PNAS
www.pnas.org

Supplementary Information for

The pentatricopeptide repeat protein Rmd9 recognizes the dodecameric element in the 3′-UTR of yeast mitochondrial mRNAs

Hauke S. Hillen, Dmitriy A. Markov, Ireneusz D. Wojtas, Katharina B. Hofmann, Michael Lidschreiber, Andrew T. Cowan, Julia L. Jones, Dmitry Temiakov, Patrick Cramer, and Michael Anikin

Patrick Cramer, Michael Anikin
Email: patrick.cramer@mpibpc.mpg.de; anikinmi@rowan.edu

This PDF file includes:

Supplementary Materials and Methods

Preparation of the plasmids used for the expression of recombinant proteins. Recombinant Rmd9 was expressed in *E. coli* using the plasmid pMA27O2THP. To prepare this plasmid, the coding sequence corresponding to Δ51-Rmd9 was PCR-amplified from yeast chromosomal DNA (strain BY4743) with primers Rmd-e-m-pciI and Rmd-e-xhoI and ligated into pSTBlue-1 (Novagen). The gene was excised with *Pci*I and *Xho*I and inserted into a pTrcHisC vector (Invitrogen) at *Nco*I and *Xho*I sites. The resulting construct expressed MSH_6-Δ51-Rmd9, which is referred to as Rmd9 in all *in vitro* experiments in this study. For tighter control of the target gene, two additional elements were incorporated in the plasmid by sequential site-directed mutagenesis using a QuikChange II kit (Agilent). A copy of the lac operon O2 operator (1, 2) was added downstream of the RMD9 gene (primers Trc-O2 and Trc-O2c) and a tHP transcription terminator (3) was inserted following the *lacI* gene (primers TrcTHP and TrcTHPc), resulting in pMA27O2THP.

The yeast mitochondrial degradosome, mtEXO, comprises two major subunits, the 3′-exoribonuclease Dss1 and the NTP-dependent helicase Suv3. Dss1 was produced by expressing the plasmid pMA31 in *E. coli*. To prepare the plasmid, a segment of *S. cerevisiae* BY4743 genomic DNA was amplified by PCR using primers Dss1-PciI and Dss1-PstI. The amplicon was ligated into vector pT7Blue (Novagen) to give plasmid pMA33. The gene was cut out with *Pci*I and *Pst*I enzymes and ligated into pTrcHisA (Invitrogen) at *Nco*I and *Pst*I sites to give plasmid pDW1. Sequencing of the plasmid revealed a point mutation in the target gene (A219P) relative to the reference NP_014014. The mutation was corrected by site-directed mutagenesis using a QuikChange II kit (Agilent) and primers Fix_Dss1 and Fix_Dss1c. This resulted in pMA31, which encoded an N-terminally truncated and His-tagged form of Dss1 (MSH_6-Δ51-Dss1). *E. coli* expression of Suv3 was directed by the plasmid pMA35. A fragment of yeast genomic DNA carrying the *SUV3* gene was PCR-amplified using primers Suv3-NcoI and Suv3-XhoI. The amplicon was inserted into pTrcHisA vector (Invitrogen) at *Nco*I and *Xho*I sites resulting in plasmid pMA32. The sequence encoding the first 50 amino acids of Suv3 in pMA32 was deleted by site-directed mutagenesis using a QuikChange II kit and mutagenesis primers Del50-Suv3 and Del50c-Suv3. This gave plasmid pMA35, which directed the expression of the Δ50-Suv3 truncation fused at the N-terminus to the purification tag MAH_6.

Crystallization of Rmd9-RNA complexes. Rmd9 was purified by heparin affinity chromatography as described in Methods. A portion of the eluate containing 8.25 mg of the protein was concentrated on an Ultracel®-50K centrifugal filter (Merck Millipore) to a volume of 0.4 ml and then diluted with 1.2 ml of 10 mM Tris·HCl (pH 7.9) to reduce the concentration of NaCl. RNA20 was added (43 μl of 3.1 mM solution; 10 % molar excess) and the mixture was incubated for 20 min at 30 °C. The formed complex was separated from the excess of free RNA on a HiLoad 16/600 Superdex 200 PG column (GE Healthcare) eluted with gel filtration buffer (20 mM Tris·HCl pH 7.9, 100 mM NaCl, 5 mM $MgCl_2$, 5 % glycerol, 10 mM DTT). The fractions containing pure Rmd9-RNA20 complex were concentrated on a centrifugal filter to approximately 10 mg/ml, the concentrate was distributed in 25 μl aliquots, flash-frozen in liquid N_2, and stored at -80 °C. The Rmd9-RNA16 complex was prepared in the same way. The crystals were grown at 20 °C using the sitting drop vapor diffusion method. To crystallize the Rmd9-RNA16 complex, 3 μl of the well solution (33 mM Na cacodylate, pH 6.5, 67 mM NaCl, 32 mM Li_2SO_4, 66 mM calcium acetate, 3.4 mM $MgCl_2$, 10 % glycerol, 13 mM β-mercaptoethanol, 5.9 % PEG 8000) was mixed with 3 μl of the complex solution (7 mg/ml in 14 mM Tris·HCl, pH 7.9, 100 mM NaCl, 5 mM $MgCl_2$, 8 % glycerol, 7 mM DTT, 6 mM β-mercaptoethanol). The drop was incubated over 0.5 ml of the well solution. The crystals appeared overnight and were allowed to grow for 8 days. The liquid in the drop was gradually replaced with cryoprotectant solution (50 mM Na cacodylate, pH 6.5, 100 mM NaCl, 10 mM calcium acetate, 15 % PEG 8000, 25 % glycerol) and the crystals were frozen in liquid N_2. To crystallize the Rmd9-RNA20 complex, 2 μl of well solution (36 mM Na cacodylate, pH 6.5, 65 mM NaCl, 32 mM Li_2SO_4, 71 mM calcium acetate, 3.2 mM $MgCl_2$, 6.5 % glycerol, 6.5 mM DTT, 6.4 % PEG 8000) was mixed with 2 μl of the complex solution (7.5 mg/ml in 20 mM Tris·HCl, pH 8.0, 100 mM NaCl, 5 mM $MgCl_2$, 5 % glycerol, 10 mM DTT). The drop was incubated over 0.5 ml of the well solution. The crystals appeared overnight and were allowed to grow for 7 days. The liquid in the drop was gradually substituted with cryoprotectant solution (50

mM Na cacodylate pH 6.5, 100 mM NaCl, 15 % PEG 8000, 25 % glycerol) and the crystals were frozen in liquid N_2.

