Supporting Information

MIEF1 Microprotein Regulates Mitochondrial Translation

Annie Rathore[†,§], Qian Chu[†], Dan Tan[†], Thomas F. Martinez[†], Cynthia J. Donaldson[†], Jolene K. Diedrich[±,‡], John R. Yates, III[‡], and Alan Saghatelian[,†]*

† Clayton Foundation Laboratories for Peptide Biology, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, United States

[±] Mass Spectrometry Core for Proteomics and Metabolomics, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, United States

§ Division of Biological Sciences, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093, United States

‡ Department of Molecular Medicine, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, United States

*To whom correspondence should be addressed. Email: asaghatelian@salk.edu (A.S.)

METHODS

Measuring Mitochondrial Complex 1 Activity. HeLa cells transfected with MIEF1-MP-FLAG construct (with pcDNA control) and MIEF1 siRNA (with siRNA control) were tested for the mitochondrial OXPHOS Complex I enzyme activity using the ab109721 assay kit (Abcam). Protein extracts from the samples were loaded to the microplate wells which were pre-coated with Complex 1 capture antibodies and incubated for 3 hours at room temperature. After the immobilization of the target in the wells, they were washed and assay solution containing NADH and Dye was added. Complex I activity was measured by the oxidation of NADH to NAD+ and the simultaneous reduction of a dye which led to increased absorbance at OD=450 nm. Bovine heart mitochondrial extract (ab110338) was used as a positive control.

Lipidomics Analysis. HeLa cells were transfected with MIEF1-MP-FLAG construct (with pcDNA control) and MIEF1 siRNA (with siRNA control) in triplicates and grown in DMEM media with delipidated fetal bovine serum (Cocalico Biologicals, #55-0115). The cell pellet was then lysed in MeOH: PBS: $CHCl_3$ (1:1:2) with d7 cholesterol and 13C-palmitic acid (PA) as internal standards. The samples were then homogenized followed by centrifugation at 2400g, 5 mins, 4°C to collect the bottom layer containing lipids for proteomics analysis. LC-MS/MS analyses were carried out on a Vanquish UHPLC system interfaced to a Q-Exactive Plus mass spectrometer (Thermo Fisher Scientific, Waltham, MA). Lipids were separated using a Bio-Bond 5U C4 column (Dikma). A 70 min gradient was used as follows: 0% B hold 5 min, 0-20% B in 0.1 min, 20-100% B in 50 min, hold 8 min, 100-0% B in 0.1 min, then hold 7 min. Solvent A consisted of 95:5 water: methanol, and solvent B was 60:35:5 isopropanol: methanol: water. 0.1% formic acid and 5 mM ammonium formate were added for positive ionization mode and 0.03% ammonium hydroxide was added for negative ionization mode. MS/MS spectra were

acquired in the data dependent mode with one full MS scan (resolution, 70K; AGC target, 1e6; mass range, 200-1500) followed by 5 MS2 scans (resolution, 17.5K; AGC target, $1e^5$; stepped normalized collision energy of 20, 30, 40). Lipids were identified using Lipid Search[1]. Precursor tolerance and product tolerance were set to 5 ppm and 10 ppm respectively. M-score threshold was set to 3. Quantification was performed using an in-house software.

