers4Yeast (http://wws.weizmann.ac.il/Primers-4-Yeast)(Yofe and Schuldiner, 2014) or manually if the gene was truncated. Primers used in this study are listed in Table S4.

Yeast growth

Yeast cells were grown on either liquid media or solid media that contained 2.2% agar. For selection only for antibiotic resistance yeast cells were grown on YPD media (2% peptone, 1% yeast extract, 2% glucose) supplemented with nourseothricin (NAT) to 0.2 g/l. For auxotrophic selections yeast were grown in SD media (0.67% yeast nitrogen base without amino acids and with ammonium sulfate, 2% glucose, and OMM amino acid mix (Hanscho et al., 2012)) if necessary supplemented with the same amount of antibiotic.

Growth assays

For growth assays with plate reader the strains were inoculated in 96-well plate 1 day before the experiment start if grown in fermentative media (YPD, see above) or 2 days before the experiment if grown in fermentative media YPGlycerol (2% peptone, 1% yeast extract, 2% glycerol) at 30˚C with 500 rpm shaking in automated Liconic incubator. On the day of experiment the saturated cultures were diluted 1:50 in fresh media using EVO Freedom liquid handler (Tecan) and incubated at 30˚C with shaking. The optical density measurements at 600 nm were taken every 30 min with the SPARK plate reader (Tecan). Each assay was repeated 2 times (Fig. S5).

For drop dilution assay, the yeast cultures in the phase of exponential growth (OD_{600} ~ 0.6) were sedimented and diluted in fresh media to the OD=0.1. This suspension was serially diluted ×10 and 2.5 µl of each dilution was spotted on the agar plate using a multichannel pipette. The plates were incubated for 2 days at 24˚, 30˚, and 37˚C and imaged using smartphone camera.

Analysis of translation start site using ribosome profiling

To check for possible mis-annotation of the translation start sites of MRPs in the *Saccharomyces* genome database, we re-analyzed data from a ribosome profiling study that was specifically designed to detect translation initiation sites (ref. Knöringer, Groh, Krämer et al., in preparation; data available at GEO with the accession number GSE172017, access code available upon request). Briefly, ribosome profiling libraries of yeast cells (YPH499) were prepared as described (Stein et al., 2019) with the following modification: In one

replicate, 100 μg/ml cycloheximide was added to the yeast culture 2 min before harvesting and lysis, while in the other replicate, cells were not in contact with cycloheximide prior to cell lysis. Cycloheximide inhibits translation elongation, but not translation initiation, which results in an enrichment of ribosome footprints at translation initiation sites in CHX-treated samples compared to untreated samples. Ribosome footprints were sequenced and aligned to the yeast genome. Footprint densities along the annotated open reading frames of MRP genes were analyzed and the translation start site was re-annotated based on the following criteria: (1) the presence of contiguous ribosome footprints in all samples (previously annotated introns were taken into consideration); (2) the presence of an enrichment of footprints in the CHX-treated sample at the very 5' position of the suspected ORF; (3) the presence of an AUG codon at the suspected start site.

Fluorescence microscopy

The day before the experiment yeast were grown to saturation. On the day of experiment the saturated culture was diluted 1:50 in SD media without selection or with just auxotrophic selections. Cells were grown for 4 h and applied on glass-bottom plates coated with concanavalin A and left for 20 min to adhere. If mitochondria needed to be visualized, the solution was removed from the wells and 50 nM MitoTracker Orange CMTMRos (ThermoFisher #M7510) diluted in imaging media (SD media with complete set of amino acids but without riboflavin) was placed in the wells for 10 min. Imaging was performed in fresh imaging media. Cells were imaged using VisiScope Confocal Cell Explorer system consisting of Olympus IX83 microscope, Zeiss Yokogawa spinning disk scanning unit equipped with PCO-Edge sCMOS camera controlled by VisiView sofware. Images were recorded with 488 nm laser illumination for GFP channel and 561 nm laser illumination for MitoTracker Orange and 60× oil objective. Micrographs were cropped, and slightly adjusted for brightness and contrast using Fiji (Schindelin et al., 2012).

Data analysis

The data on the first amino acid in the mitoribosome structures were extracted directly from the mmCIF files of PDB entries 6WD0 (Loveland et al., 2020), 5MRC (Desai et al., 2017), 6YWS, 6YW5 (Itoh et al., 2020), 6NU2 (Koripella et al., 2019), 6GAW (Kummer et al., 2018), 6XYW (Waltz et al., 2020), 6HIV (Ramrath et al., 2018), 6ZP1 (Tobiasson and Amunts,

2020) using PDBeCIF (https://pypi.org/project/PDBeCif/). The number of atoms around each C-alpha atom were calculated from this data using Python (McKinney, 2010).

The cleavage site in MRP N-termini was annotated from the following sources with priority as listed: (1) original publication with N-terminal sequencing as cited by UniProt (Boguta et al., 1992; Dang and Ellis, 1990; Davis et al., 1992; Graack et al., 1988; Graack et al., 1991; Grohmann et al., 1989; Grohmann et al., 1991; Kitakawa et al., 1990; Kitakawa et al., 1997; Matsushita and Isono, 1993; Matsushita et al., 1989), (2) high-throughput N-terminal proteomic dataset (Vögtle et al., 2009), (3) cleavage site prediction by UniProt, (4) cleavage cite prediction using MitoFates (Fukasawa et al., 2015). If the cleavage site annotation didn't agree with the structural data (cleavage annotated after the amino acid actually present in the structure), we took the annotation from the source of next priority. If none of the annotations agreed with the structure, the N-terminus cleavage site was marked as "NA" (Table S1).

Mitochondrial targeting sequence prediction scores were calculated using MitoProt (Claros and Vincens, 1996), TargetP1 (Emanuelsson et al., 2000), TargetP2 (Armenteros et al., 2019), and MitoFates (Fukasawa et al., 2015) as described. The profiles and propensities of iMTSLs were calculated as described (Backes et al., 2018; Boos et al., 2018). Protein charge and hydrophobic moments were calculated using EMBOSS suite (Rice, 2000).

Structures were visualized using ChimeraX (Goddard et al., 2018), and electrostatic potentials using Chimera (Pettersen et al., 2004). Structural alignments were performed using FATCAT (Li et al., 2020), sequence alignment was performed and visualized using UniPro UGENE (Okonechnikov et al., 2012).

All other data analysis was performed using MATLAB (MathWorks). Plots were produced using MATLAB, Microsoft Excel, R (R Core Team, 2020), and GraphPad Prism.

Miscellaneous

The following procedures were carried out as published before: isolation of mitochondria and *in vitro* import experiments (Peleh et al., 2015) and plasmid shuffling (Backes et al., 2019).

List of Supplementary Materials

Figs. S1-S12

Table S1. Manual annotation of cleavable MTS's and other properties of yeast MRPs, Legend appears in Table.

Table S2. *In vivo* characterization of 15 MRP N-termini targeting capacities. Legend appears in table.

Table S3. Yeast strains used in this study.

Table S4. Plasmids used in this study.

Table S5. Primers for yeast transformation used in this study.

References

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

Armenteros, J. J. A., Salvatore, M., Emanuelsson, O., Winther, O., Heijne, G. von, Elofsson, A. and Nielsen, H. (2019). Detecting sequence signals in targeting peptides using deep learning. *Life Science Alliance* **2**,.

Backes, S., Hess, S., Boos, F., Woellhaf, M. W., Gödel, S., Jung, M., Mühlhaus, T. and Herrmann, J. M. (2018). Tom70 enhances mitochondrial preprotein import efficiency by binding to internal targeting sequences. *The Journal of Cell Biology* **217**, 1369–1382.

Backes, S., Garg, S. G., Becker, L., Peleh, V., Glockshuber, R., Gould, S. B. and Herrmann, J. M. (2019). Development of the Mitochondrial Intermembrane Space Disulfide Relay Represents a Critical Step in Eukaryotic Evolution. *Molecular Biology and Evolution* **36**, 742–756.

Ban, N., Nissen, P., Hansen, J., Moore, P. B. and Steitz, T. A. (2000). The Complete Atomic Structure of the Large Ribosomal Subunit at 2.4 Å Resolution. *Science* **289**, 905–920.

Ban, N., Beckmann, R., Cate, J. H., Dinman, J. D., Dragon, F., Ellis, S. R., Lafontaine, D. L., Lindahl, L., Liljas, A., Lipton, J. M., et al. (2014). A new system for naming ribosomal proteins. *Curr Opin Struct Biol* **24**, 165–169.

Bedwell, D. M., Strobel, S. A., Yun, K., Jongeward, G. D. and Emr, S. D. (1989). Sequence and structural requirements of a mitochondrial protein import signal defined by saturation cassette mutagenesis. *Molecular and Cellular Biology* **9**, 1014–1025.

Bertgen, L., Mühlhaus, T. and Herrmann, J. M. (2020). Clingy genes: Why were genes for ribosomal proteins retained in many mitochondrial genomes? *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1861**, 148275.

Boczonadi, V. and Horvath, R. (2014). Mitochondria: Impaired mitochondrial translation in human disease. *The International Journal of Biochemistry & Cell Biology* **48**, 77–84.

Boguta, M., Dmochowska, A., Borsuk, P., Wrobel, K., Gargouri, A., Lazowska, J., Slonimski, P. P., Szczesniak, B. and Kruszewska, A. (1992). NAM9 nuclear suppressor of mitochondrial ochre mutations in Saccharomyces cerevisiae codes for a protein homologous to S4 ribosomal proteins from chloroplasts, bacteria, and eucaryotes. *Molecular and Cellular Biology* **12**, 402–412.

591 **Bonn, F., Tatsuta, T., Petrungaro, C., Riemer, J. and Langer, T.** (2011). Presequence-
592 dependent folding ensures MrpL32 processing by the m-AAA protease in
593 mitochondria. *The EMBO journal* **30**, 2545–2556.

594 **Boos, F., Mühlhaus, T. and Herrmann, J. M.** (2018). Detection of Internal Matrix Targeting
595 Signal-like Sequences (iMTS-Ls) in Mitochondrial Precursor Proteins Using the
596 TargetP Prediction Tool. *Bio-protocol* **8**, e2474–e2474.

597 **Brachmann, C. B., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P. and Boeke, J. D.**
598 (1998). Designer deletion strains derived from Saccharomyces cerevisiae S288C: A
599 useful set of strains and plasmids for PCR-mediated gene disruption and other
600 applications. *Yeast* **14**, 115–132.

601 **Bykov, Y. S., Rapaport, D., Herrmann, J. M. and Schuldiner, M.** (2020). Cytosolic Events
602 in the Biogenesis of Mitochondrial Proteins. *Trends in Biochemical Sciences* **45**,
603 650–667.

604 **Carter, A. P., Clemons, W. M., Brodersen, D. E., Morgan-Warren, R. J., Wimberly, B. T.**
605 **and Ramakrishnan, V.** (2000). Functional insights from the structure of the 30S
606 ribosomal subunit and its interactions with antibiotics. *Nature* **407**, 340–348.

607 **Claros, M. G. and Vincens, P.** (1996). Computational Method to Predict Mitochondrially
608 Imported Proteins and their Targeting Sequences. *European Journal of Biochemistry*
609 **241**, 779–786.

610 **Dang, H. and Ellis, S. R.** (1990). Structural and functional analyses of a yeast mitochondrial
611 ribosomal protein homologous to ribosomal protein S15 of Escherichia coli. *Nucleic*
612 *Acids Res* **18**, 6895–6901.

613 **Davis, S. C., Tzagoloff, A. and Ellis, S. R.** (1992). Characterization of a yeast
614 mitochondrial ribosomal protein structurally related to the mammalian 68-kDa high
615 affinity laminin receptor. *J. Biol. Chem.* **267**, 5508–5514.

616 **Desai, N., Brown, A., Amunts, A. and Ramakrishnan, V.** (2017). The structure of the
617 yeast mitochondrial ribosome. *Science* **355**, 528–531.

618 **Desmond, E., Brochier-Armanet, C., Forterre, P. and Gribaldo, S.** (2011). On the last
619 common ancestor and early evolution of eukaryotes: reconstructing the history of
620 mitochondrial ribosomes. *Research in Microbiology* **162**, 53–70.

621 **Emanuelsson, O., Nielsen, H., Brunak, S. and von Heijne, G.** (2000). Predicting
622 Subcellular Localization of Proteins Based on their N-terminal Amino Acid Sequence.
623 *Journal of Molecular Biology* **300**, 1005–1016.

624 **Fox, G. E.** (2010). Origin and Evolution of the Ribosome. *Cold Spring Harb Perspect Biol* **2**,.

625 **Fukasawa, Y., Tsuji, J., Fu, S.-C., Tomii, K., Horton, P. and Imai, K.** (2015). MitoFates:
626 Improved Prediction of Mitochondrial Targeting Sequences and Their Cleavage
627 Sites. *Mol Cell Proteomics* **14**, 1113–1126.

628 **Gietz, R. D. and Woods, R. A.** (2006). Yeast Transformation by the LiAc/SS Carrier
629 DNA/PEG Method. In *Yeast Protocol* (ed. Xiao, W.), pp. 107–120. Totowa, NJ:
630 Humana Press.

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

634 **Gornicka, A., Bragoszewski, P., Chroscicki, P., Wenz, L.-S., Schulz, C., Rehling, P. and**
635 **Chacinska, A.** (2014). A discrete pathway for the transfer of intermembrane space
636 proteins across the outer membrane of mitochondria. *MBoC* **25**, 3999–4009.

637 **Graack, H.-R., Grohmann, L. and Choli, T.** (1988). Mitochondrial ribosomes of yeast:
638 Isolation of individual proteins and N-terminal sequencing. *FEBS Letters* **242**, 4–8.

639 **Graack, H. R., Grohmann, L. and Kitakawa, M.** (1991). The nuclear coded mitoribosomal
640 proteins YmL27 and YmL31 are both essential for mitochondrial function in yeast.
641 *Biochimie* **73**, 837–844.

642 **Greber, B. J. and Ban, N.** (2016). Structure and Function of the Mitochondrial Ribosome.
643 *Annual Review of Biochemistry* **85**, 103–132.

644 **Grohmann, L., Graack, H.-R. and Kitakawa, M.** (1989). Molecular cloning of the nuclear
645 gene for mitochondrial ribosomal protein YmL31 from Saccharomyces cerevisiae.
646 *European Journal of Biochemistry* **183**, 155–160.

647 **Grohmann, L., Graack, H.-R., Kruft, V., Choli, T., Goldschmidt-Reisin, S. and Kitakawa,**
648 **M.** (1991). Extended N-terminal sequencing of proteins of the large ribosomal subunit
649 from yeast mitochondria. *FEBS Letters* **284**, 51–56.

650 **Hamelryck, T.** (2005). An amino acid has two sides: A new 2D measure provides a different
651 view of solvent exposure. *Proteins: Structure, Function, and Bioinformatics* **59**, 38–
652 48.

653 **Hanscho, M., Ruckerbauer, D. E., Chauhan, N., Hofbauer, H. F., Krahulec, S., Nidetzky,**
654 **B., Kohlwein, S. D., Zanghellini, J. and Natter, K.** (2012). Nutritional requirements
655 of the BY series of Saccharomyces cerevisiae strains for optimum growth. *FEMS*
656 *Yeast Research* **12**, 796–808.