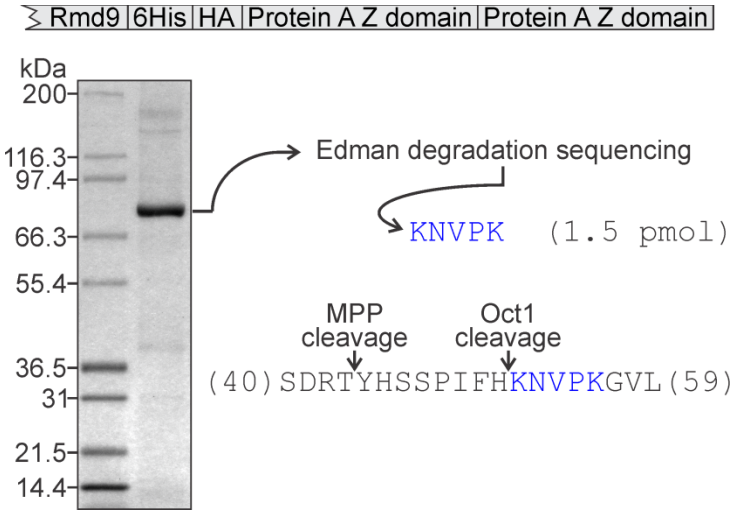


Fig. S1. Position K52 is the N-terminus of the mature form of Rmd9. The cytosol-translated Rmd9 precursor is expected to undergo processing upon import into the mitochondria. We determined the N-terminus of the mature protein by taking the approach we described previously (4). Briefly, BY4743 yeast cells were transformed with plasmid BG1805-Rmd9 (GE Healthcare Open Biosystems) encoding Rmd9 fused to a C-terminal purification tag (the composition of the tag is indicated at the top). The cells were grown, and the Rmd9 fusion was partially purified from the mitochondrial fraction under denaturing conditions using Ni-beads. The proteins in the preparation were then resolved by SDS-PAGE and stained with coomassie (the image on the left). The protein in the indicated band was submitted to Edman N-terminal sequencing, which returned a sequencing read as shown on the right. The read (blue) corresponded to the K52-K56 fragment of the Rmd9 sequence as indicated in the scheme below. The scheme shows a region of the Rmd9 sequence between amino acids 40 and 59. These data suggest that Rmd9 may undergo two-step processing, in which the precursor is first cleaved by MPP after T43 and then an additional octapeptide is removed by Oct1 cleaving after H51. K52 was also identified as the N-terminus of mature Rmd9 by previous high-throughput LC-MS/MS analysis (5).

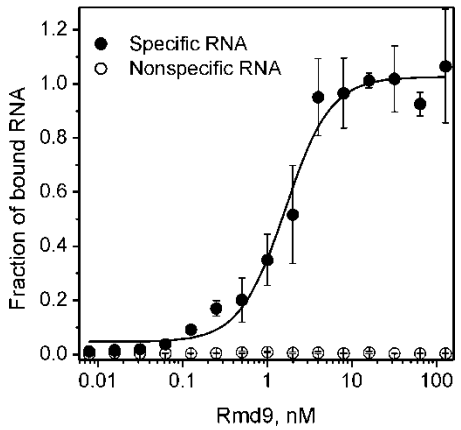


Fig. S2. Rmd9 binds to a dodecamer-containing RNA with an affinity in the low-nM range. We compared the effect of recombinant Rmd9 on the stability of two short RNA probes in the presence of RNase I. The probes were 5′-[^{32}P]-labeled and carried either the dodecameric element (RNA20) or a scrambled sequence (RNA20c). The radioactive reporter was placed within the Rmd9 RNase I footprint (see Fig. 1c), presumably making it inaccessible by the RNase while the probe was specifically associated with the protein. Rmd9 was added at varying concentrations (0.008 to 128 nM) to each probe (1 nM) in 20 μl of buffer containing 100 mM NaCl, 50 mM Tris·HCl (pH 7.9), 10 mM $MgCl_2$, and 1 mM DTT. The mixtures were incubated for 20 min at 30 °C in the presence of 200 nM nonspecific competitor (yeast tRNA) and the formed protein-RNA complexes were challenged with 25 U of RNase I (New England Biolabs) for 15 s. Under these conditions, the RNase could digest about 95 % of unprotected RNA. The digestion reactions were stopped by the addition of 20 μl of preheated gel loading solution (90% formamide and 50 mM EDTA spiked with bromophenol blue and xylene cyanol FF) and incubation at 95 °C for 2 min. The products of the reactions were resolved by 7M urea denaturing PAGE (10% acrylamide:bis-acrylamide, 19:1), visualized by phosphor imaging on a Typhoon 9200 scanner (GE Healthcare), and quantified using ImageQuant 5.2 (Molecular Dynamics). For each reaction, the fraction of bound RNA was determined as the intensity of the full-length probe divided by that in the RNase-untreated control. This allowed us to plot the binding curves shown in the graph (n=3). As seen in the plots, Rmd9 was not able to appreciably protect the nonspecific RNA probe even at high concentrations. In contrast, the dodecamer-carrying probe was perfectly protected by the protein. We estimated from the binding curve that the apparent Rmd9-dodecamer binding affinity is in the low nanomolar range.