Measuring Steady State Protein Levels using TMT Labeling. The crude mitochondrial fraction was isolated from HEK 293T cells transfected with the MIEF1-MP-FLAG construct and pcDNA control in triplicates, according to a reported protocol[2]. The mitochondrial pellet was resuspended in RIPA Buffer and sonicated followed by protein concentration normalization. The equivalent of 100 µg of each sample was precipitated using Methanol-Chloroform. Dried pellets were dissolved in 8 M urea, reduced with 5 mM tris (2-carboxyethyl) phosphine hydrochloride (TCEP), and alkylated with 50 mM chloroacetamide. Proteins were then trypsin digested overnight at 37 °C. The digested peptides were labeled with TMT 6plex (Lot RH239931) and fractionated by basic reverse phase (ThermoFisher, #84868). The TMT labeled samples were analyzed on a Fusion Lumos mass spectrometer (ThermoFisher). Samples were injected directly onto a 25 cm, 100 µm ID column packed with BEH 1.7 µm C18 resin (Waters). Samples were separated at a flow rate of 300 nL/min on a nLC 1200 (ThermoFisher). Buffer A and B were 0.1% formic acid in water and 90% acetonitrile, respectively. A gradient of 1–25% B over 120 min, an increase to 40% B over 80 min, an increase to 100% B over another 30 min and held at 90% B for a final 10 min of washing was used for 240 min total run time. Column was re-equilibrated with 20 µL of buffer A prior to the injection of sample. Peptides were eluted directly from the tip of the column and nanosprayed directly into the mass spectrometer by application of 2.8 kV voltage at the back of the column. The Lumos was operated in a data dependent mode.

Full MS1 scans were collected in the Orbitrap at 120k resolution. The cycle time was set to 3 s, and within this 3 s the most abundant ions per scan were selected for CID MS/MS in the ion trap. MS3 analysis with multinotch isolation (SPS3) was utilized for detection of TMT reporter ions at 60k resolution. Monoisotopic precursor selection was enabled, and dynamic exclusion was used with exclusion duration of 10 seconds.

Protein and peptide identification was done with Integrated Proteomics Pipeline – IP2 (Integrated Proteomics Applications). Tandem mass spectra were extracted from raw files using RawConverter[3] and searched with ProLuCID[4] against human UniProt database appended with MIEF1-MP. The search space included all fully-tryptic and half-tryptic peptide candidates. Carbamidomethylation on cysteine and TMT labels on N-terminus and lysine were considered as static modifications. Data was searched with 50 ppm precursor ion tolerance and 600 ppm fragment ion tolerance. Identified proteins were filtered to using DTASelect[5] and utilizing a target-decoy database search strategy to control the false discovery rate to 1% at the protein level[6].

Mass Spectrometry and Data Analysis. For proteomics analysis, the eluted samples were precipitated with trichloroacetic acid (TCA, MP Biomedicals #196057) overnight at 4 °C or using Methanol- Chloroform. Dried pellets were dissolved in 8 M urea, reduced with 5 mM tris(2-carboxyethyl) phosphine hydrochloride (TCEP, Thermo #20491), and alkylated with 10 mM iodoacetamide (Sigma I1149). Proteins were then trypsin (Promega V5111) digested overnight at 37 °C. The reaction was quenched with formic acid at a final concentration of 5% (v/v). Digested samples were analyzed on a Q Exactive Hybrid Quadrupole-Orbitrap Mass Spectrometer Samples were injected directly onto a 25 cm, 100 µm ID column packed with BEH 1.7 µm C18 resin (Waters). Samples were separated at a flow rate of 300nl/min on an nLC 1000

(Thermo). Buffer A and B were 0.1% formic acid in water and 0.1% formic acid in 90% acetonitrile, respectively. A gradient of 1-30% B over 110 min, an increase to 40% B over 10 min, an increase to 90% B over another 10 min and held at 90% B for 10 min was used for 140 min total run time. Column was re-equilibrated with 20 µl of buffer A prior to the injection of sample. Peptides were eluted directly from the tip of the column and nanosprayed directly into the mass spectrometer by application of 2.5kV voltage at the back of the column. The Q Exactive was operated in a data dependent mode. Full MS1 scans were collected in the Orbitrap at 70K resolution with a mass range of 400 to 1800 m/z. The 10 most abundant ions per cycle were selected for MS/MS and dynamic exclusion was used with exclusion duration of 15 seconds.