657 **Itoh, Y., Naschberger, A., Mortezaei, N., Herrmann, J. M. and Amunts, A.** (2020).
658 Analysis of translating mitoribosome reveals functional characteristics of translation
659 in mitochondria of fungi. *Nature Communications* **11**, 5187.

660 **Janke, C., Magiera, M. M., Rathfelder, N., Taxis, C., Reber, S., Maekawa, H., Moreno-**
661 **Borchart, A., Doenges, G., Schwob, E., Schiebel, E., et al.** (2004). A versatile
662 toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more
663 markers and promoter substitution cassettes. *Yeast* **21**, 947–962.

664 **Kitakawa, M., Grohmann, L., Graack, H.-R. and Isono, K.** (1990). Cloning and
665 characterization of nuclear genes for two mitochondrial ribosomal proteins in
666 Saccharomyces cerevisiae. *Nucleic Acids Res* **18**, 1521–1529.

667 **Kitakawa, M., Graack, H.-R., Grohmann, L., Goldschmidt-Reisin, S., Herfurth, E.,**
668 **Wittmann-Liebold, B., Nishimura, T. and Isono, K.** (1997). Identification and
669 Characterization of the Genes for Mitochondrial Ribosomal Proteins of
670 Saccharomyces Cerevisiae. *European Journal of Biochemistry* **245**, 449–456.

671 **Koripella, R. K., Sharma, M. R., Risteff, P., Keshavan, P. and Agrawal, R. K.** (2019).
672 Structural insights into unique features of the human mitochondrial ribosome
673 recycling. *PNAS* **116**, 8283–8288.

674 **Kummer, E. and Ban, N.** (2021). Mechanisms and regulation of protein synthesis in
675 mitochondria. *Nature Reviews Molecular Cell Biology* 1–19.

676 **Kummer, E., Leibundgut, M., Rackham, O., Lee, R. G., Boehringer, D., Filipovska, A.**
677 **and Ban, N.** (2018). Unique features of mammalian mitochondrial translation
678 initiation revealed by cryo-EM. *Nature* **560**, 263–267.

679 **Lee, C. M., Sedman, J., Neupert, W. and Stuart, R. A.** (1999). The DNA Helicase, Hmi1p,
680 Is Transported into Mitochondria by a C-terminal Cleavable Targeting Signal. *J. Biol.*
681 *Chem.* **274**, 20937–20942.

682 **Li, Z., Jaroszewski, L., Iyer, M., Sedova, M. and Godzik, A.** (2020). FATCAT 2.0: towards
683 a better understanding of the structural diversity of proteins. *Nucleic Acids Research*
684 **48**, W60–W64.

685 **Longen, S., Woellhaf, M. W., Petrungaro, C., Riemer, J. and Herrmann, J. M.** (2014).
686 The Disulfide Relay of the Intermembrane Space Oxidizes the Ribosomal Subunit
687 Mrp10 on Its Transit into the Mitochondrial Matrix. *Developmental Cell* **28**, 30–42.

688 **Longtine, M. S., Iii, A. M., Demarini, D. J., Shah, N. G., Wach, A., Brachat, A.,**
689 **Philippsen, P. and Pringle, J. R.** (1998). Additional modules for versatile and
690 economical PCR-based gene deletion and modification in Saccharomyces
691 cerevisiae. *Yeast* **14**, 953–961.

692 **Loveland, A. B., Demo, G. and Korostelev, A. A.** (2020). Cryo-EM of elongating ribosome
693 with EF-Tu•GTP elucidates tRNA proofreading. *Nature* 1–6.

694 **Lucattini, R., Likić, V. A. and Lithgow, T.** (2004). Bacterial Proteins Predisposed for
695 Targeting to Mitochondria. *Mol Biol Evol* **21**, 652–658.

696 **Lutz, T., Neupert, W. and Herrmann, J. M.** (2003). Import of small Tim proteins into the
697 mitochondrial intermembrane space. *The EMBO Journal* **22**, 4400–4408.

698 **Martijn, J., Vosseberg, J., Guy, L., Offre, P. and Ettema, T. J. G.** (2018). Deep
699 mitochondrial origin outside the sampled alphaproteobacteria. *Nature* **557**, 101–105.

700 **Matsushita, Y. and Isono, K.** (1993). Mitochondrial transport of mitoribosomal proteins,
701 YmL8 and YmL20, in Saccharomyces cerevisiae. *European Journal of Biochemistry*
702 **214**, 577–585.

703 **Matsushita, Y., Kitakawa, M. and Isono, K.** (1989). Cloning and analysis of the nuclear
704 genes for two mitochondrial ribosomal proteins in yeast. *Molec. Gen. Genet.* **219**,
705 119–124.

706 **McKinney, W.** (2010). Data Structures for Statistical Computing in Python.pp. 56–61.
707 Austin, Texas.

708 **Melnikov, S., Ben-Shem, A., Yusupova, G. and Yusupov, M.** (2015). Insights into the
709 origin of the nuclear localization signals in conserved ribosomal proteins. *Nature*
710 *Communications* **6**, 7382.

Melnikov, S., Manakongtreecheep, K. and Söll, D. (2018). Revising the Structural Diversity of Ribosomal Proteins Across the Three Domains of Life. *Mol Biol Evol* **35**, 1588–1598.

Morgenstern, M., Stiller, S. B., Lübbert, P., Peikert, C. D., Dannenmaier, S., Drepper, F., Weill, U., Höß, P., Feuerstein, R., Gebert, M., et al. (2017). Definition of a High-Confidence Mitochondrial Proteome at Quantitative Scale. *Cell Reports* **19**, 2836–2852.

Muto, T., Obita, T., Abe, Y., Shodai, T., Endo, T. and Kohda, D. (2001). NMR identification of the Tom20 binding segment in mitochondrial presequences11Edited by M. F. Sumers. *Journal of Molecular Biology* **306**, 137–143.

Nolden, M., Ehses, S., Koppen, M., Bernacchia, A., Rugarli, E. I. and Langer, T. (2005). The m-AAA Protease Defective in Hereditary Spastic Paraplegia Controls Ribosome Assembly in Mitochondria. *Cell* **123**, 277–289.

Obita, T., Muto, T., Endo, T. and Kohda, D. (2003). Peptide Library Approach with a Disulfide Tether to Refine the Tom20 Recognition Motif in Mitochondrial Presequences. *Journal of Molecular Biology* **328**, 495–504.

Ohba, M. and Schatz, G. (1987). Disruption of the outer membrane restores protein import to trypsin-treated yeast mitochondria. *The EMBO Journal* **6**, 2117–2122.

Okonechnikov, K., Golosova, O., Fursov, M., and the UGENE team (2012). Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* **28**, 1166–1167.

Ott, M., Amunts, A. and Brown, A. (2016). Organization and Regulation of Mitochondrial Protein Synthesis. *Annual Review of Biochemistry* **85**, 77–101.

Pagliarini, D. J., Calvo, S. E., Chang, B., Sheth, S. A., Vafai, S. B., Ong, S.-E., Walford, G. A., Sugiana, C., Boneh, A., Chen, W. K., et al. (2008). A Mitochondrial Protein Compendium Elucidates Complex I Disease Biology. *Cell* **134**, 112–123.

Peleh, V., Ramesh, A. and Herrmann, J. M. (2015). Import of Proteins into Isolated Yeast Mitochondria. In *Membrane Trafficking: Second Edition* (ed. Tang, B. L.), pp. 37–50. New York, NY: Springer.

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

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. and Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612.

Pilla, S. P. and Bahadur, R. P. (2019). Residue conservation elucidates the evolution of r-proteins in ribosomal assembly and function. *International Journal of Biological Macromolecules* **140**, 323–329.

Poveda-Huertes, D., Matic, S., Marada, A., Habernig, L., Licheva, M., Myketin, L., Gilsbach, R., Tosal-Castano, S., Papinski, D., Mulica, P., et al. (2020). An Early mtUPR: Redistribution of the Nuclear Transcription Factor Rox1 to Mitochondria Protects against Intramitochondrial Proteotoxic Aggregates. *Molecular Cell* **77**, 180-188.e9.

753 **R Core Team** (2020). R: A language and environment for statistical computing. R
754 Foundation for Statistical Computing, Vienna, Austria.

755 **Ramrath, D. J. F., Niemann, M., Leibundgut, M., Bieri, P., Prange, C., Horn, E. K.,**
756 **Leitner, A., Boehringer, D., Schneider, A. and Ban, N.** (2018). Evolutionary shift
757 toward protein-based architecture in trypanosomal mitochondrial ribosomes. *Science*
758 **362**,.

759 **Rice, P.** (2000). EMBOSS: The European Molecular Biology Open Software Suite. *TIG* **16**,
760 276–277.

761 **Roger, A. J., Muñoz-Gómez, S. A. and Kamikawa, R.** (2017). The Origin and
762 Diversification of Mitochondria. *Current Biology* **27**, R1177–R1192.

763 **Sagan, L.** (1967). On the origin of mitosing cells. *Journal of Theoretical Biology* **14**, 225-IN6.

764 **Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T.,**
765 **Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al.** (2012). Fiji: an open-
766 source platform for biological-image analysis. *Nature Methods* **9**, 676–682.

767 **Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan,**
768 **A., Bartels, H., Agmon, I., Franceschi, F., et al.** (2000). Structure of Functionally
769 Activated Small Ribosomal Subunit at 3.3 Å Resolution. *Cell* **102**, 615–623.

770 **Shakya, V. P., Barbeau, W. A., Xiao, T., Knutson, C. S., Schuler, M. H. and Hughes, A.**
771 **L.** (2021). A nuclear-based quality control pathway for non-imported mitochondrial
772 proteins. *eLife* **10**, e61230.

773 **Sluis, V. D., O, E., Bauerschmitt, H., Becker, T., Mielke, T., Frauenfeld, J.,**
774 **Berninghausen, O., Neupert, W., Herrmann, J. M. and Beckmann, R.** (2015).
775 Parallel Structural Evolution of Mitochondrial Ribosomes and OXPHOS Complexes.
776 *Genome Biol Evol* **7**, 1235–1251.

777 **Smits, P., Smeitink, J. A. M., van den Heuvel, L. P., Huynen, M. A. and Ettema, T. J. G.**
778 (2007). Reconstructing the evolution of the mitochondrial ribosomal proteome.
779 *Nucleic Acids Research* **35**, 4686–4703.

780 **Stein, K. C., Kriel, A. and Frydman, J.** (2019). Nascent Polypeptide Domain Topology and
781 Elongation Rate Direct the Cotranslational Hierarchy of Hsp70 and TRiC/CCT.
782 *Molecular Cell* **75**, 1117-1130.e5.

783 **Tobiasson, V. and Amunts, A.** (2020). Ciliate mitoribosome illuminates evolutionary steps
784 of mitochondrial translation. *eLife* **9**, e59264.

785 **Vestweber, D. and Schatz, G.** (1988). Point mutations destabilizing a precursor protein
786 enhance its post-translational import into mitochondria. *The EMBO Journal* **7**, 1147–
787 1151.

788 **Vishwanath, P., Favaretto, P., Hartman, H., Mohr, S. C. and Smith, T. F.** (2004).
789 Ribosomal protein-sequence block structure suggests complex prokaryotic evolution
790 with implications for the origin of eukaryotes. *Molecular Phylogenetics and Evolution*
791 **33**, 615–625.

792 **Vögtle, F.-N., Wortelkamp, S., Zahedi, R. P., Becker, D., Leidhold, C., Gevaert, K.,**
793 **Kellermann, J., Voos, W., Sickmann, A., Pfanner, N., et al.** (2009). Global

794 Analysis of the Mitochondrial N-Proteome Identifies a Processing Peptidase Critical
795 for Protein Stability. *Cell* **139**, 428–439.

796 **Vögtle, F.-N., Burkhart, J. M., Gonczarowska-Jorge, H., Kücükköse, C., Taskin, A. A.,**
797 **Kopczynski, D., Ahrends, R., Mossmann, D., Sickmann, A., Zahedi, R. P., et al.**
798 (2017). Landscape of submitochondrial protein distribution. *Nature Communications*
799 **8**, 290.

800 **von Heijne, G.** (1986). Mitochondrial targeting sequences may form amphiphilic helices.
801 *EMBO J* **5**, 1335–1342.

802 **Waltz, F. and Giegé, P.** (2019). Striking Diversity of Mitochondria-Specific Translation
803 Processes across Eukaryotes. *Trends in Biochemical Sciences* **0**,.

804 **Waltz, F., Soufari, H., Bochler, A., Giegé, P. and Hashem, Y.** (2020). Cryo-EM structure of
805 the RNA-rich plant mitochondrial ribosome. *Nature Plants* **6**, 377–383.

806 **Woellhaf, M. W., Hansen, K. G., Garth, C. and Herrmann, J. M.** (2014). Import of
807 ribosomal proteins into yeast mitochondria. *Biochemistry and Cell Biology* **92**, 489–
808 498.

809 **Yamano, K., Yatsukawa, Y., Esaki, M., Hobbs, A. E. A., Jensen, R. E. and Endo, T.**
810 (2008). Tom20 and Tom22 Share the Common Signal Recognition Pathway in
811 Mitochondrial Protein Import. *J. Biol. Chem.* **283**, 3799–3807.

812 **Yofe, I. and Schuldiner, M.** (2014). Primers-4-Yeast: a comprehensive web tool for
813 planning primers for Saccharomyces cerevisiae. *Yeast* **31**, 77–80.

814 **Zaremba-Niedzwiedzka, K., Caceres, E. F., Saw, J. H., Bäckström, D., Juzokaite, L.,**
815 **Vancaester, E., Seitz, K. W., Anantharaman, K., Starnawski, P., Kjeldsen, K. U.,**
816 **et al.** (2017). Asgard archaea illuminate the origin of eukaryotic cellular complexity.
817 *Nature* **541**, 353–358.