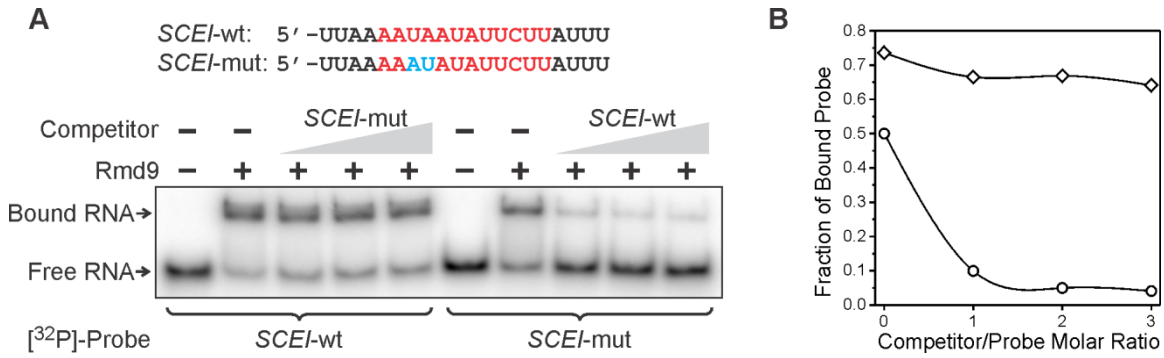
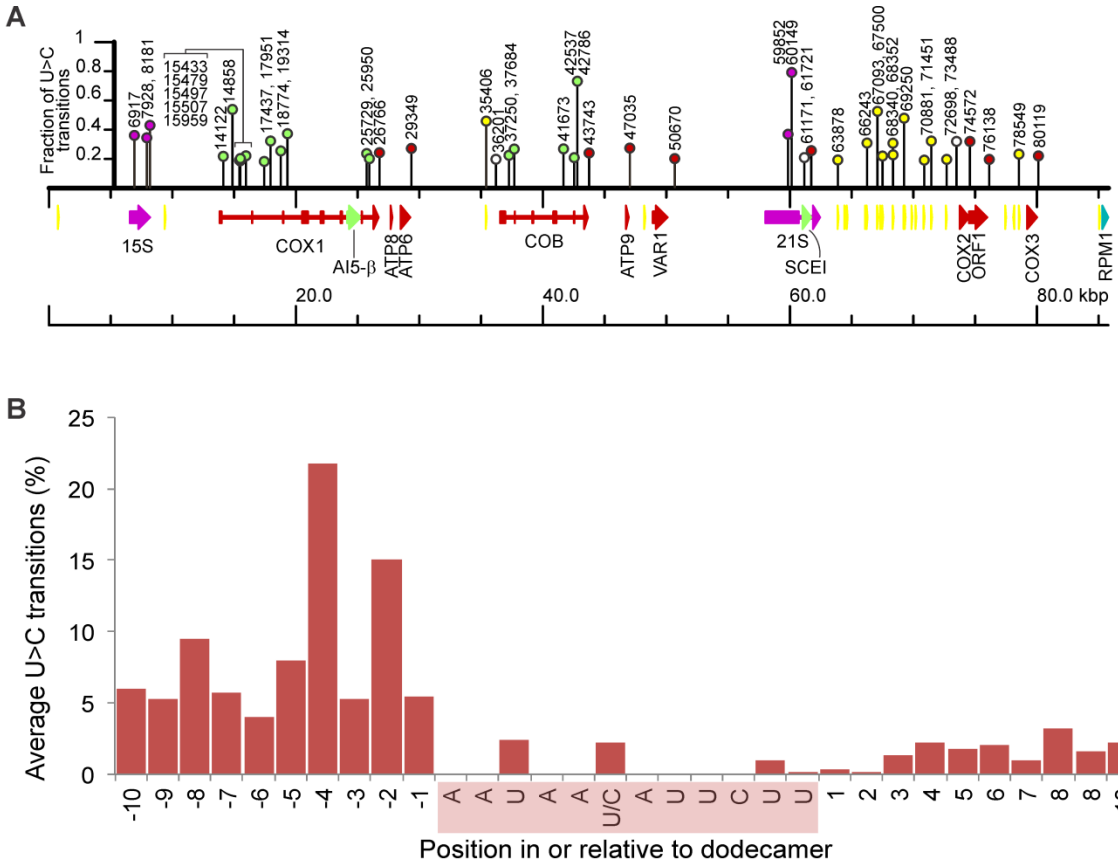


Fig. S3. Mutations within the dodecameric sequence can significantly weaken its interaction with Rmd9. (A) Two RNA probes were allowed to compete for binding to Rmd9. One of them, *SCEI*-wt, represented a dodecamer-containing fragment of the sequence at the end of the *SCEI* ORF (the sequence is shown at the top of the panel with the dodecamer element indicated in red). The second probe, *SCEI*-mut, contained a 3U>A,4A>U mutation within the dodecamer as indicated in blue. In the five left-side lanes, formation of the Rmd9-5′-[^{32}P]-*SCEI*-wt complex was examined by EMSA in the absence or presence of *SCEI*-mut. In a reciprocal experiment, the influence of *SCEI*-wt on binding of 5′-[^{32}P]-*SCEI*-mut to Rmd9 was tested (the right side of the image). For this experiment, native 10 % acrylamide:bisacrylamide (37.5:1) gels were cast in 0.5× TBE using a BioRad Mini-PROTEAN Tetra system. Rmd9-RNA complexes were allowed to form for 20 min at 30 °C in 10 μl of binding buffer (20 mM Tris·HCl, pH 7.9, 50 mM NaCl, 5 mM $MgCl_2$, 10 % glycerol, 2 mM β-mercaptoethanol). Rmd9 and the 5′-[^{32}P]-RNA probes were present at 300 nM concentration each. Where indicated, the unlabeled RNA competitors were added (300, 600, and 900 nM). The gels were run in 0.5× TBE at room temperature (100 V, 15 min) and the labeled RNA species were visualized by phosphor imaging using a Typhoon 9200 scanner (GE Healthcare). (B) The RNA bands in the image in panel A were quantified using ImageQuant 5.2 software (Molecular Dynamics) and the efficiency of complex formation was calculated and plotted as the fraction of Rmd9-bound [^{32}P]-RNA at different competitor/probe ratios. The diamonds indicate the experiment in which [^{32}P]-*SCEI*-wt competed with *SCEI*-mut and the circles indicate the reciprocal experiment.



