Mitochondrial COA7 is a heme-binding protein involved in the early stages of complex IV assembly.

Luke E. Formosa^{a*}, Shadi Maghool^{b*}, Alice J. Sharpe^a, Boris Reljic^{a,c}, Linden Muellner-Wong^{a,c}, David A. Stroud^c, Michael T. Ryan^{a,†} and Megan J. Maher^{b,d,†}

^a*Department of Biochemistry and Molecular Biology, Biomedicine Discovery Institute, Monash University, Clayton 3800, Australia.*

^b*School of Chemistry and Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Parkville, Victoria 3010, Australia.*

^c*Department of Biochemistry and Pharmacology and The Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Parkville, Victoria 3010, Australia.*

^d*Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne 3086, Australia.*

* *These authors contributed equally to this work.*

†Corresponding authors: michael.ryan@monash.edu and megan.maher@unimelb.edu

Classification

Biological Sciences; Biochemistry

Keywords

COA7, mitochondria, cytochrome *c* oxidase, COX, heme, X-ray crystallography

Author Contributions

LEF, SM, MJM and MTR designed research; LEF, SM, AJS, BR, LM-W, DAS, MJM and MTR performed research; LEF, SM, BR, DAS, MJM and MTR analyzed data; LEF, SM, MJM and MTR wrote the paper.

Competing Interests Statement

The authors declare no competing interests.

Abstract

Cytochrome *c* oxidase assembly factor 7 (COA7) is a metazoan-specific assembly factor, critical for the biogenesis of mitochondrial complex IV (cytochrome *c* oxidase). Although mutations in COA7 have been linked in patients to complex IV assembly defects and neurological conditions such as peripheral neuropathy, ataxia and leukoencephalopathy, the precise role COA7 plays in the biogenesis of complex IV is not known. Here we show that the absence of COA7 leads to arrest of the complex IV assembly pathway at the initial step where the COX1 module is built, which requires incorporation of copper and heme cofactors. In solution, purified COA7 binds heme with micromolar affinity, through axial ligation to the central iron atom by histidine and methionine residues. Surprisingly, the crystal structure of COA7, determined to 2.4 Å resolution, reveals a ‘banana-shaped’ molecule composed of five helix-turn-helix (α/α) repeats, tethered by disulfide bonds, with a structure entirely distinct from proteins with characterized heme binding activities. We therefore propose a role for COA7 in heme binding/chaperoning in the mitochondrial intermembrane space, this activity being crucial for and providing a missing link in complex IV biogenesis.

Significance Statement

Assembly factors play key roles in the biogenesis of many mitochondrial protein complexes regulating their stability, activity and incorporation of essential cofactors. COA7 is a metazoan-specific assembly factor, the absence or mutation of which in humans accompanies complex IV assembly defects and neurological conditions. Here we report the crystal structure of COA7 to 2.4 Å resolution, revealing a ‘banana-shaped’ molecule composed of five helix-turn-helix (α/α) repeats, tethered by disulfide bonds. Characterization of pathogenic variants reveals significantly lower stabilities, correlating with the associated disease outcomes. Fascinatingly, COA7 binds heme with micromolar affinity, despite the fact that the protein structure does not resemble previously characterized heme-binding proteins. This provides a possible missing link for heme handling in the mitochondrial intermembrane space.

Introduction

In eukaryotes, the bulk of ATP generation occurs through oxidative phosphorylation (OXPHOS) that involves five multi-subunit protein complexes in the inner mitochondrial membrane (IMM), known as complexes I to V (1). Complex IV (cytochrome *c* oxidase or COX) is the copper and heme *a* containing terminal oxidase of the mitochondrial respiratory chain, which catalyzes electron transfer from cytochrome *c* in the intermembrane space (IMS) to molecular oxygen and contributes to the generation of the proton gradient required to power ATP synthesis. In mammals, complex IV is composed of 14 subunits, including three core subunits (COX1, COX2 and COX3) that are encoded by mitochondrial DNA (mtDNA) (2-4).

The assembly of complex IV is complicated, but is thought to involve the biogenesis of different modules harboring both mtDNA and nuclear encoded subunits (5, 6). A central module is seeded by membrane integration and maturation of COX1 that then incorporates separate modules containing COX2 and COX3. The incorporation of three copper ions, heme *a* and heme *a*₃ cofactors is necessary for the biogenesis of the COX1 and COX2 subunits and assembly of complex IV. In addition, complex IV requires the additional action of more than 30 assembly factors (6). While not part of the final enzyme, these assembly factors have been shown to perform critical roles in subunit maturation, co-factor incorporation and stabilization of intermediate assemblies of complex IV in humans. Loss-of-function mutations in both mtDNA and nuclear DNA genes encoding mitochondrial proteins, including complex IV subunits and assembly factors, lead to complex IV deficiency and mitochondrial disease. Complex IV is also found in a number of supercomplexes, which are large assemblies containing complex IV along with complexes I and III, including the complex I-complex III₂-complex IV (CI/III₂/IV) respirasome (7-9).