818

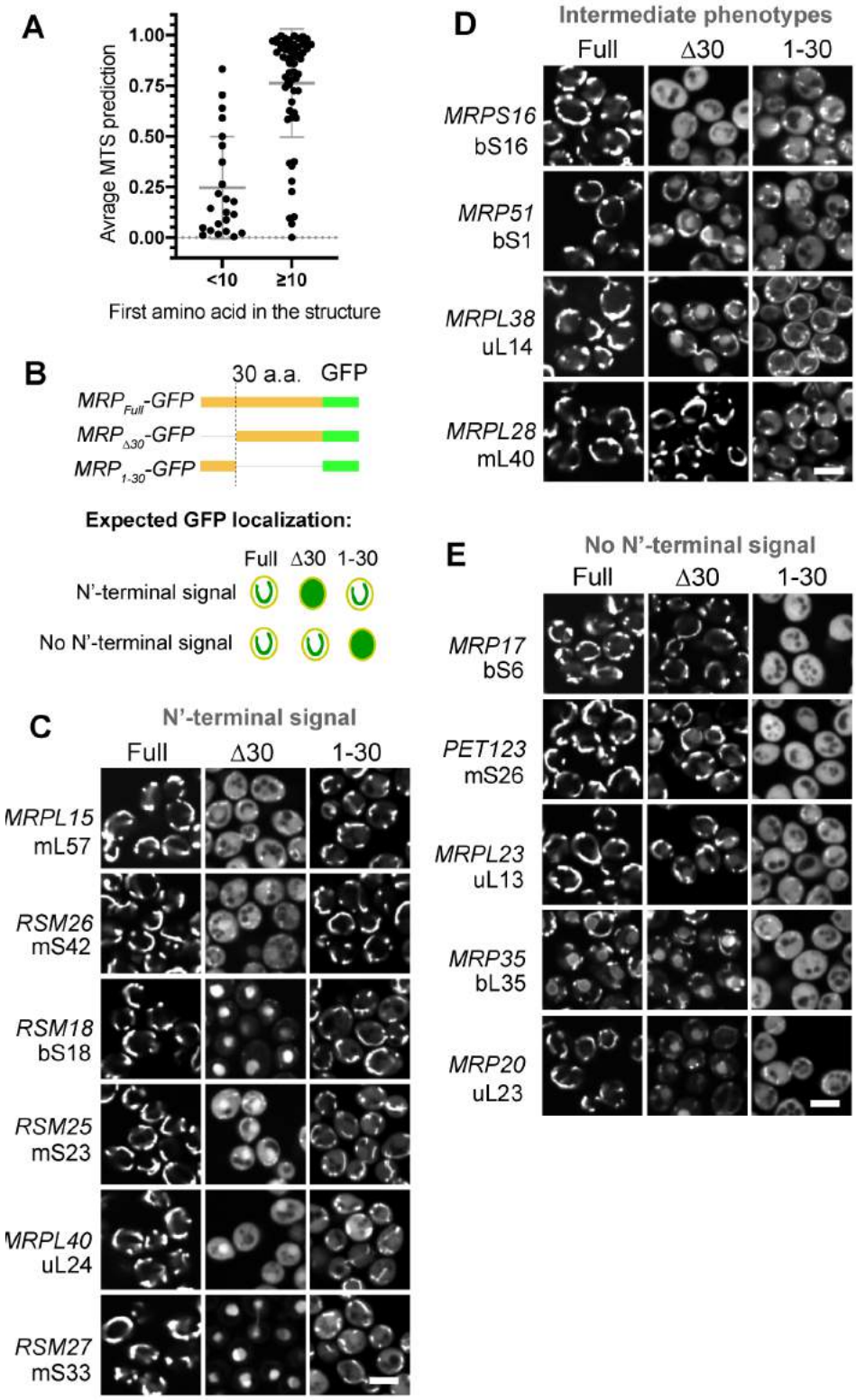


Fig.1. Mitochondrial Ribosomal Proteins (MRPs) have various types of targeting signals. (A) Yeast MRPs having uncleaved N-termini that can be tracked in the mitoribosome structure (PDB:5MRC) score much lower with MTS prediction algorithms (average of TargetP2 and MitoFates) compared to other MRPs that have their N-termini cleaved off or are not present in the structure (so might be flexible and outside the mitoribosome body). (B) Schematic of MRP truncations used to characterize targeting properties of MRP N-termini: MRPFull as control, MRPΔ30 to check if the N-terminus is necessary, MRP1-30 to check if it is sufficient (top) and the schematics of expected GFP localization in case the N-terminus is MTS-like (necessary and sufficient) or not (bottom); (C)-(E) Micrographs collected in the GFP channel for each truncation (columns) of each studied MRP (rows) grouped by the N-terminus targeting properties based on theoretical expectation summarized in (B) with the MRPs possessing MTS-like N-termini in panel (C), MRPs with intermediate phenotype in panel (D) and MRPs without N-terminal signal in panel (D), for each MRP a yeast gene name and new nomenclature protein name is shown on the left. Scale bar for all micrographs is 5 μm.

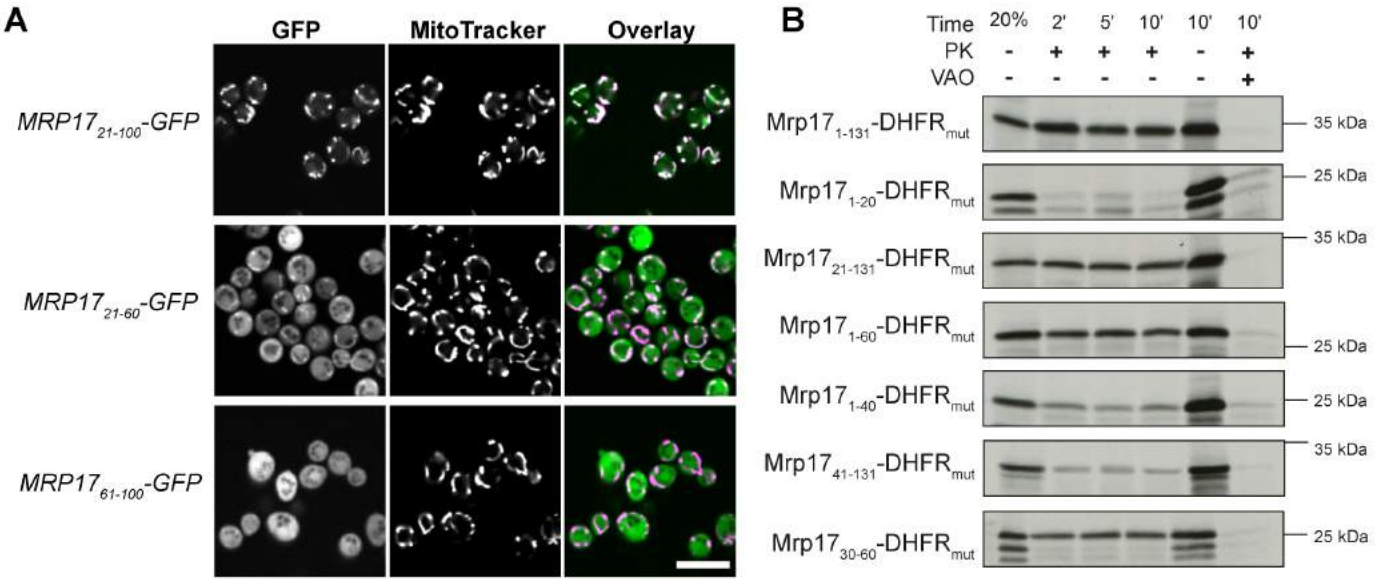


Fig. 2. The non-canonical targeting and translocation signal of Mrp17 is located between amino acids 30 and 60. (A) In vivo characterization of mitochondrial targeting capacity of different Mrp17 truncations fused to GFP visualized by fluorescent microscopy with MitoTracker Orange straining. Scale bar for all micrographs is 10 µm. (B) Characterization of Mrp17 translocation signal using an in vitro import assay: shown are autoradiographs of full-length Mrp17 or its truncations fused to DHFRmut, translated in vitro with radiolabeled amino-acids, incubated with isolated yeast mitochondria for 2, 5, or 10 min, treated with proteinase K (PK) to remove nonimported proteins and visualized by SDS-PAGE/autoradiography. As a negative control, mitochondria were treated with valinomycin, antimycin and oligomycin (VAO) that eliminate membrane potential. For comparison, 20% of the protein used per import reaction was loaded on the first lane.

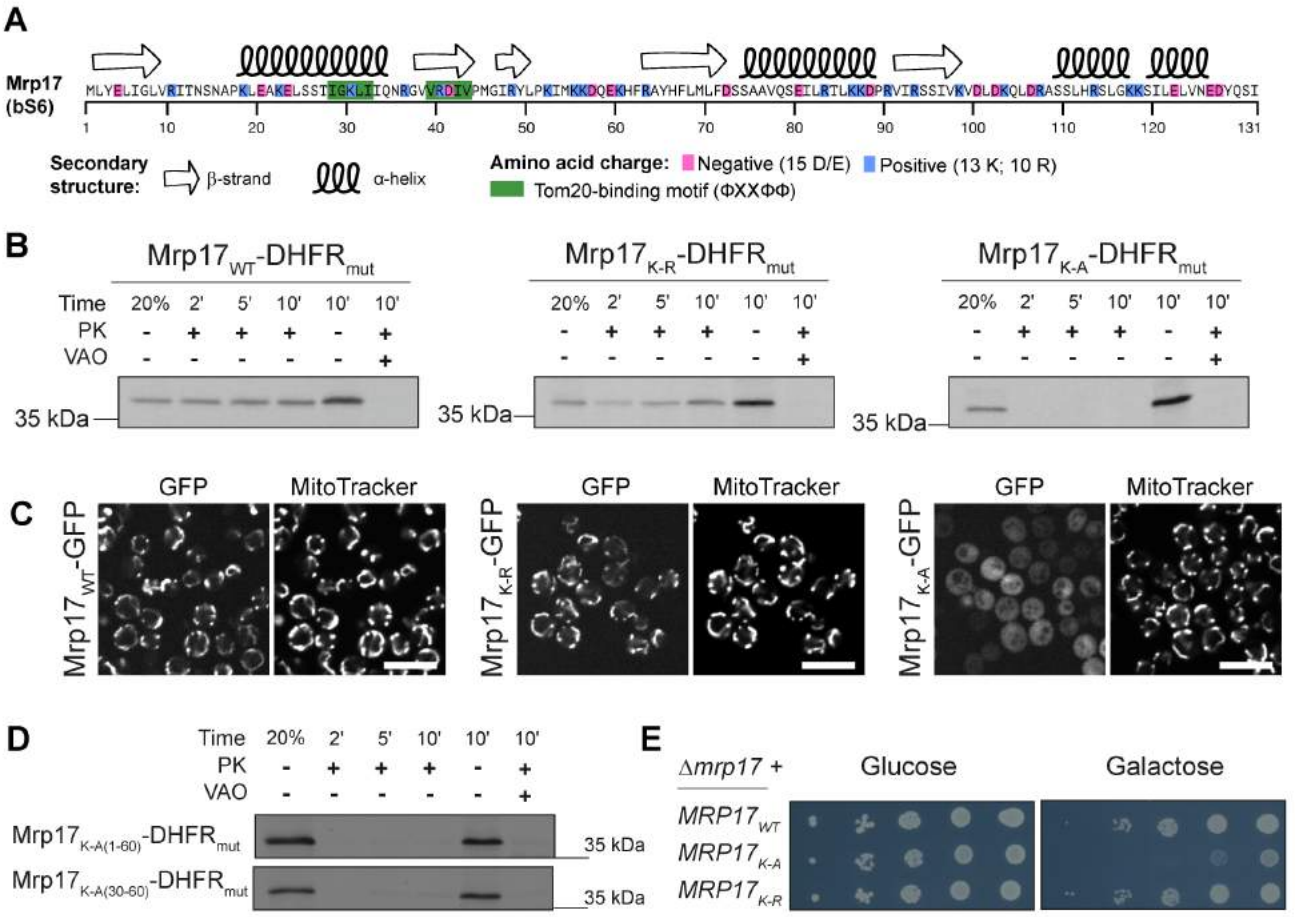


Fig. 3. Positive charge and Tom20-binding motifs are important features of the Mrp17 targeting signal. (A) Primary sequence of Mrp17 highlighting charged amino-acids, Tom20-binding motifs and secondary structure (from PDB:5MRC); (B) In vitro mitochondria translocation capacity of WT Mrp17, Mrp17K-R with all lysines (K) substituted with arginines (R), and Mrp17K-A with all lysines (K) substituted with alanines (A) fused to DHFRmut. Import was performed as described in Fig. 2 legend. (C) In vivo mitochondrial targeting capability of WT Mrp17, Mrp17K-R, and Mrp17K-A fused to GFP. (D) Substitution of lysines with alanines in the regions 1-60 or 30-60 of Mrp17 abolishes mitochondrial import capacity. (E) Plasmids carrying the indicated Mrp17 variants were transformed into a Δmrp17 strain that harbored an MRP17 gene on an URA3-containing plasmid. The latter plasmid was removed by plasmid shuffling and the growth of the strains was assessed on glucose and galactose media. Scale bar for all micrographs is 10 µm.

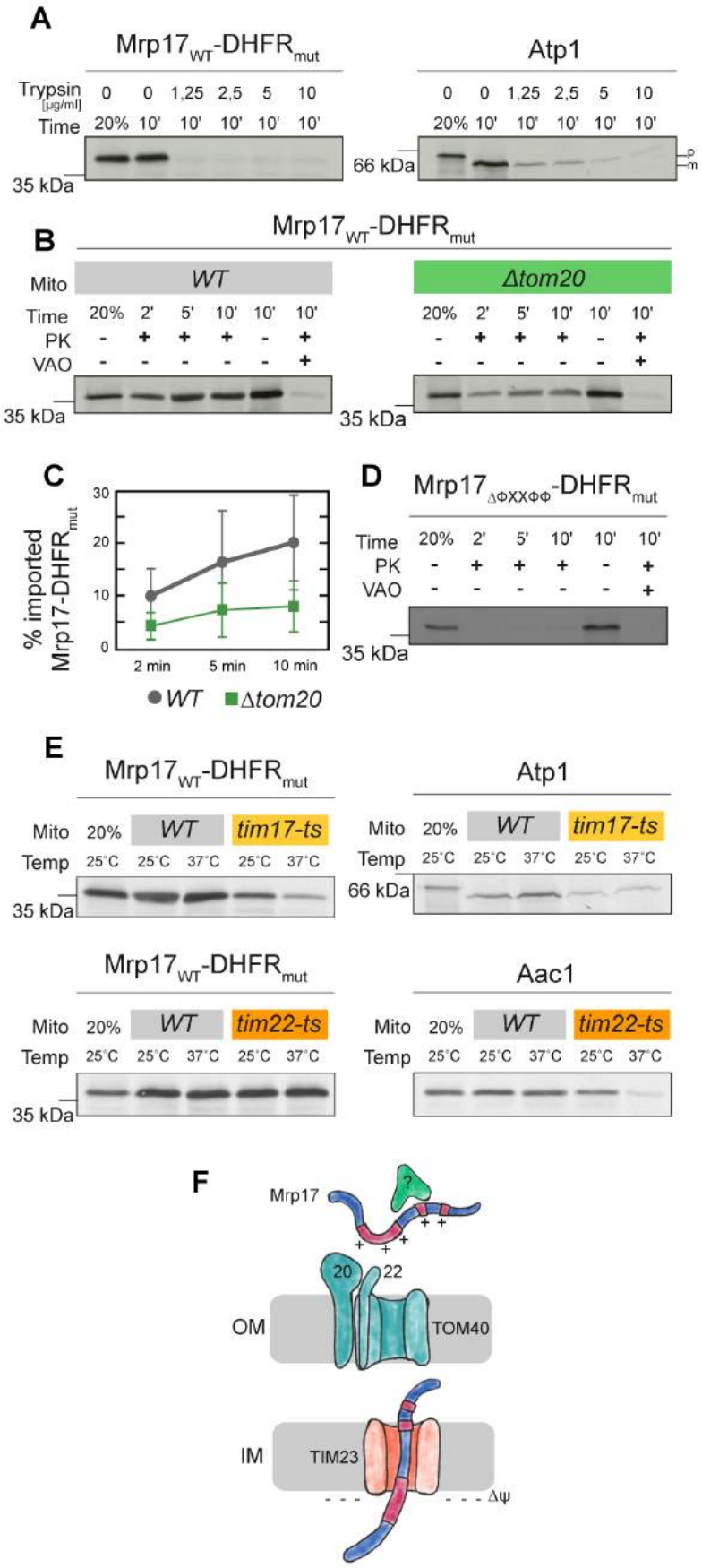


Fig. 4. Mrp17 requires Tom20 for translocation into isolated mitochondria. (A) The in vitro import of Mrp17-DHFRmut is sensitive to the elimination of outer mitochondrial membrane proteins by trypsinization (left), more so than the model import substrate Atp1 (right). Mitochondria were incubated with the indicated concentrations of trypsin and the import assay was performed as described in the legend for Fig. 2. Proteinase K (PK) was not added to the sample without mitochondria. (B) In vitro translocation of Mrp17-DHFRmut into mitochondria isolated from WT and Δtom20 yeast showing reduced translocation in Δ tom20 background. (C) Quantification of the experiment in panel (B) repeated 3 times, whiskers correspond to minimum and maximum values. (D) Substitution of both Tom20-binding motifs (ΦXXΦΦ) with alanines abolishes Mrp17 mitochondrial import capacity. (E) Import of Mrp17-DHFRmut and control proteins Atp1 and Aac2 into WT, tim17-ts and tim22-ts mitochondria, import was performed at indicated temperatures, proteinase K was added to all samples except the loading control in the first lane of each autoradiograph (20% of protein amount used for each import reaction). (F) Mrp17 (blue) targeting and translocation model: positively charged internal signals (magenta) might bind yet unidentified targeting factors (green) and then are recognized at the outer membrane of mitochondria by the same receptors Tom20 and Tom22 as a regular MTS, then imported through TOM40 (teal) and TIM23 (orange) translocons.