	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	A	A	U	A	A	U/C	A	U	U	C	U	U	1	2	3	4	5	6	7	8	9	10
AI5β	-	8	-	3	1	-	-	2	4	-			3			2		0	0		1	1	1	1	1	3	1	1	2	1	-	1
COX1	-	10	-	-	10	-	24	-	-	7			2			2		0	0		0	0	-	0	-	-	-	-	-	15	-	0
ATP6	-	-	-	-	-	-	27	3	8	4			3			3		0	0		1	0	0	0	-	0	-	3	-	-	2	-
COB	-	1	-	-	3	-	-	16	24	14			3			3		0	0		0	0	-	-	-	-	-	3	-	-	1	-
ATP9	10	-	-	3	3	-	-	-	27	6			3			-		0	0		0	0	-	0	-	-	0	0	-	2	0	5
VAR1	-	-	-	20	-	-	-	12	-	-			4			2		0	0		1	0	-	-	3	2	3	1	1	-	3	3
SCEI	-	-	16	5	3	8	25	1	-	-			2			1		0	0		2	0	-	0	0	1	2	3	-	1	0	2
COX2	2	-	-	2	-	-	32	2	-	2			1			2		0	0		1	0	-	-	-	5	-	2	-	-	3	-
ORF1	-	-	3	-	-	-	19	4	12	4			2			3		0	0		1	0	0	0	-	0	3	-	0	0	0	-
COX3	-	2	-	2	4	-	22	2	-	1			1			2		0	0		1	0	-	-	-	-	-	4	-	0	-	-
Average %	6	5.3	9.5	5.7	4	8	21.7	5.3	15	5.4			2.4			2.2		0	0		1	0.2	0.3	0.2	1.3	2.2	1.8	2	1	3.2	1.6	2.2

Fig. S4. Rmd9 contacts RNA in the vicinity of the dodecameric sequence *in vivo*. Yeast cells were grown in the presence of 4-thiouracil to photosensitize the RNA. The cells expressed chromosomally encoded C-terminally TAP-tagged Rmd9. The addition of this tag to Rmd9 has been previously reported not to disrupt the function of the protein (6). The cells were UV-irradiated, and the photo cross-linked RNA was co-immunoprecipitated with anti-TAP antibodies. The RNA was trimmed by limited RNase T1 digestion, converted to cDNA, and deep-sequenced. The sequencing reads were aligned with the yeast mitochondrial genomic sequence and analyzed for the presence of U>C substitutions that map the locations of the sites of cross-linking. (A) The sites of high rate of U>C transitions are mapped onto a scheme of the *Saccharomyces cerevisiae* mitochondrial genome (NC_001224). The genomic elements are color-coded (yellow, tRNA; magenta, rRNAs; red, ORFs; green, ORFs inside introns). The exon structures of *COX1*, *COB*, and 21S rRNA are indicated by wider bars. The positions of the high-rate U>C transition sites are shown as lollipops. The lollipops colored red are associated with the dodecameric elements of mRNAs. Other lollipops are found in rRNAs (magenta), tRNAs (yellow), and introns (green). The three lollipops without color are located in the 5′-UTR of *COB*, within the *SCE1* ORF, and just upstream of *COX2* mRNA. The genomic positions of the sites are indicated. (B) The

graph shows the percentage of U>C substitutions in the sequencing reads within the region comprising the dodecameric motif and 10 nucleotides upstream and downstream, which was averaged over the 3′-UTRs of *AI5β*, *ATP6*, *ATP9*, *COB*, *COX1*, *COX2*, *COX3*, *ORF1*, *SCE1*, and *VAR1* mRNAs. The values were calculated for each position of the region as the sum of the percentages of U>C transitions in all of the above mRNAs divided by the number of Us found in the mRNAs at that position. The table below the graph shows the percentage of U>C transitions at all U-containing positions in the indicated mRNAs. The dodecameric sequence is highlighted in pink and additional ten nucleotides upstream and downstream of the dodecamer are included. For each mRNA, the positions that contain nucleotides other than U are indicated as "-".