Protein and peptide identification were done with Integrated Proteomics Pipeline – IP2 (Integrated Proteomics Applications). Tandem mass spectra were extracted from raw files using RawConverter[3] and searched with ProLuCID[4] against human UnitProt database appended with MIEF1-MP. The search space included all fully-tryptic and half-tryptic peptide candidates. Carbamidomethylation on cysteine was considered as a static modification. Data was searched with 50 ppm precursor ion tolerance and 600 ppm fragment ion tolerance. Identified proteins were filtered to using DTASelect[5] and utilizing a target-decoy database search strategy to control the false discovery rate to 1% at the protein level[6].

MATERIALS

Antibodies. Primary antibodies: mouse anti-FLAG (F1804, Sigma), rabbit anti-TOM20 (sc-11415, Santa Cruz Biotechnology), goat anti-TIMM50 (ab23938, Abcam), rabbit anti-HSP60 (ab46798, Abcam), mouse anti-TOM20 (sc-17764, Santa Cruz Biotechnology), rabbit anti-myc tag (#2272, Cell Signaling Technology), rabbit anti-MRPS27 (#17280-1-AP, Proteintech), rabbit

anti-MRPL4 (HPA051261, Atlas Antibodies), rabbit anti-biotin (D5A7) (#5597S, Cell Signaling Technology).

Secondary antibodies: goat anti-rabbit IRDye® 800CW (926-32211, LI-COR), goat anti-mouse IRDye® 800CW (926-32210, LI-COR), donkey anti-goat IRDye® 800CW (926-32214, LI-COR), Goat anti-mouse Alexa Fluor 488 (A11001, Life Technologies), Goat anti-rabbit Alexa Fluor 546 (A11010, Life Technologies), Goat anti-rabbit Alexa Fluor 488 (A11008, Life Technologies) and Goat anti-mouse Alexa Fluor 546 (A11003, Life Technologies).

Plasmids. pcDNA3-APEX-NES was a gift from Alice Ting (Addgene plasmid #49386). APEX coding sequence was subcloned into pcDNA3.1(+) vector with an N-terminal myc-tag to generate pcDNA3.1-myc-APEX construct. MIEF microprotein coding sequence was obtained by PCR amplification of HEK293T cDNA pool. MIEF-APEX-myc was amplified by overlap extension PCR and subcloned into pcDNA3.1(+) vector. The conserved L-Y-R (leucine/tyrosine/arginine) and downstream F (phenylalanine) residues in the Wild Type MIEF sequence were mutant to A (alanine) using the Q5 Site-Directed Mutagenesis Kit (Biolabs E0552S) making the Mutant MIEF-APEX-myc construct in pcDNA3.1(+) vector.

FIGURES AND TABLES

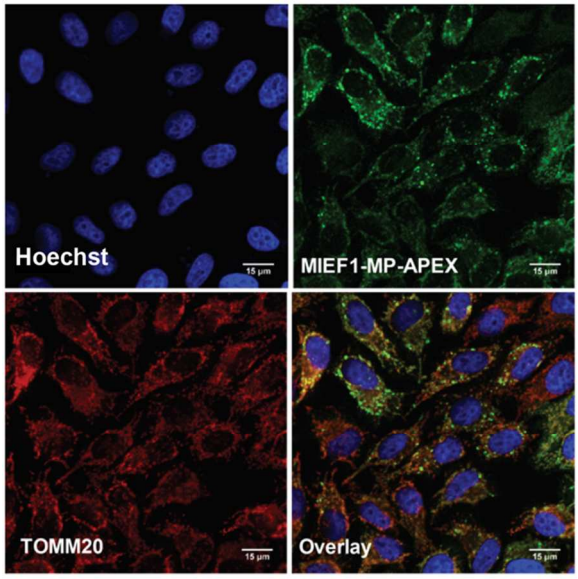

Figure S1. MIEF1-MP-APEX-MYC localizes in mitochondria. MIEF1-MP-APEX-MYC colocalizes with the mitochondrial marker TOM20 in Hela cells verifying its mitochondrial localization (Hoechst (blue) nuclear stain, TOM20 (red), MIEF1-MP-APEX-MYC (green), yellow/orange overlap).