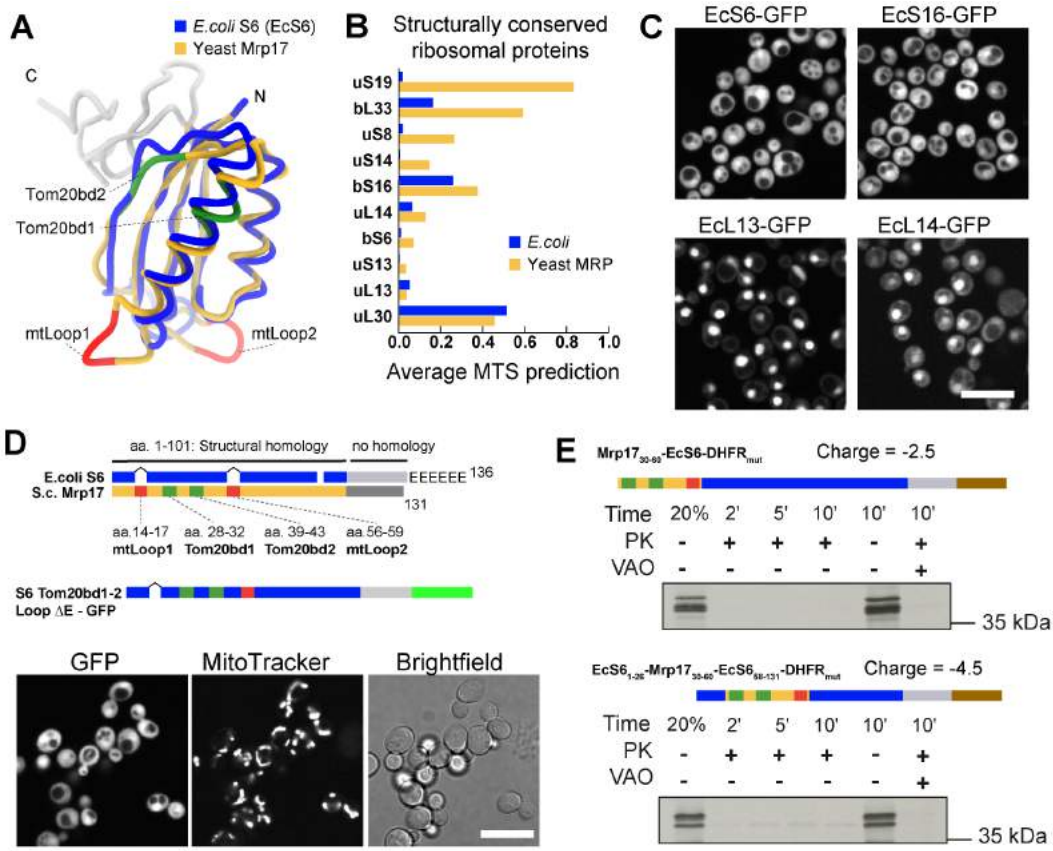


Fig. 5. Bacterial homologs of MRPs have no mitochondrial import predisposition. (A) Structural alignment of the yeast Mrp17 (from PDB:5MRC) and its E. coli homolog ribosomal protein S6 (from PDB: 6WD0) viewed from its cytosol-facing side, highlighting regions of structural homology (blue and yellow) and non-homologous C-terminal regions (grey), Tom20-binding domains in Mrp17 are highlighted in green and expanded loops in red, schematic summaries of such structural alignments for all MRPs and their bacterial homologs are shown in Fig. S10. (B) Average MTS prediction scores (MitoFates and TargetP2) for structurally conserved yeast MRPs and their bacterial homologs sorted by the difference between bacteria and mitochondria. (C) Expression of structurally conserved bacterial RPs fused to GFP in yeast cells. (D) Schematic summary of structural alignment of bacterial S6 and yeast Mrp17 shown in panel (A) highlighting additional features of Mrp17 structure (two loops and two Tom20-binding sites) and six glutamate residues (E's) at the C-terminus of bacterial S6 (top), schematic of chimeric construct of EcS6 that lacks C-terminal glutamines, but incorporates structural features of Mrp17 and is fused to GFP shown bright green (middle), and the result of expression of this chimeric construct in yeast (bottom). (E) Chimeric variants of EcS6-DHFRmut with Mrp17 amino acids 30-60 fused to it at the N-terminus or incorporated instead of the homologous region of S6 itself (top), and in vitro import of these constructs into isolated yeast mitochondria. Lanes are labeled as described in Fig. 2 legend (bottom). Scale bar in all micrographs is 10 μm.

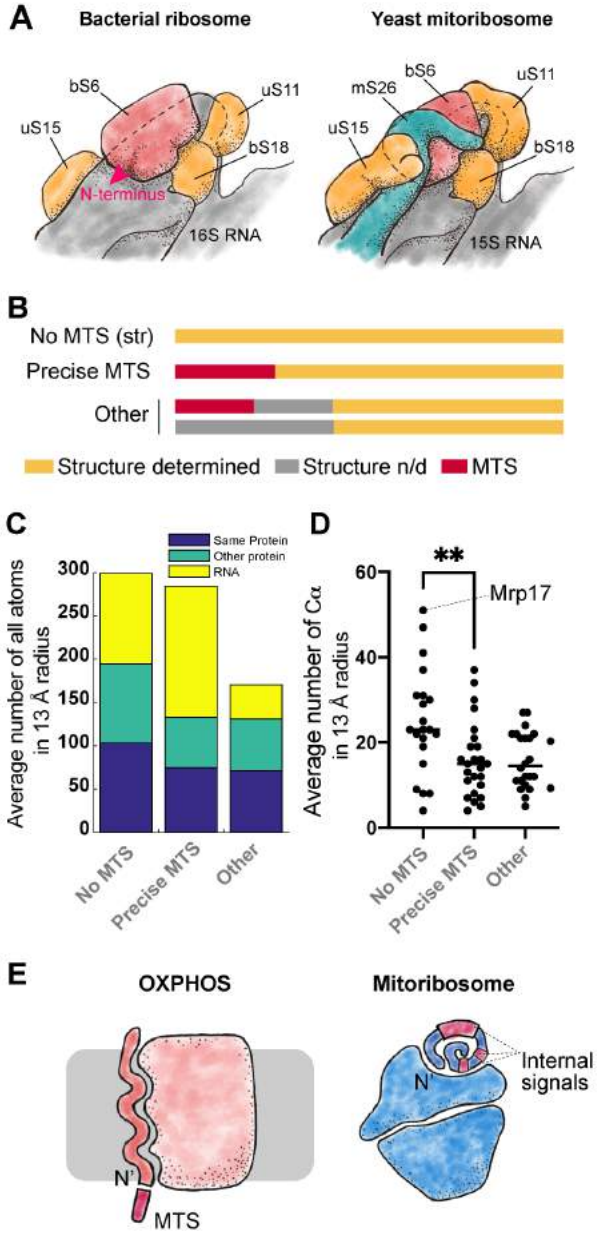


Fig. 6. Structural constraints for the evolution of N-terminal mitochondrial targeting signals. (A) Schematic depiction of the S6 protein (red) in the context of bacterial ribosome (after PDB:6WD0, left) and yeast mitochondrial ribosome (after PDB:5MRC, right) highlighting the position of the N-terminus of S6 (red arrow, only shown on the left image). Proteins homologous between bacterial and yeast structures are shown in orange, mitochondria-specific protein mS26 in teal and rRNA in grey. (B) MRPs divided in three groups depending on relative position of the MTS and structured residues: No MTS (str) – structure starts at the very N-terminus without cleavable MTS, Precise MTS – MTS is cleaved off 1-2 amino acids before the structure starts, Other – unstructured uncleaved amino acids present before the structure start. (C) MRP N-termini solvent exposure: average number of atoms belonging to the same protein, other protein chains or RNA in the 13Å radius from the first structured amino acid Cα atom, plotted separately for proteins with their most N-terminus appearing in the structure vs proteins with precisely cleaved MTS (first structured amino acid is within 3 or less amino acids from the end of the MTS) and other proteins which have a longer unstructured extension calculated based on PDB:5MRC. (D) Same as (E) but only coordination number (number of Cα atoms in the 13Å radius (Hamelryck, 2005)) is calculated. (F) Compared to OXPHOS complex components that usually have their N-termini free to develop an MTS (left), mitoribosomes often employ the MRP N-termini in binding interfaces which puts structural constraints on the development of N-terminal MTS and promotes the evolution of internal targeting sequences that can also take part in ribosome assembly (right).

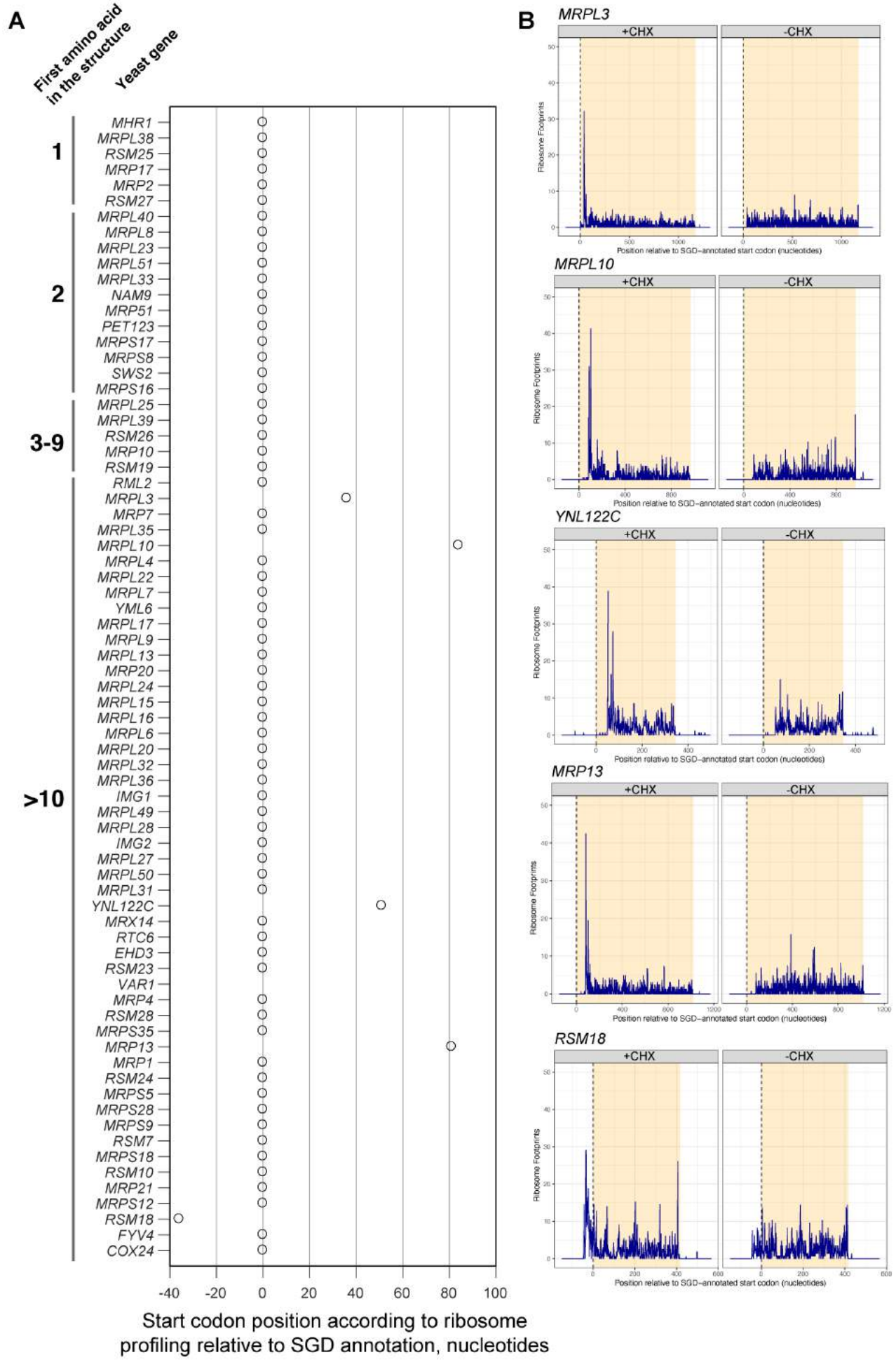


Fig. S1 - legend next page

Fig. S1. RiboSeq reveals no mis-annotated upstream translation starts that can account for the missing N-terminal signals. (A) For each MRP the translation start measured by RiboSeq is plotted relative to SGD annotation, revealing that only one protein has an N-terminal extension while 4 others have shorter N-termini, all of the misannotated proteins are reported to have an MTS (Table S1); no MRPs with their N-termini in the structure (first structured amino acid <10) had misannotated translation starts. (B) Ribosome footprints for each ORF with misannotated start codons in the presence of cycloheximide (+CHX, left) and without cycloheximide (-CHX, right), plotted relative to the ORF annotation in SGD (shaded in yellow).