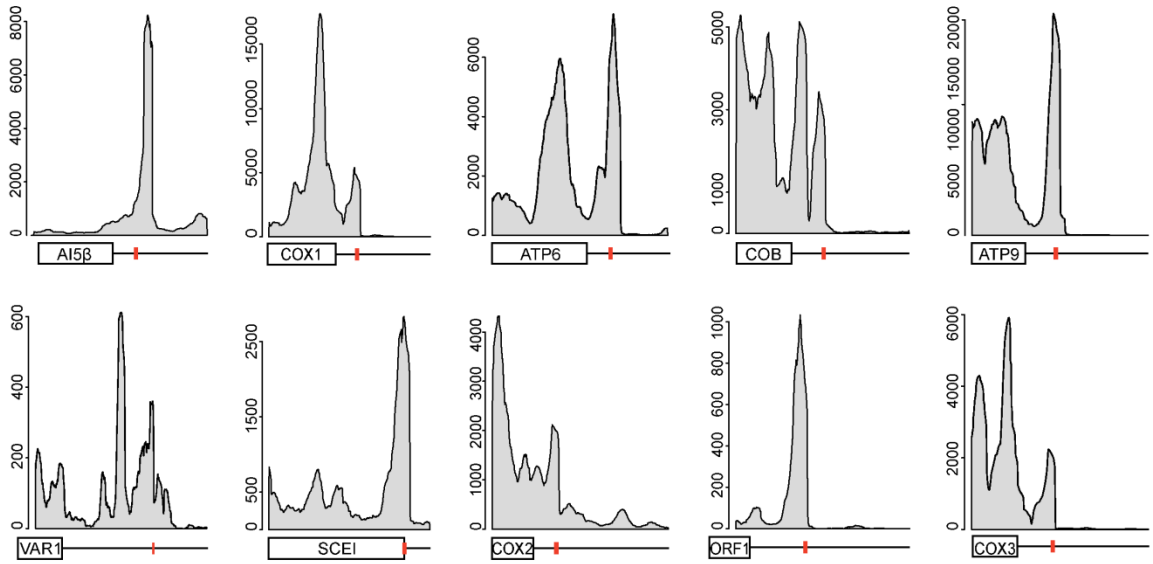


Fig. S5. Distribution of Rmd9 PAR-CLIP sequencing reads relative to the dodecamer elements. PAR-CLIP experiment was performed as in Fig. S4 and the resulting sequencing reads were aligned to the sequence of *Saccharomyces cerevisiae* mitochondrial genome. Portions of the alignment are plotted as the number of PAR-CLIP reads per base pair. The regions that correspond to ORFs are shown as boxes and genes are identified inside the boxes. Positions of the dodecamer elements in the 3′-UTRs are indicated as red bands. The PAR-CLIP sequence coverage is shown in a range of 500 base pairs for all genes except VAR1 (1000 base pairs).

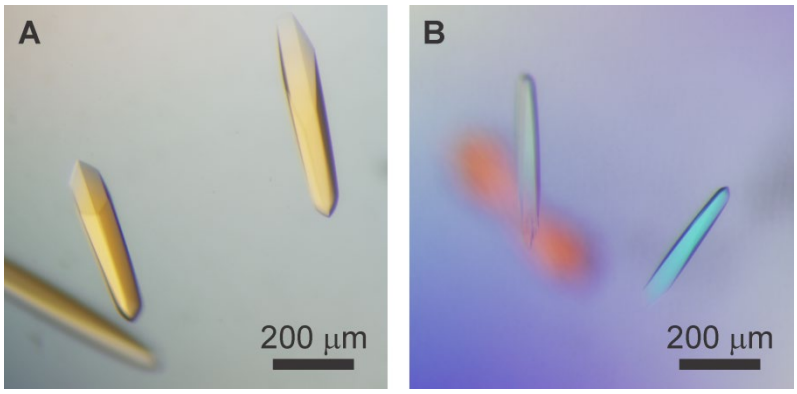


Fig. S6. Polarized light images of the crystals of the Rmd9-RNA complexes. (A) Rmd9-RNA16 complex. (B) Rmd9-RNA20 complex.

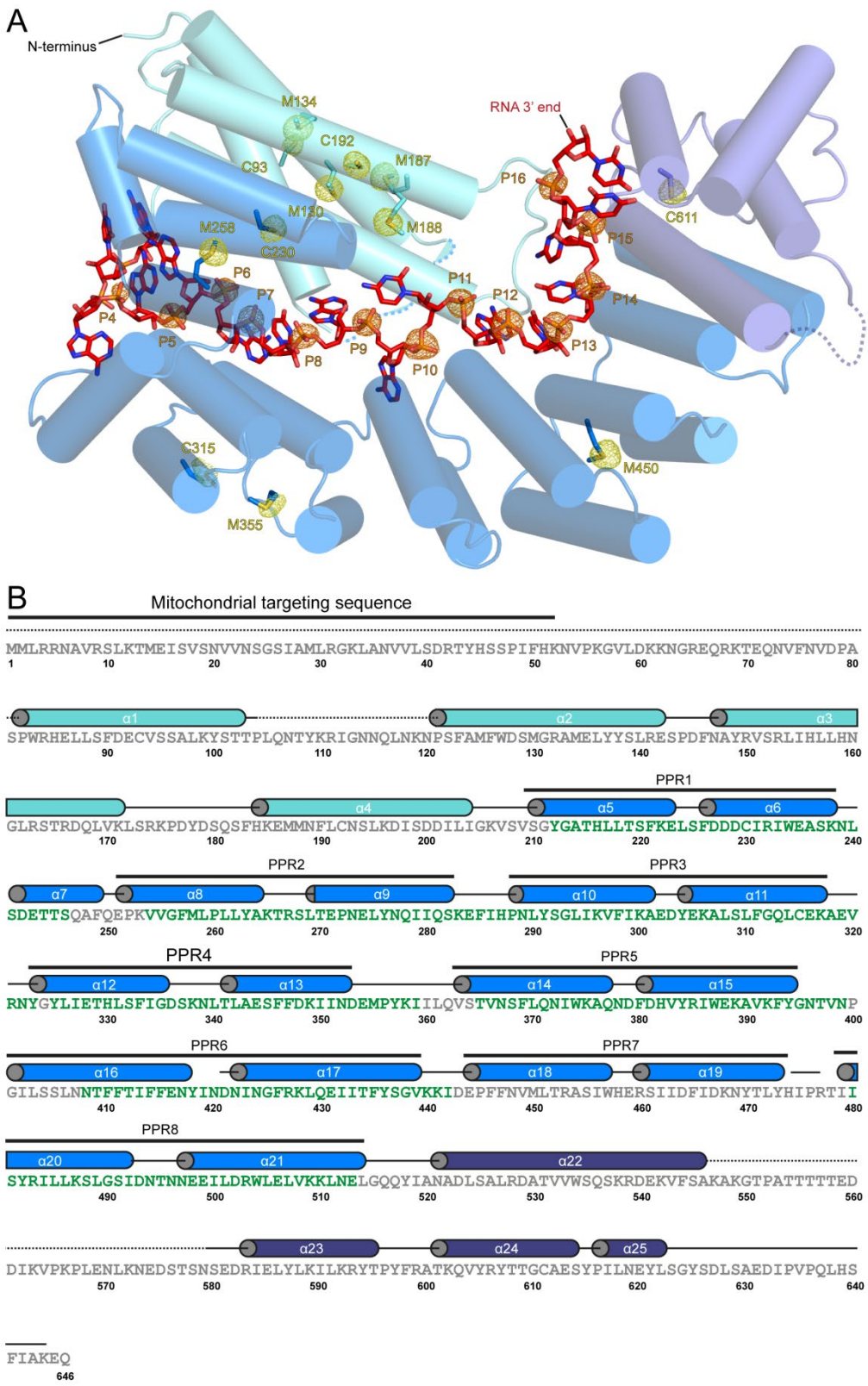


Fig. S7. Structure determination and secondary structure of Rmd9. (A) An anomalous difference Fourier map calculated using the refined model and the dataset used for phasing by native-SAD