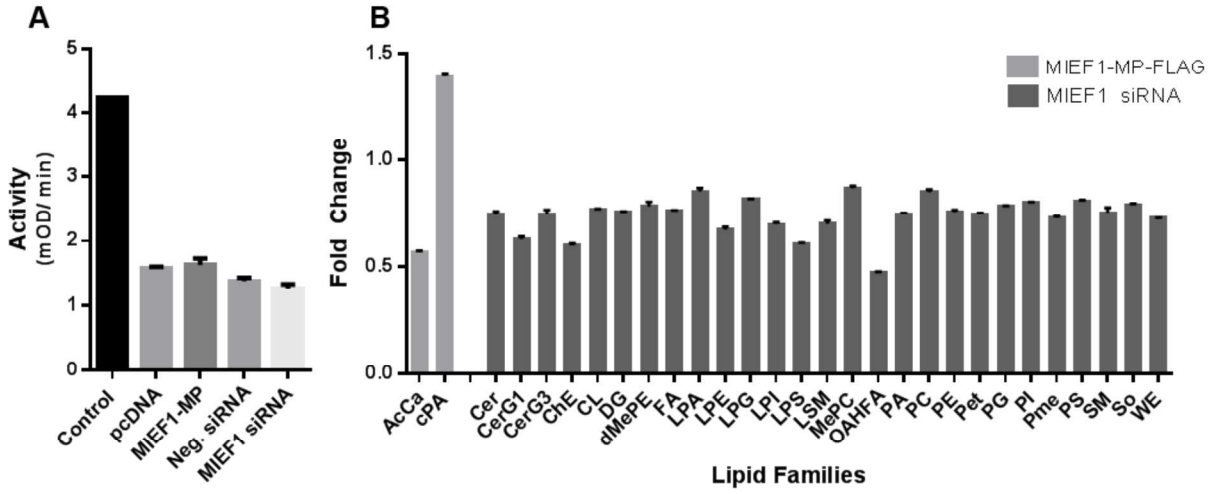

Figure S2. Effect of MIEF1-MP on ACPM/NDUFAB1-related cellular functions. **(A)** Mitochondrial Complex 1 Activity measured (in mOD/min using abcam 109721 assay kit) for HEK293T cells overexpressing or knocking down MIEF1-MP, Bovine heart mitochondrial extract was used as a positive control. **(B)** Fold change in steady-state lipid levels in HEK293T cells after overexpression or knockdown of MIEF1-MP (all indicated lipid families show fold change with significant p-value < 0.05)