		Bacteria	Fungi		Mammals		Plant	Kinetoplastida	Alveolata
Protein	Group	*E.coli*	*S.cerevisiea*	*N.crassa*	*H.sapiens*	*S.scrofa*	*B.oleracea*	*T.brucei*	*T.termophyla*
uL1	u/b	3	NA	59	NA	NA	NA	NA	NA
uL2	u/b	1	61	54	61	61	56	NA	1
uL3	u/b	1	21	63	45	42	64	43	32
uL4	u/b	1	29	62	45	45	89	18	22
uL5	u/b	1	19	44	NA	NA	3	NA	1
uL6	u/b	1	17	52	NA	NA	2	NA	1
bL9	u/b	1	10	52	53	53	44	9	12
uL10	u/b	1	NA	54	30	29	3	4	1
uL11	u/b	1	NA	12	18	17	15	14	1
bL12	u/b	NA	NA	NA	NA	10	NA	NA	12
uL13	u/b	1	2	1	2	2	15	11	28
uL14	u/b	1	1	1	31	31	47	NA	1
uL15	u/b	2	58	47	10	9	65	10	33
uL16	u/b	1	38	61	47	30	37	25	1
bL17	u/b	1	2	2	9	9	4	40	2
uL18	u/b	2	NA	NA	42	38	8	NA	NA
bL19	u/b	1	16	41	74	53	101	8	1
bL20	u/b	1	NA	NA	10	10	1	10	2
bL21	u/b	1	44	85	49	55	110	56	30
uL22	u/b	1	85	43	47	NA	103	2	23
uL23	u/b	1	55	13	2	2	4	61	6
uL24	u/b	1	2	1	15	11	10	1	45
bL25	u/b	1	NA	NA	NA	NA	34	NA	26
bL27	u/b	7	35	67	34	39	55	27	38
bL28	u/b	1	22	42	2	2	41	10	23
uL29	u/b	1	15	44	63	65	31	9	21
uL30	u/b	1	2	2	35	35	17	51	NA
bL31	u/b	NA	20	48	34	35	32	2	NA
bL32	u/b	1	72	64	79	79	87	26	NA
bL33	u/b	3	6	7	14	13	8	36	6
bL34	u/b	1	60	95	47	50	NA	NA	NA
bL35	u/b	1	54	NA	94	94	110	40	50
bL36	u/b	1	56	79	67	63	2	59	1
mL37	e	NA	NA	NA	31	30	NA	NA	NA
mL38	e	NA	20	83	27	27	NA	41	NA
mL39	e	NA	NA	NA	36	30	NA	NA	NA
mL40	e	NA	34	248	83	83	145	33	23
mL41	e	NA	17	9	15	14	22	47	11
mL42	e	NA	NA	NA	35	35	NA	51	NA
mL43	e	NA	2	2	2	2	4	2	2
mL44	e	NA	67	55	31	31	NA	NA	NA
mL45	e	NA	NA	NA	117	38	NA	NA	NA
mL46	e	NA	23	82	43	43	26	18	66
mL48	e	NA	NA	NA	48	48	NA	NA	NA
mL49	e	NA	41	82	38	34	NA	9	1
mL50	e	NA	38	112	56	51	NA	NA	NA
mL51	e	NA	NA	NA	32	32	NA	NA	NA
mL52	e	NA	NA	NA	24	23	NA	19	NA
mL53	e	NA	NA	1	13	2	5	37	59
mL54	e	NA	NA	77	114	71	NA	NA	40
mL57	e	NA	29	61	NA	NA	NA	NA	NA
mL58	e	NA	19	41	NA	NA	NA	NA	NA
mL59	e	NA	3	2	25	25	40	27	NA
mL60	e	NA	13	13	NA	NA	23	NA	NA
mL61	e	NA	NA	NA	NA	NA	NA	NA	NA
MHR1	e	NA	1	33	NA	NA	NA	NA	NA
mL62	e	NA	NA	NA	38	38	NA	NA	NA
mL63	e	NA	NA	NA	9	9	NA	17	NA
mL65	e	NA	NA	NA	41	40	NA	NA	NA
mL66	e	NA	NA	NA	35	35	NA	NA	NA

		Bacteria	Fungi		Mammals		Plant	Kinetoplastida	Alveolata
Protein	Group	*E.coli*	*S.cerevisiea*	*N.crassa*	*H.sapiens*	*S.scrofa*	*B.oleracea*	*T.brucei*	*T.termophyla*
bS1	u/b	NA	2	6	77	NA	NA	NA	NA
uS2	u/b	1	128	107	58	55	7	NA	1
uS3	u/b	1	30	17	36	36	3	4	1
uS4	u/b	1	2	1	NA	NA	18	NA	1
uS5	u/b	9	14	56	88	88	297	10	16
bS6	u/b	1	1	1	2	2	2	2	2
uS7	u/b	1	87	87	35	35	8	NA	1
uS8	u/b	1	2	2	NA	NA	2	10	NA
uS9	u/b	3	35	69	71	70	306	1	23
uS10	u/b	5	16	81	61	50	124	10	1
uS11	u/b	12	70	253	59	60	191	10	13
uS12	u/b	1	29	45	31	31	2	54	5
uS13	u/b	1	2	2	NA	NA	33	NA	3
uS14	u/b	1	1	2	28	28	69	10	1
uS15	u/b	1	34	38	66	65	316	69	18
bS16	u/b	1	2	2	10	NA	2	9	1
uS17	u/b	3	2	6	5	2	4	10	2
bS18	u/b	9	40	108	47	47	164	7	27
uS19	u/b	2	9	9	NA	NA	128	31	1
bS21	u/b	3	86	149	2	2	NA	10	1
mS22	e	NA	NA	NA	67	88	NA	10	NA
mS23	e	NA	1	2	2	2	6	9	1
mS25	e	NA	NA	NA	2	2	NA	NA	NA
mS26	e	NA	2	47	27	27	43	10	25
mS27	e	NA	42	40	36	68	NA	NA	NA
mS29	e	NA	50	57	51	45	128	9	22
mS31	e	NA	NA	NA	276	NA	NA	NA	182
mS33	e	NA	1	2	5	4	12	11	2
mS34	e	NA	NA	NA	6	1	50	10	10
mS35	e	NA	47	42	52	51	289	30	12
mS37	e	NA	5	2	2	2	30	20	3
mS38	e	NA	78	312	128	128	NA	141	NA
mS39	e	NA	NA	NA	56	55	NA	NA	NA
mS40	e	NA	NA	NA	52	49	NA	NA	NA
mS41	e	NA	30	43	NA	NA	41	10	24
mS42	e	NA	8	43	NA	NA	NA	9	NA
mS43	e	NA	18	NA	NA	NA	NA	9	NA
mS45	e	NA	27	66	NA	NA	143	NA	18
mS46	e	NA	197	223	NA	NA	NA	NA	NA
mS47	e	NA	36	42	NA	NA	42	26	24

First amino acid in the structure

- 1 to 9
- 10 to 14
- 15 and more
- Protein absent from the species or the structure

Protein origin

- u/b — Universal or specific for bacterial and and mitochondrial ribosomes
- e — Specific for mitochondria (any eukaryotic lineage)

Fig. S2. The first amino acids of MRPs found in structures of different mitoribosomes are moderately conserved. The first amino acid numbers were extracted directly from the PDB entries 6WDO (Escherichia coli), 5MRC (Saccharomyces cerevisiae), 6YWS (large subunit Neurospora crassa), 6YW5 (small subunit N. crassa), 6NU2 (Homo sapiens), 6GAW (Sus scrofa), 6XYW (Brassica oleracea), 6HIV (Trypanosoma brucei), 6ZP1 (Tetrahymena thermophila). Only MRPs occurring in two or more structures are shown.

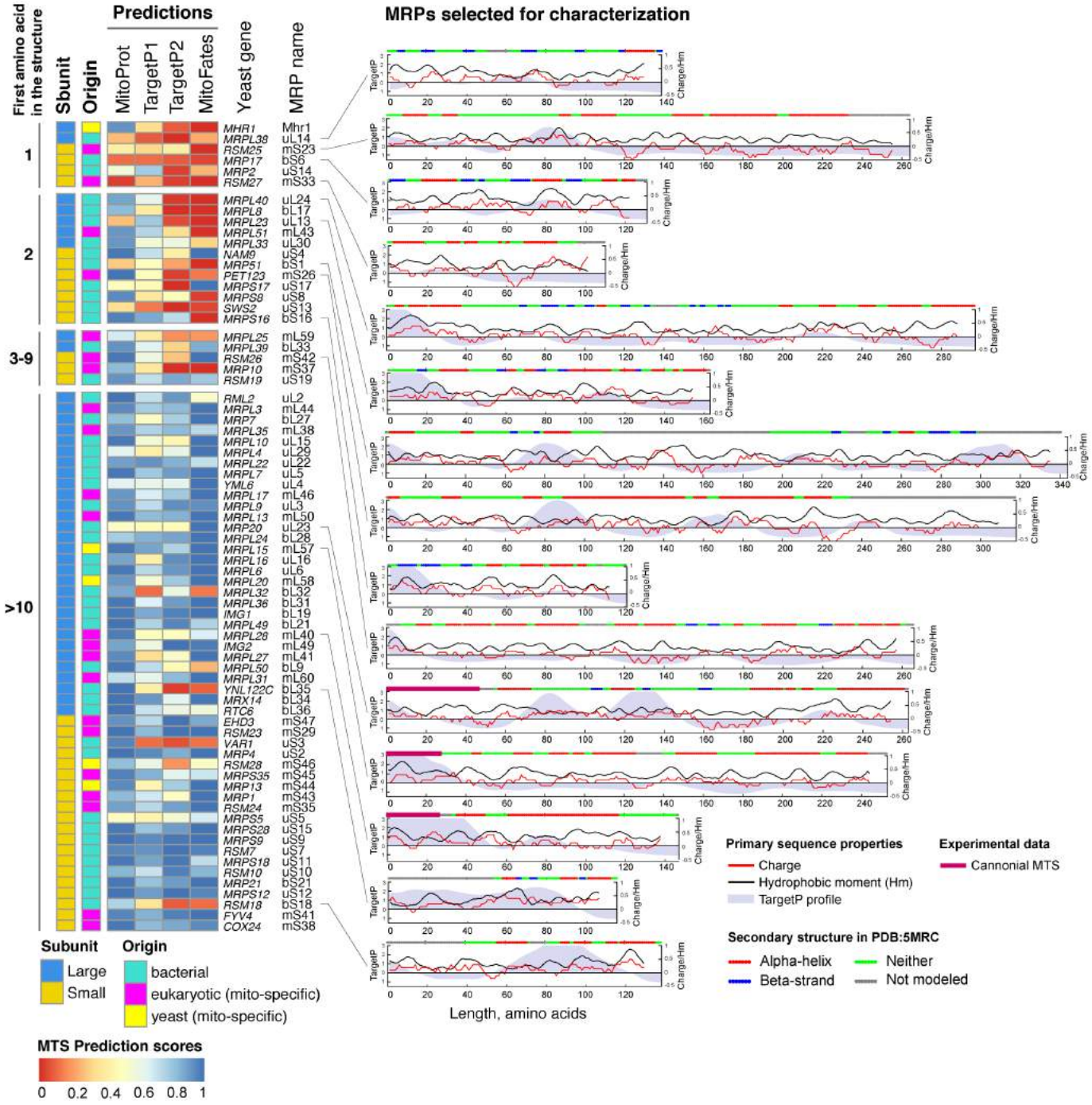


Fig. S3. MRPs can be classified according to the presence of their most N-terminus inside the mitoribosome structure. Left – MTS prediction scores for yeast MRPs first sorted in groups by the first amino acid with reported atomic coordinates in the structure PDB:5MRC, then by subunit and then by length with protein origin and subunit noted for each MRP. Right – primary and secondary sequence properties for 15 MRPs selected for further characterization showing a variety of N-terminal and internal targeting signal predictions, overall positive charge, presence of documented cleavable MTS, and a variety of N-terminal secondary structures. Universal ribosomal protein nomenclature is used (Ban et al. 2014), except for Mhr1 which is a yeast-specific MRP.