is shown as mesh at 5.0σ and a carve of 2 Å. Peaks originating from sulfur atoms are colored in yellow and peaks originating from phosphorous atoms are colored in orange. The Rmd9-RNA structure is also shown with α-helices presented as cylinders and the RNA as sticks. The protein structure is shown at 50% transparency. (B) Primary sequence of Rmd9 with α-helices depicted as cylinders and colored as in Fig. 3. The pairs of the helices comprising the PPR motifs are indicated and numbered. Rmd9 was previously predicted to contain seven PPR motifs (7). Positions of the predicted motifs (amino acids 212-246, 254-288, 289-323, 325-359, 365-399, 408-422, and 480-513) are indicated in the sequence by green lettering. The regions visible in the structure and the missing regions are indicated (solid and dashed lanes, respectively).

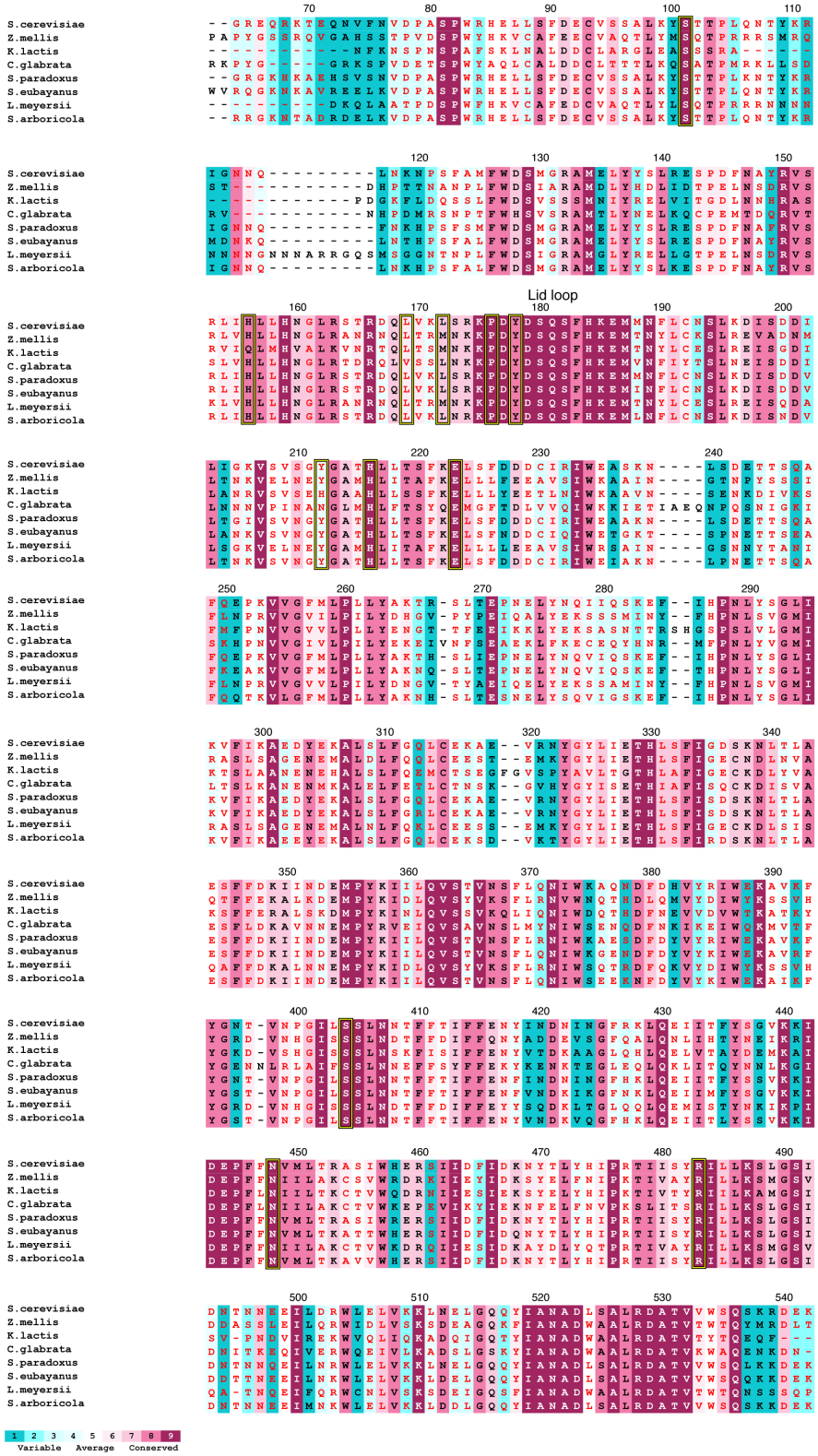


Fig. S8. Alignment of sequences of Rmd9 from eight fungal species that show the presence of dodecameric elements in their mitochondrial mRNAs. The following organisms and accession

numbers were used: *Saccharomyces cerevisiae* (P53140), *Zygosaccharomyces mellis* (GCE98470), *Kluyveromyces lactis* (XP_455556), *Candida glabrata* (XP_446292), *Saccharomyces paradoxus* (QHS73229), *Saccharomyces eubayanus* (XP_018222131), *Lachancea meyersii* (SCU84038), *Saccharomyces arboricola* (putative Rmd9 ORF located on the negative strand of chromosome 7, nucleotides 301215-303140; accession CM001569). The sequences were aligned using ClustalΩ within the MPI bioinformatics toolkit (8). The amino acids that are observed to coordinate the RNA in the structure of Rmd9-RNA complex (Fig. 3C) are boxed. The color scheme used to convey the degree of sequence conservation is indicated at the bottom.