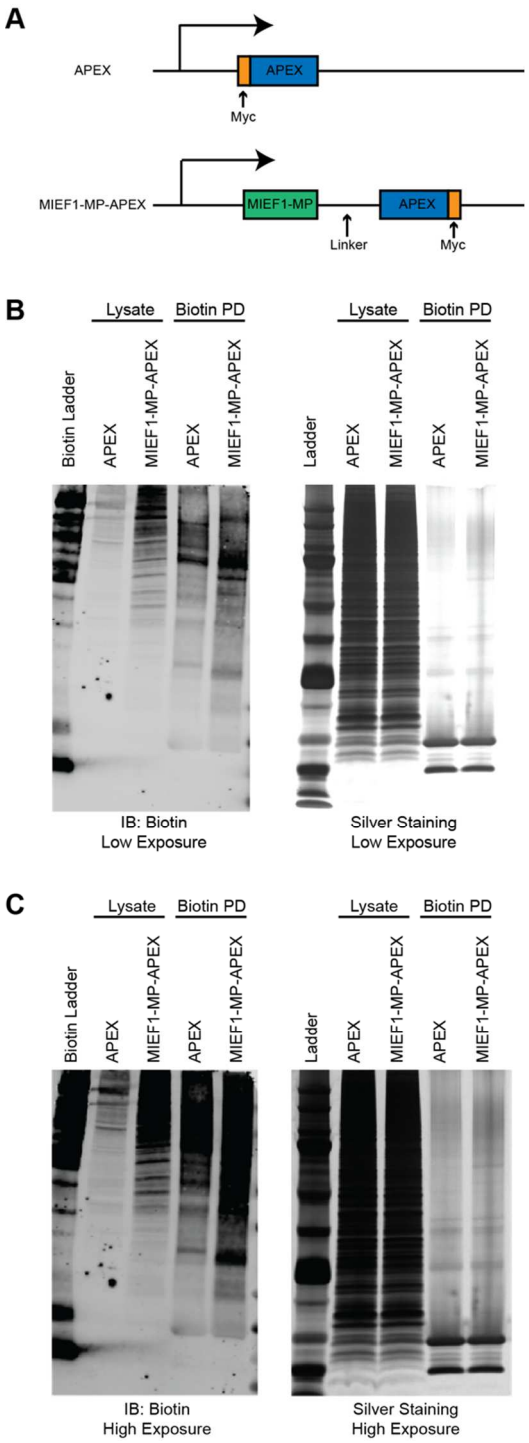

Figure S3. Constructs and controls used for the MIEF1-MP-APEX experiment in Figure 4C. **(A)** Schematic Illustration of APEX Control and MIEF1-MP-APEX constructs. Total lysates and the biotinylated proteomes (enriched by streptavidin beads) of the APEX and MIEF-APEX transfected HEK293T cells were analyzed for biotinylation (by western blotting using anti-biotin

antibody) and loading control (using silver staining) with **(B)** low exposure and **(C)** high exposure.