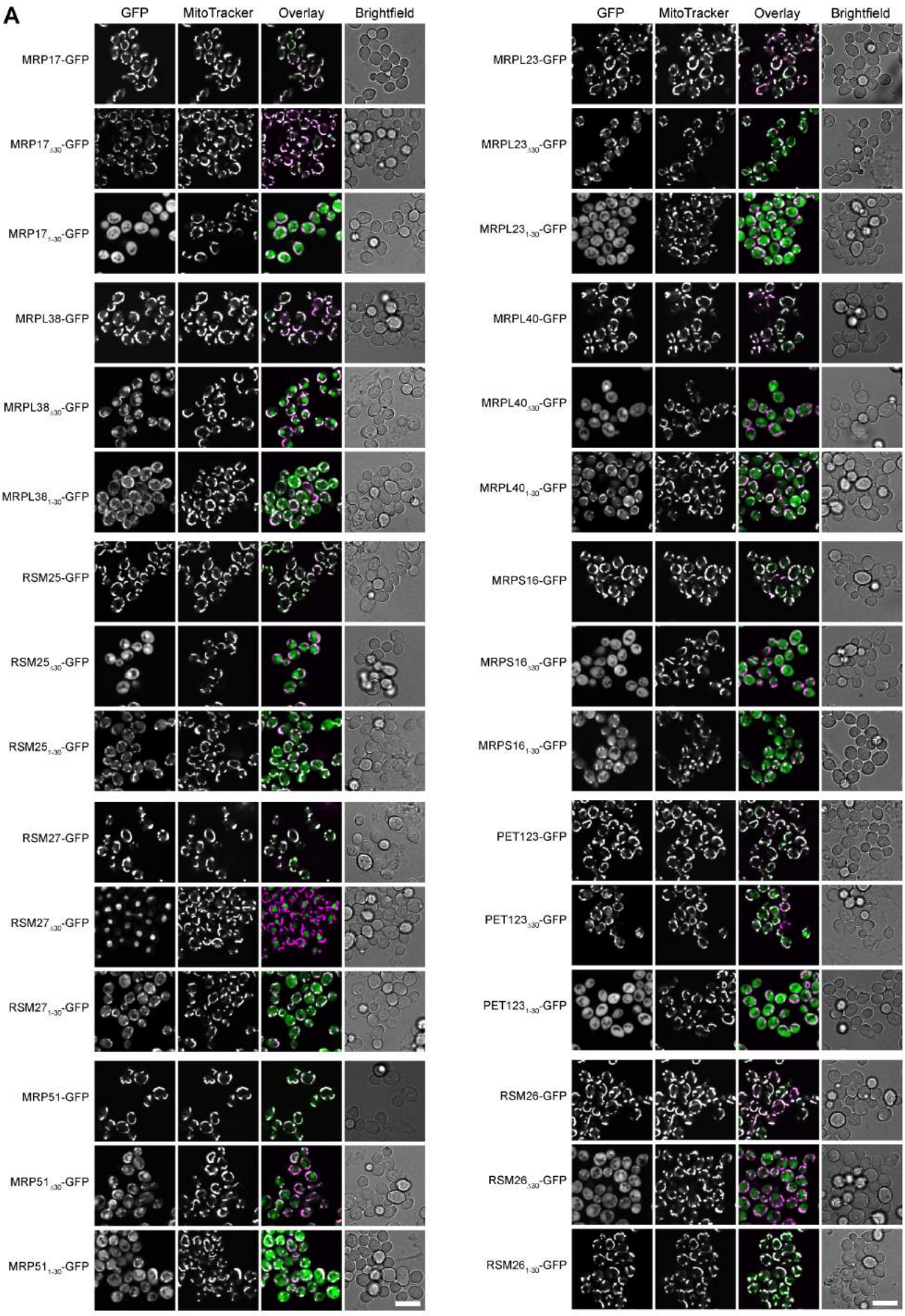


Fig. S4 - continues on the next page

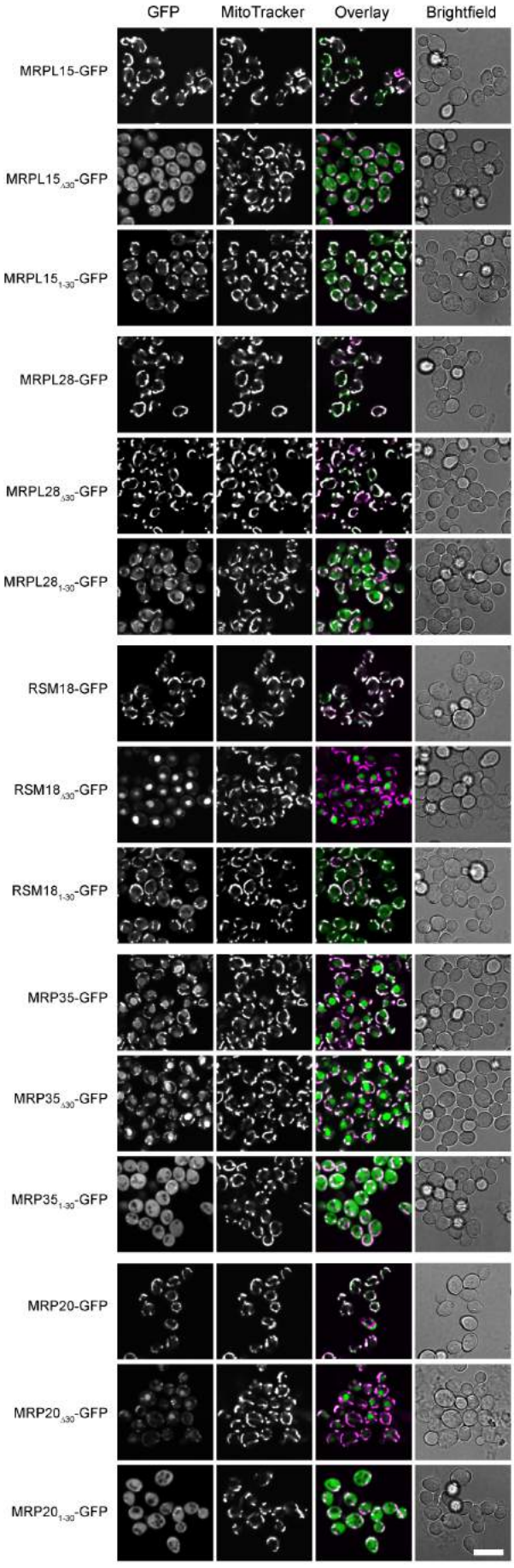
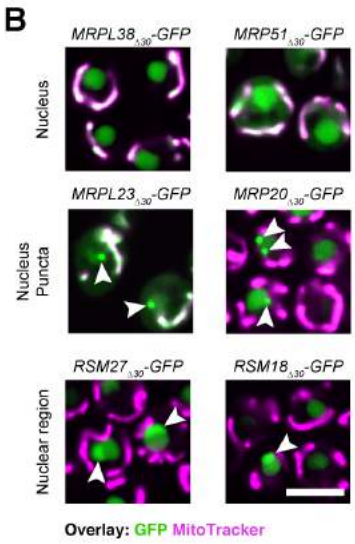


Fig. S4 - end.



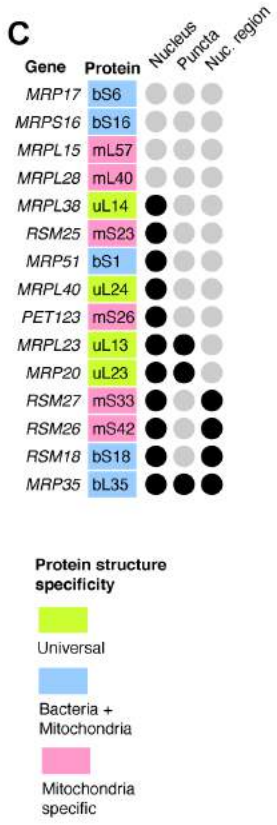


Fig. S4. MRPs have N-termini with various targeting properties. (A) Same micrographs as in Fig. 1C shown in different fluorescent channels for all the MRP. (B) Examples of mistargeting destinations of some MRPs with deleted 1-30 amino acids from their N-termini: top row – uniform nuclear signal, middle row – nuclear signal with bright perinuclear puncta, bottom row – nuclear signal with GFP enriched in non-punctate nuclear region, shown as enlarged overlays of GFP (green) and MitoTracker (magenta) channels cropped from micrographs in panel (A). (C) Summary of mistargeting locations for one or more truncations of each MRP, if the location is observed for any of the MRP truncation, it is marked with black circle. Scale bars are 10 μm in (A) and 5 μm in (B).

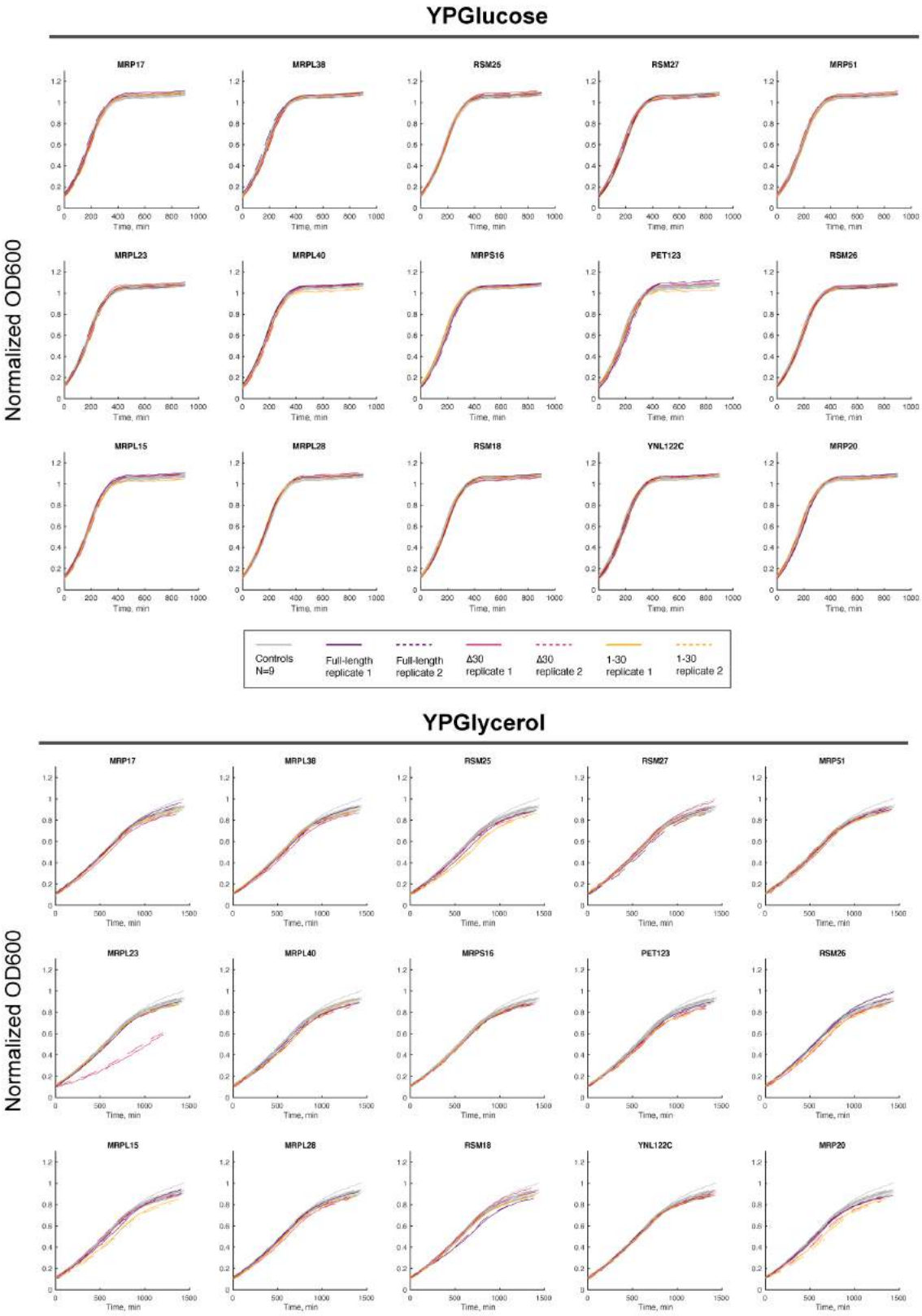


Fig. S5. MRP truncations in the genomic locus of one allele in diploid yeast strains have no dominant negative growth defects in fermentative media and rare defects in respiratory media despite having non-mitochondrial localizations (see Figure 1). Growth curves in fermentative (YPDextrose) and respiratory media (YPGlycerol) for yeast expressing each of the MRP truncations (in 2 replicates) are plotted besides WT controls not expressing any construct (grey lines); Curves are normalized by subtracting the minimal value of each to compensate for the background OD and then by plotting all curves only from the timepoint when the normalized OD reached 0.1 to compensate for different lag-times before the exponential phase; thus only the amplitudes and slopes of the growth curves can be compared but not the lag-times which are mostly attributed to inaccuracies of preparing the starting culture dilution.

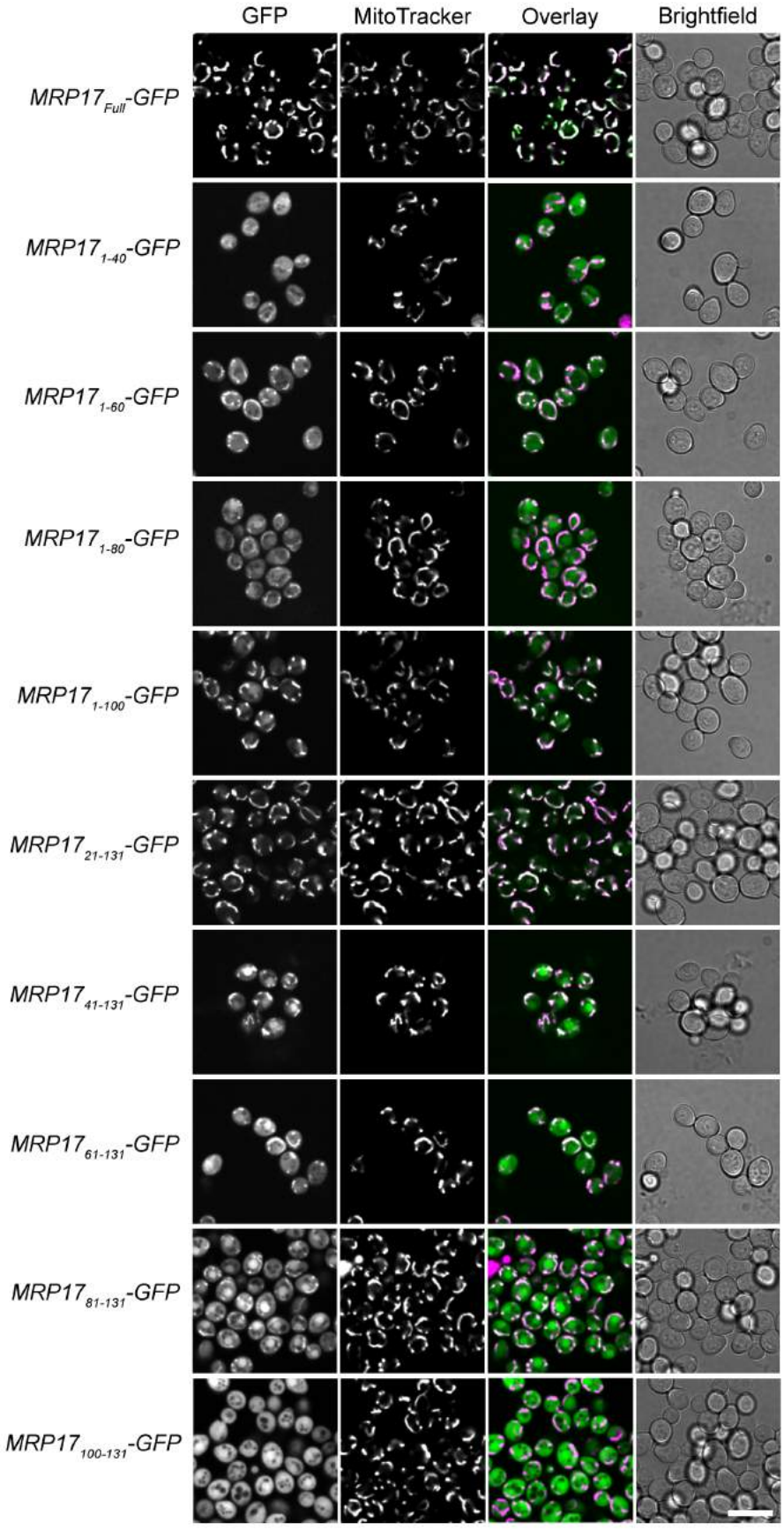


Fig. S6 - beginning, continues on the next page.

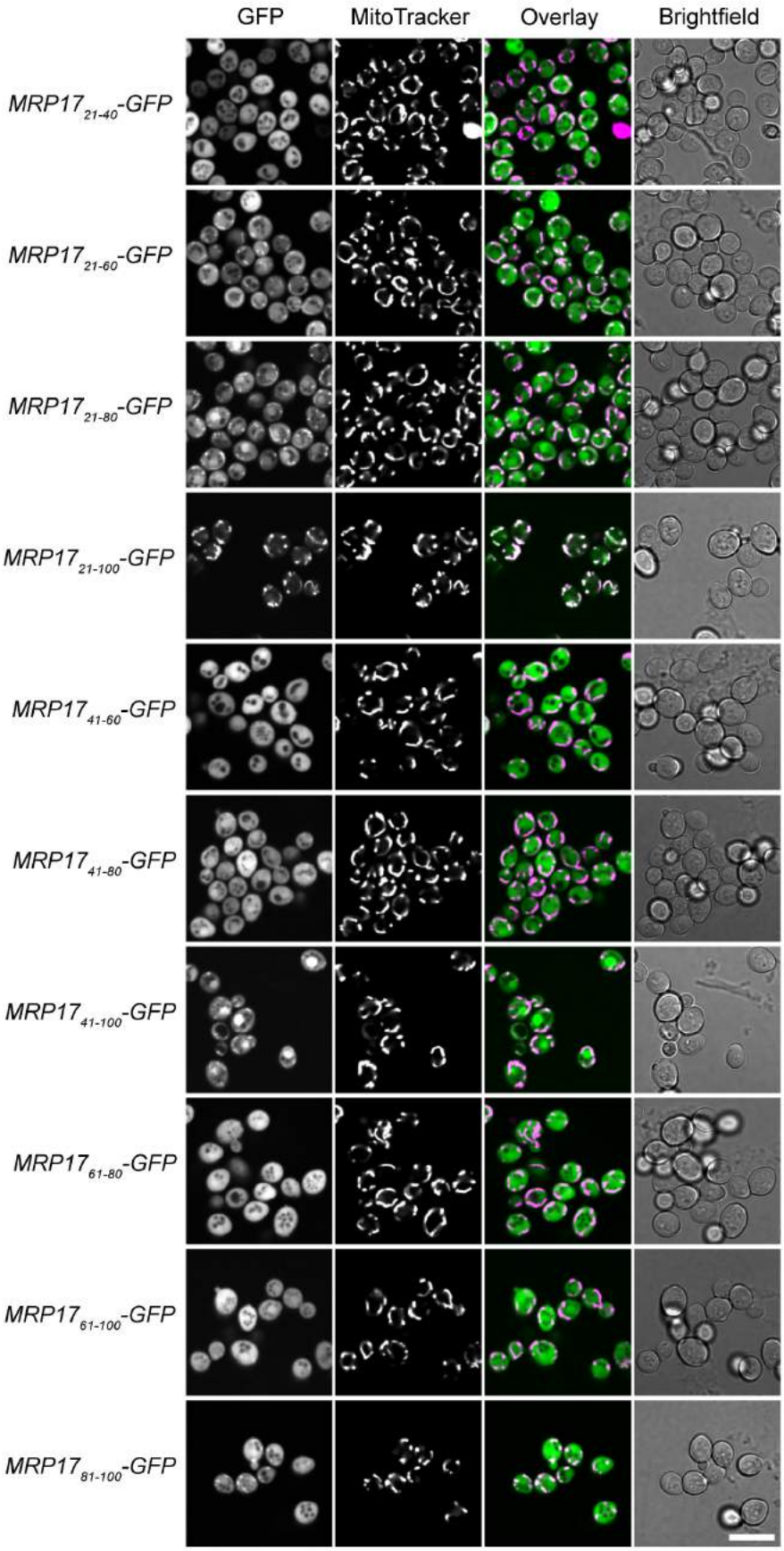


Fig. S6. The ability of different Mrp17 truncations to target GFP to mitochondria. Full sets of micrographs in different fluorescent channels for all the Mrp17 truncations including the ones shown in Fig. 2A. Scale bar 10 µm.