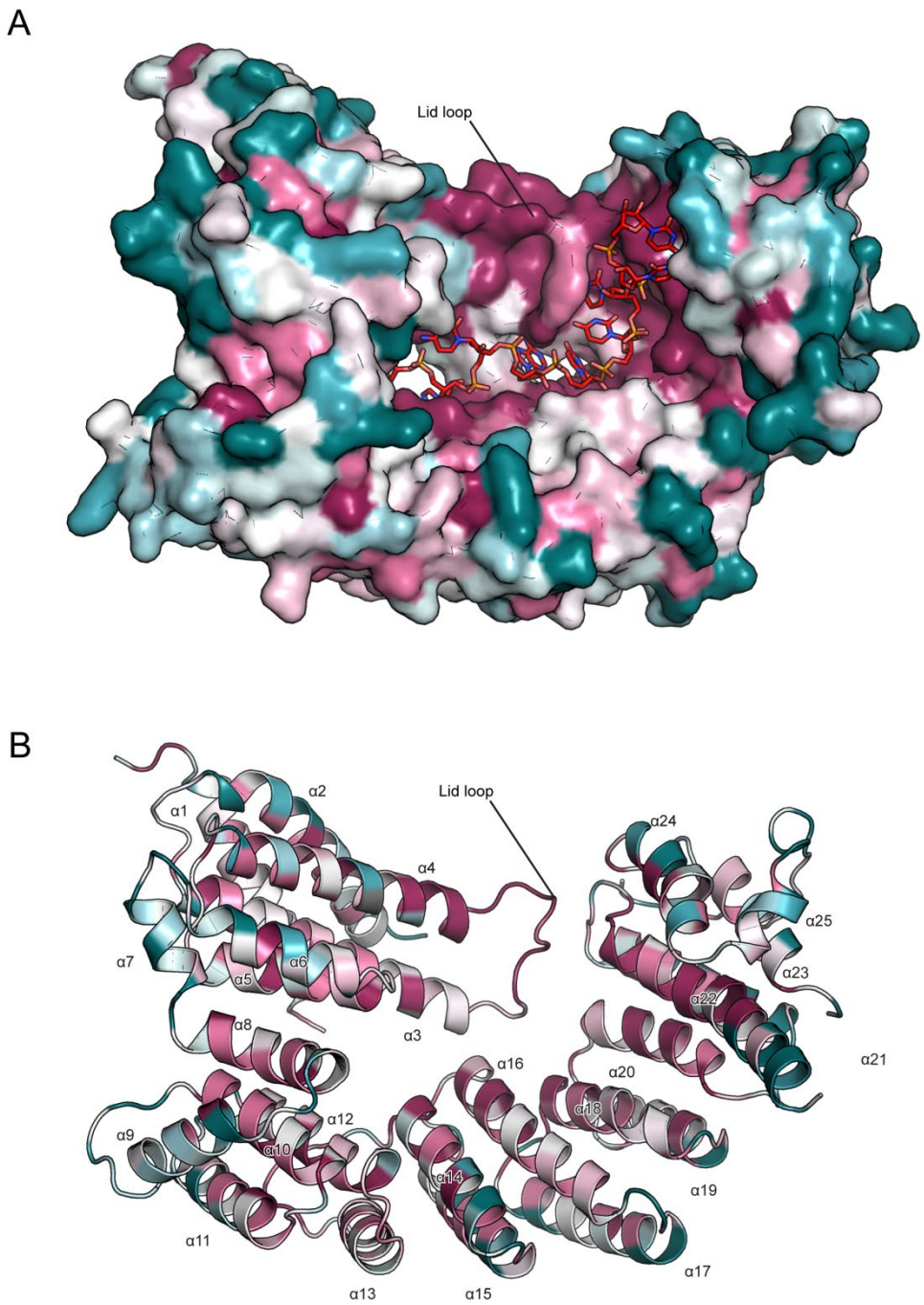


Fig. S9. Protein sequence conservation (see Fig. S8) transposed into the structure of Rmd9 using ConSurf (9). Both panels show the same orientation of the structure presented as a surface (A) or a cartoon (B). The RNA (red, blue, and orange in panel A) is not shown in panel B for clarity. The color scheme is as in Fig. S8.