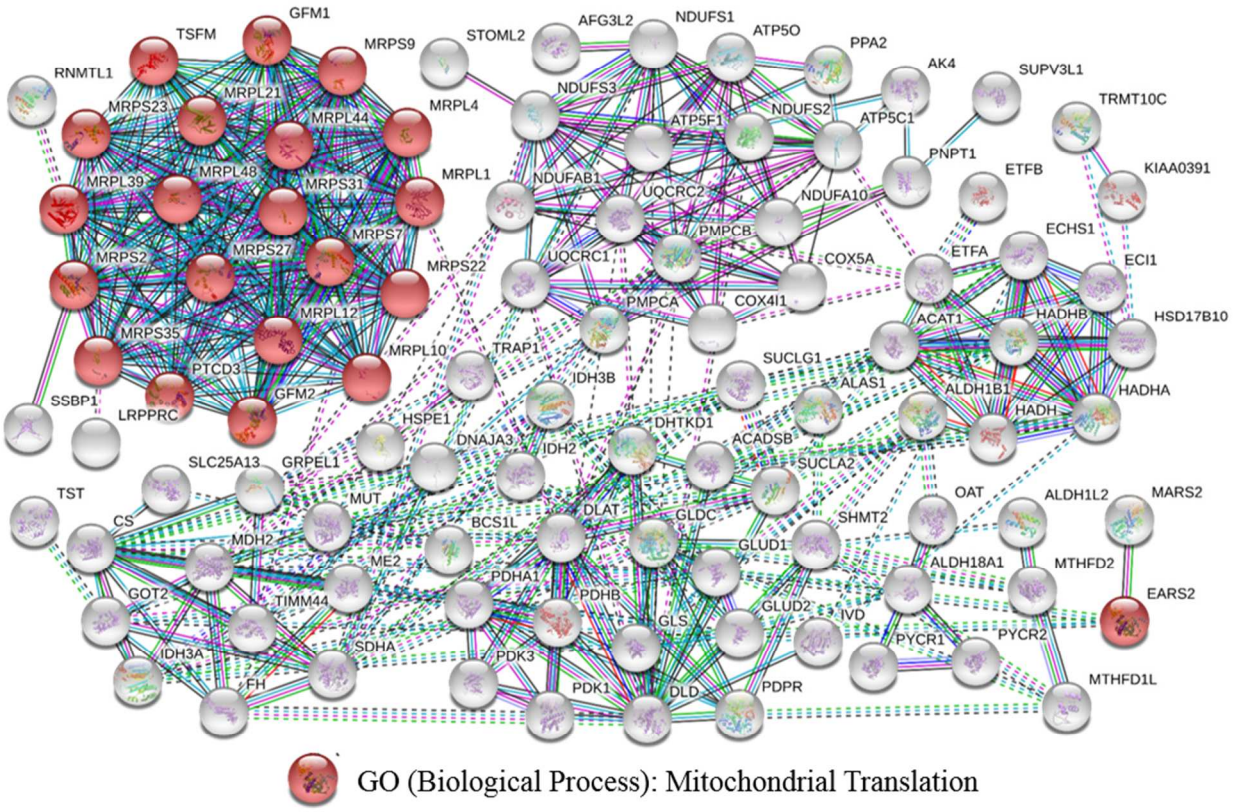

Figure S4. Gene Ontology analysis of the MIEF1-MP STRING interaction network (Figure 4A). The interaction with the mitoribosome suggested a role of MIEF1 microprotein in regulating mitochondrial translation. Proteins highlighted in blue/purple have biological functions related to mitochondrial translation.

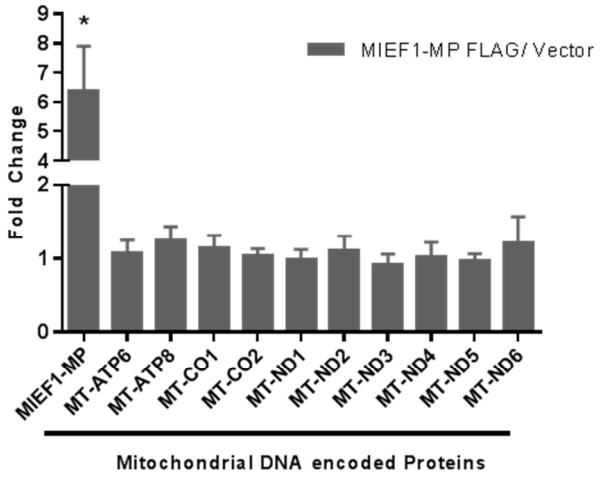

Figure S5. Effect of MIEF1-MP on the mtDNA encoded protein steady state levels. The mtDNA encoded total protein levels were quantified using TMT labeling and proteomic analysis with MIEF1 microprotein overexpression in HEK293T cells. Fold change of protein levels in MIEF1-MP expressing cells compared to the vector control is shown along y axis. * indicates p-value< 0.05