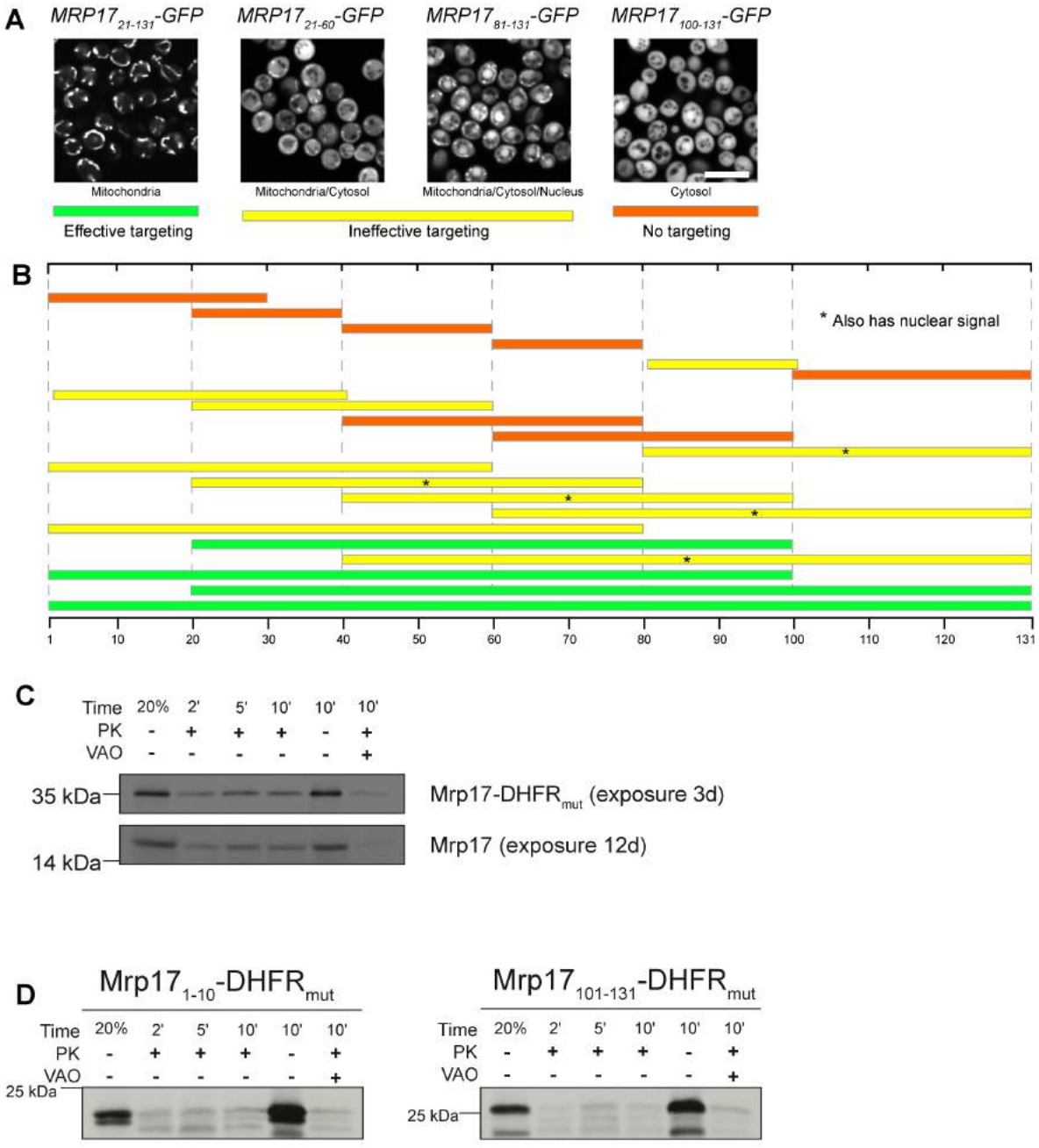


Fig. S7. The non-canonical targeting and translocation signal of Mrp17 is located between amino acids 30 and 60. (A and B) In vivo characterization of mitochondrial targeting capacity of different Mrp17 truncations fused to GFP: (A) GFP localization examples taken from Fig. S6 and demonstrating effective targeting with only mitochondrial GFP signal, ineffective targeting with mitochondrial GFP signal accompanied by strong cytosolic signal, ineffective targeting with additional nuclear signal, and no detectable mitochondrial targeting with exclusively cytosolic GFP. Scale bar is 10 µm. (B) localization summary of different Mrp17 truncations fused to GFP and colored according to the color-code for effective, ineffective, and no targeting introduced in panel (A), truncations additionally targeted to the nucleus are marked with asterisks. (C) Mrp17-DHFRmut is translocated into isolated mitochondria at the same rate as WT Mrp17 but gives better signal in the autoradiograph. (D) In vitro import assays for additional truncations of Mrp17 fused to DHFRmut not shown in Fig. 2B, import was performed as described in the legend for Fig.2.

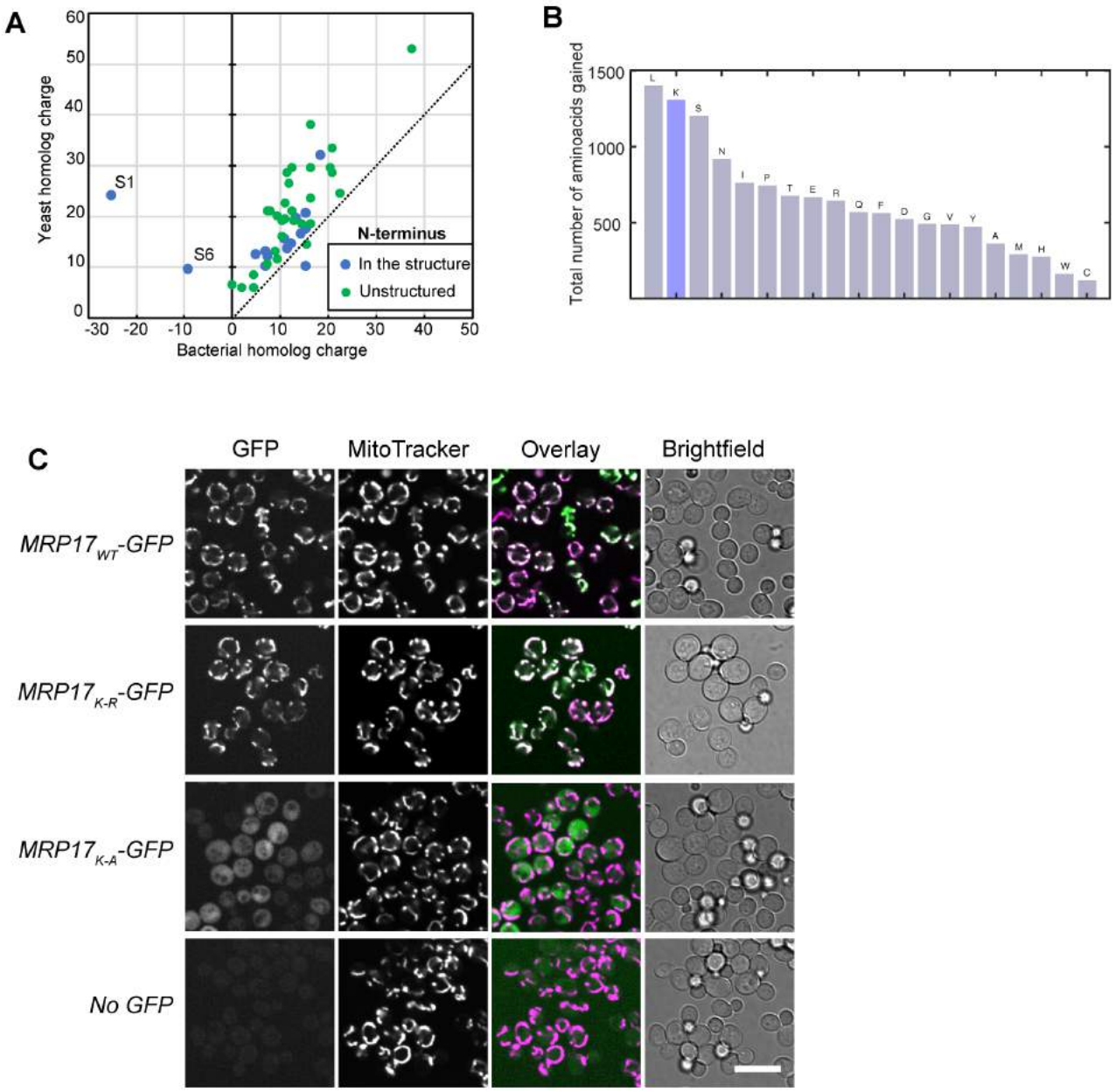


Fig. S8. Mrp17 sequence features important for targeting and translocation to mitochondria.
(A) Ribosomal proteins are positively charged and mitochondrial proteins acquired even more positive net charge compared to their bacterial homologs; (B) Total amino acid gain of MRPs (calculated as the difference between total count of each amino acid in all yeast MRPs, including mitochondria-specific, and all bacterial RPs) compared to bacterial RPs shows over-representation of lysines (K); (C) Lysines in Mrp17 are not important for mitochondrial targeting and can be substituted with arginines, same micrographs as in Fig. 3C shown in all channels beside micrographs of yeast not expressing any GFP (bottom row) as control for autofluorescence relative to cytosolic signal. All micrographs in the GFP channel are shown at the same contrast and brightness for comparison; Scale bar of all micrographs is 10 μm.

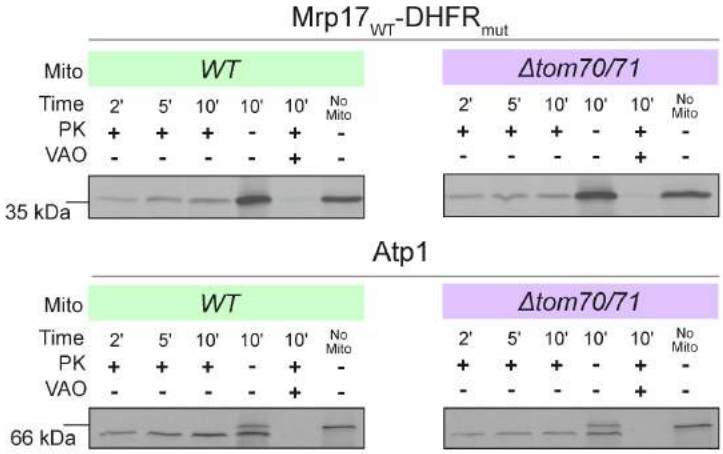


Fig. S9. The translocation pathway of Mrp17 does not involve Tom70/71. *In vitro* translocation of Mrp17-DHFRmut and control protein Atp1 into mitochondria isolated from WT and a *Δtom70/71* mutant showing no Tom70/71-dependence for Mrp17 translocation, import was performed as described in the legend for Fig.2.