Table S1. Sequences of oligonucleotides

Oligonucleotide	Sequence (5′ to 3′)
DNA oligonucleotides	
Rmd-e-m-pcil	CCATAAACATGTCACATCATCACCACCATCATAAAAATGTACCTAAGGGCG
Rmd-e-xhol	GCCCGCCTCGAGTTACTATTGCTCCTTAGCAATAAAGC
Trc-O2	CTGGTACCATATGGGAATTCGGGTTGTTACTCGCTCACATTTAAGCTTGGCTGT TTTGGCGG
Trc-O2c	CCGCCAAAACAGCCAAGCTTAAATGTGAGCGAGTAACAACCCGAATTCCCATAT GGTACCAG
TrcTHP	GGAAAGCGGGCAGTGAGCGGTACCCGATAAAAGCGGCTTCCTGACAGGAGGC CGTTTTGTTTTGCAGCCCACCTCAACGCAATTAATGTGAGTTAGC
TrcTHPc	GCTAACTCACATTAATTGCGTTGAGGTGGGCTGCAAAACAAAACGGCCTCCTGT CAGGAAGCCGCTTTTATCGGGTACCGCTCACTGCCCGCTTTCC
Dss1-Pcil	CAATGAACATGTCTCATCATCATCACCATCATAATGATCAAGCTACAGAGACAG
Dss1-PstI	CATATTCTGCAGTTATTATAGCTTTTCCAACTCTAACATTC
Fix_Dss1	GATTCCTCACAAATTACCTGCTGGAATCCATTCACTTATAC
Fix_Dss1c	GTATAAGTGAATGGATTCCAGCAGGTAATTTGTGAGGAATC
Suv3-NcoI	CTTAATCCATGGCACATCATCATCATCATCACATGGCACTTGTCAAATACAG
Suv3-XhoI	CTAATACTCGAGTTATTATGTACGCAATCTTCTTCTCG
Del50-Suv3	CACATCATCATCATCATCACAATTTACCCAAAAATGAACATTC
Del50c-Suv3	GAATGTTCATTTTTGGGTAAATTGTGATGATGATGATGATGTG
RNA oligonucleotides	
RNA16	AAAAUAACAUUCUUAA
RNA20	AUAAAAUAACAUUCUUAAUU
RNA20c	AUAACUAUAUUCAAAUAAUU
SCEI-wt	UUAAAAUAAUAUUCUUAUUU
SCEI-mut	UUAAAAAUAUAUUCUUAUUU
RNA48	AAUAAUAAUAAUAAUUAUAAUAAUAUUCUUAAAUAUAAUAAAGAUAUA
RNA54	AAUAAUUAUAAUAAUAUUCUUAAAUAUAAUAAAGAUAUAGAUUUAUAUUCUAUU
Chimeric DNA/RNA oligonucleotides	
RNA-bait-DD	5Biosg-$(dA)_{30}$r(AUAAAAUAACAUUCUUAAUU)[a]
RNA-bait-NS	5Biosg-$(dA)_{30}$r(AUAACUAUAUUCAAAUAAUU)

[a] 5Biosg, 2-(N-biotinyl-2-aminoethoxy)-ethoxyphosphoryl

Table S2. X-Ray data collection and refinement statistics.

	Rmd9-RNA20 native-SAD[b]	Rmd9-RNA20 native (PDB 7A9W)	Rmd9-RNA16 native (PDB 7A9X)
Data collection			
Space group	$P3_121$	$P3_121$	$P3_121$
Cell dimensions			
a, *b*, *c* (Å)	106.68, 106.68, 128.05	106.67, 106.67, 128.16	106.25, 106.25, 128.52
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120
Wavelength	2.0751	1.0000	1.0000
Resolution (Å)[a]	49.24 - 2.80 (2.89 - 2.80)	49.24 – 2.55 (2.61 – 2.55)	46.00 – 2.45 (2.50 – 2.45)
R_{meas}	0.28 (9.04)	0.09 (3.01)	0.08 (2.81)
$I/\sigma(I)$	45.52 (1.40)	18.16 (0.87)	18.11 (0.85)
$CC_{1/2}$	100.0 (78.1)	100.0 (37.3)	100.0 (37.8)
Completeness (%)	99.9 (99.7)	99.9 (999)	100.0 (100.0)
Redundancy	277.2	10.8	10.7
Refinement			
Resolution (Å)		46.19 – 2.55	46.01 – 2.45
No. reflections		27978	31347
R_{work} / R_{free}		19.42 / 23.47	19.81 / 22.95
No. atoms		4593	4536
Macromolecules		4506	4439
Ligand/ion		19	28
Water		68	69
B factors			
Macromolecules		96.40	90.02
Ligand/ion		131.18	112.10
Water		84.59	76.20
r.m.s deviations			
Bond lengths (Å)		0.002	0.002
Bond angles (°)		0.39	0.43
Ramachandran Preferred/allowed/ disallowed (%)		97.01 / 2.79 / 0.20	97.99 / 1.61 / 0.40

[a] Values in parentheses are for highest-resolution shell.
[b] A total of 27 datasets from two crystals were merged.

SI References

1. S. Oehler, E. R. Eismann, H. Krämer, B. Müller-Hill, The three operators of the lac operon cooperate in repression. *The EMBO Journal* **9**, 973-979 (1990).
2. J. Muller, S. Oehler, B. Müller-Hill, Repression of lac promoter as a function of distance, phase and quality of an auxiliary lac operator. *J Mol Biol* **257**, 21-29 (1996).
3. A Krebber, J Burmester, A Pluckthun, Inclusion of an upstream transcriptional terminator in phage display vectors abolishes background expression of toxic fusions with coat protein g3p. *Gene* **178**, 71-74 (1996).
4. J. L. Jones*, et al.* Yeast mitochondrial protein Pet111p binds directly to two distinct targets in *COX2* mRNA, suggesting a mechanism of translational activation. *Journal of Biological Chemistry* **294**, 7528-7536 (2019).
5. F. N. Vögtle*, et al.* Global analysis of the mitochondrial N-proteome identifies a processing peptidase critical for protein stability. *Cell* **139**, 428-439 (2009).
6. E. H. Williams, C. A. Butler, N. Bonnefoy, T. D. Fox, Translation initiation in *Saccharomyces cerevisiae* mitochondria: Functional interactions among mitochondrial ribosomal protein Rsm28p, initiation factor 2, methionyl-tRNA-formyltransferase and novel protein Rmd9p. *Genetics* **175**, 1117-1126 (2007).
7. K.A. Lipinski, O. Puchta, V. Surendranath, M. Kudla, P. Golik, Revisiting the yeast PPR proteins—application of an iterative hidden Markov model algorithm reveals new members of the rapidly evolving family. *Molecular Biology and Evolution* **28**, 2935-2948 (2011).
8. L. Zimmermann*, et al.* A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J Mol Biol* **430**, 2237-2243 (2018).
9. H. Ashkenazy*, et al.* ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic acids research* **44**, W344-W350 (2016).