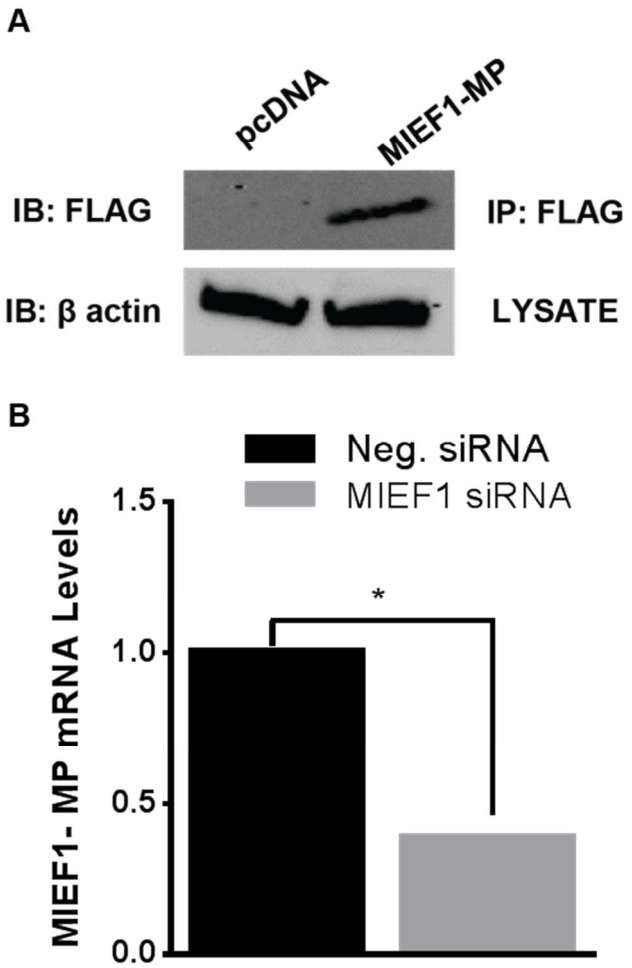

Figure S6. Controls for BONCAT method measuring mitochondrial translation rate in Figure 5. **(A)** Western Blot using Anti-FLAG antibody to confirm FLAG tagged MIEF1 microprotein overexpression in the FLAG Immunoprecipitated eluent normalized with total cell lysate. Anti- β actin antibody was used as a loading control. **(B)** RT-qPCR using primers targeted to the MIEF1 microprotein sequence to confirm the knockdown of the MIEF1 mRNA transcript with siRNA treatment. (*, p-value <0.05 using Student's t-test).

Protein Name	MIEF1- MP Spectral Counts
MRPL4	9
MRPL10	5
MRPL12	3
MRPL16	0
MRPL21	3
MRPL28	0
MRPL34	0
MRPL35	5
MRPL37	0
MRPL38	0
MRPL39	4
MRPL41	4
MRPL42	10
MRPL43	4
MRPL46	0
MRPL47	8
MRPL50	5
MRPS2	0
MRPS5	0
MRPS7	0
MRPS11	0
MRPS15	0
MRPS18A	2
MRPS24	0
MRPS27	0
MRPS30	4
MRPS31	0
MRPS34	0

Table S1. Identification of additional MIEF1 microprotein- mitoribosome protein interactors. The table indicates the number of MIEF1-MP spectral counts identified in the raw Affinity Purification- Mass Spectrometry (AP- MS) data for all the mitoribosome proteins available in the Bioplex database, allowing identification of new mitoribosome interactors highlighted in blue in addition to the mitoribosome interactors identified by MIEF1-MP-APEX PD and proteomic analysis in grey.

REFERENCES

[1] Taguchi, R., and Ishikawa, M. (2010) Precise and global identification of phospholipid molecular species by an Orbitrap mass spectrometer and automated search engine Lipid Search, *Journal of Chromatography A 1217*, 4229-4239.

[2] Wieckowski, M. R., Giorgi, C., Lebiedzinska, M., Duszynski, J., and Pinton, P. (2009) Isolation of mitochondria-associated membranes and mitochondria from animal tissues and cells, *Nature Protocols 4*, 1582.

[3] He, L., Diedrich, J., Chu, Y.-Y., and Yates, J. R. (2015) Extracting Accurate Precursor Information for Tandem Mass Spectra by RawConverter, *Analytical Chemistry 87*, 11361-11367.

[4] Xu, T., Park, S. K., Venable, J. D., Wohlschlegel, J. A., Diedrich, J. K., Cociorva, D., Lu, B., Liao, L., Hewel, J., Han, X., Wong, C. C. L., Fonslow, B., Delahunty, C., Gao, Y., Shah, H., and Yates, J. R. (2015) ProLuCID: an improved SEQUEST-like algorithm with enhanced sensitivity and specificity, *Journal of proteomics 129*, 16-24.

[5] Tabb, D. L., McDonald, W. H., and Yates, J. R. (2002) DTASelect and Contrast: Tools for Assembling and Comparing Protein Identifications from Shotgun Proteomics, *Journal of proteome research 1*, 21-26.

[6] Peng, J., Elias, J. E., Thoreen, C. C., Licklider, L. J., and Gygi, S. P. (2003) Evaluation of Multidimensional Chromatography Coupled with Tandem Mass Spectrometry (LC/LC−MS/MS) for Large-Scale Protein Analysis: The Yeast Proteome, *Journal of Proteome Research 2*, 43-50.