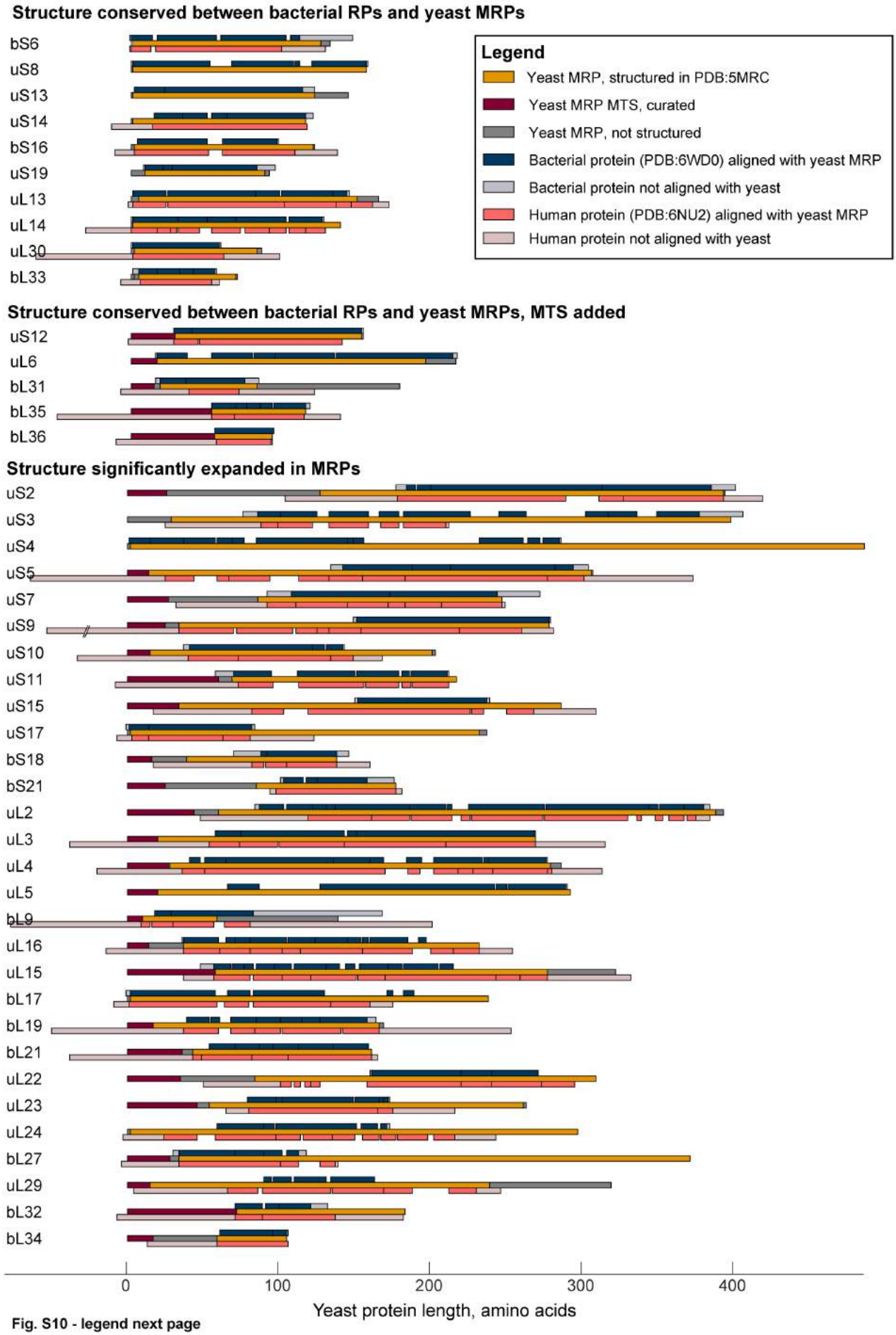
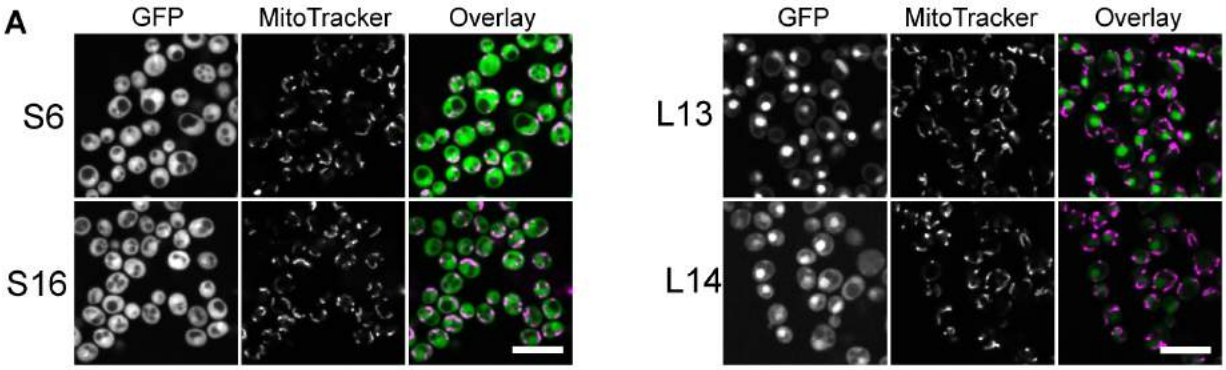
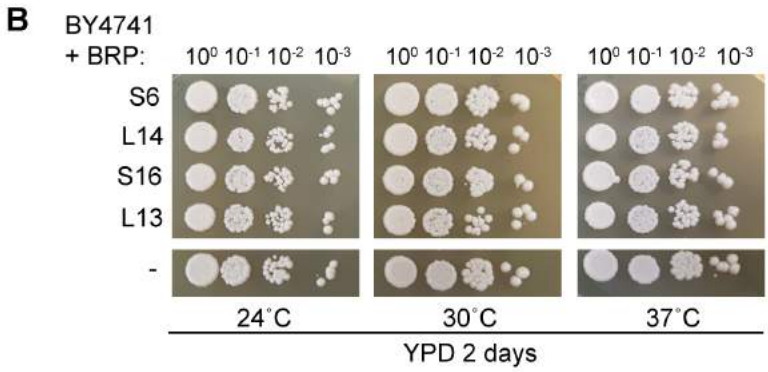


Fig. S10 - legend next page

Fig. S10. Structural alignments of yeast and human MRPs and their bacterial homologs. The structures of yeast MRPs that have bacterial homologs were aligned to their bacterial and human (when present) homologs using flexible structural alignment tool FATCAT (Li et al., 2020), and the resulting alignment was plotted relative to the length of yeast proteins, deletions in the yeast proteins relative to human and bacterial homologs are not plotted and the positions of corresponding insertions within human and bacterial sequences are indicated as solid black lines. MTS of yeast proteins were annotated according to Table S1. MPRs are grouped by structure conservation relative to bacterial homologs and then sorted by name.





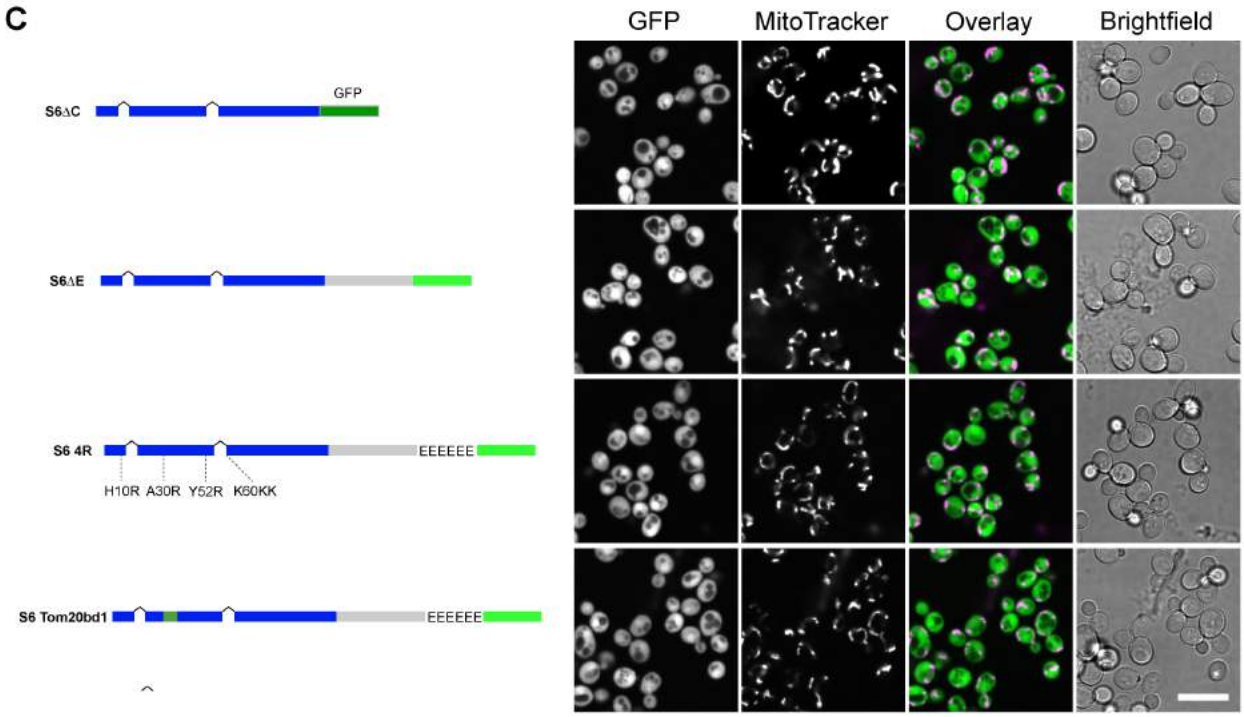


Fig. S11 - continues on the next page

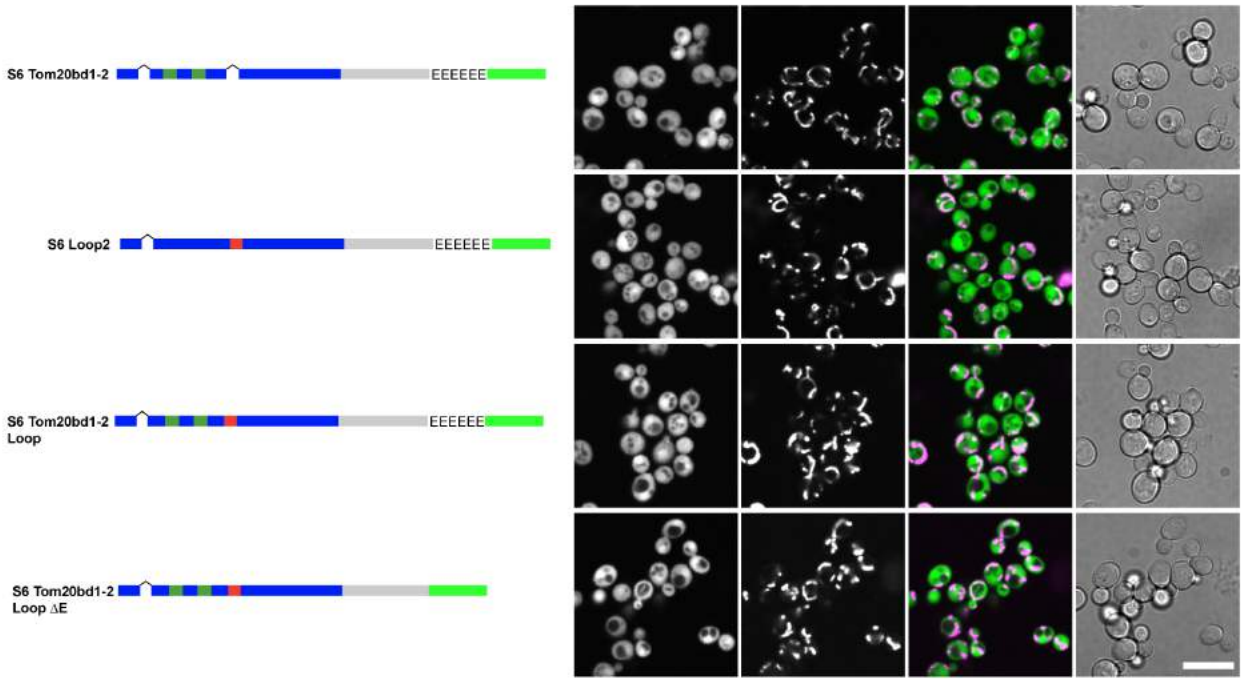


Fig. S11. Comparison of Mrp17 and its bacterial homolog EcS6 in the structural context. (A) Expression of bacterial homologs of MRPs in yeast, same micrographs as in Fig. 5C shown in all channels. (B) Drop dilution growth assay for all the strains from panel A and WT control performed on rich fermentative media at different temperatures. (C) Schematic representations of chimeric constructs incorporating Mrp17 features into EcS6 (corresponds to schematic on Fig. 5D, top) and fluorescence micrographs of yeast expressing each respective construct. Scale bar in all panels is 10 μm.

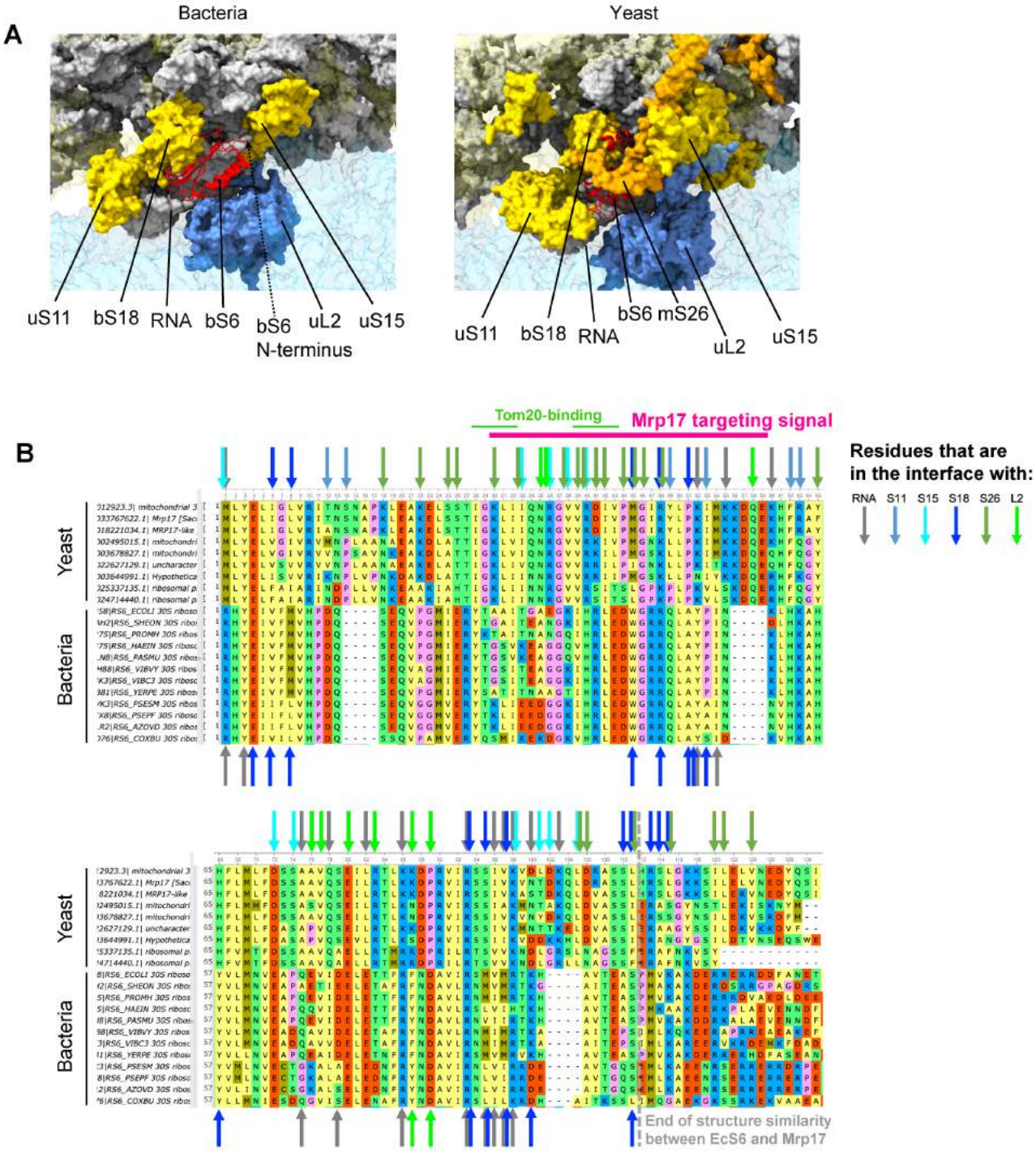


Fig. S12. Mrp17 and EcS6 in the context of ribosome structures. (A) A comparison of the structural environment of the S6 protein in bacterial ribosome (PDB:6WD0, left) and yeast mitoribosome (PDB:5MRC, right). RNA is shown in grey, small subunit (SSU) proteins in yellow (highlighted are solid and others are transparent), S6 is red, all large subunit (LSU) components are in blue (highlighted are solid and others are transparent), mitochondria-specific protein mS26 is in orange. (B) Structural alignment of Mrp17 (first sequence) and its yeast homologs with EcS6 (first out of bacterial) and its bacterial homologs highlighting involvement of each amino acid in protein-protein and protein-RNA interfaces (arrows, see legend on the figure) in yeast mitoribosome (on top) and bacterial ribosome (in the bottom), amino acids are numbered relative to Mrp17, thus the N-terminal Met in bacterial proteins is not shown.