The cytochrome *c* oxidase assembly factor 7 (COA7, also referred to as RESA1) is an IMS localized (10, 11), metazoan-specific complex IV assembly factor. The function of COA7 in complex IV assembly is unknown. COA7 is a cysteine-rich protein (13 cysteine residues) which based on sequence analysis, contains five Sell-like repeat domains. Sell-like repeat (SLR) proteins are a sub-class of the tetratricopeptide repeat (TPR) proteins, which belong to the α/α repeat family of solenoid proteins (12, 13). Available structures of SLR proteins

show large solvent accessible surfaces, suitable for the binding of both large and small substrates (12-14). Accordingly, solenoid proteins often function in DNA-, peptide- and protein-protein interactions (12-14), which facilitate cellular processes such as signaling, the cell-cycle and inter-compartmental transport (13, 15-17).

A number of pathogenic mutations in the gene encoding COA7 have been identified in patients exhibiting neurological symptoms of peripheral neuropathy, ataxia and leukoencephalopathy (18, 19). The first of these studies described an isolated deficiency of complex IV in patient skin fibroblasts and skeletal muscle in the presence of compound heterozygous mutations in *COA7* (p. Tyr137Cys and *COA7*-exon2 Δ), resulting in the expression of the COA7 mutant proteins Y137C and an in-frame deletion mutant, missing residues 37-83. A second study examined a cohort of Japanese patients with recessive mutations in *COA7*, where skin fibroblasts from these patients showed significant decreases in complex I and complex IV activities. Patients had either homozygous single amino acid p.Asp6Gly (D6G) mutations, compound D6G with p.Ser149Ile (S149I) or a missense p.Gly144fs mutation, and compound p.Arg39Trp (R39W) and *COA7*-exon2 Δ mutations (19). The Arg39, Tyr137 and Ser149 residues are highly conserved among COA7 proteins from different species (**Fig. S1**).

In the present study, we used structural and functional approaches to investigate the role of COA7 in mitochondrial complex biogenesis. Loss of COA7 blocks complex IV assembly steps subsequent to formation of the COX1 intermediate module, with COX2 being highly labile. We report the crystal structure of human COA7, to 2.4 Å resolution, which allows us to propose the molecular origins of the pathogeneses observed for the patient mutations. Furthermore, characterization of recombinant COA7 shows that the protein binds heme *in vitro* specifically and with micromolar affinity, through axial His/Met coordination to the iron atom, despite no resemblance between the COA7 structure and well characterized heme-binding proteins. These findings point to a role for COA7 in binding and/or incorporation of heme into complex IV.

Results and Discussion

Loss of COA7 results in reduced levels of respiratory chain enzymes, with a severe complex IV assembly defect

In order to investigate the function of COA7 in the biogenesis of the respiratory chain, CRISPR-Cas9 mediated gene editing (20) was used to generate COA7^{KO} HEK293T cell lines. Following transfection, cells were sorted to generate individual clones. Two clones (termed COA7^{KO}-1 and COA7^{KO}-2) were found to contain frameshift deletions in the *COA7* gene where the gRNA was targeted. Mitochondria were isolated from both control and COA7^{KO} cells, solubilized in digitonin and subjected to blue native polyacrylamide gel electrophoresis (BN-PAGE) and immunoblotted for components of the respiratory chain complexes (**Fig. 1A**). Probing for the complex IV subunit COX4, revealed a severe complex IV defect, with no complex IV holoenzyme nor higher complex IV-containing supercomplexes (CI/III₂/IV or III₂/IV) detected (**Fig. 1A**). Consistent with this, probing for the complex I subunit NDUFA9 revealed the loss of the respirasome (CI/III₂/IV) while a faster migrating species corresponding to the CI/III₂ supercomplex was seen (**Fig. 1A**). Probing for the complex III subunit UQCRC1, revealed the loss of the III₂/IV supercomplex and the accumulation of the CI/III₂ supercomplex (**Fig. 1A**). The levels of complex V (analyzed using antibodies for the subunit ATP5A) were unchanged (**Fig. 1A**). The defects observed in the supercomplexes and complex IV were rescued upon re-expression of COA7 with a C-terminal Flag epitope (+COA7^{Flag}), indicating the complex IV defect observed was indeed due to the loss of COA7 (**Fig. 1A**).

To delineate the effect of COA7^{KO} on assembly of complex IV, we analyzed representative subunits of each of the three assembly modules: COX1 and COX2 and COX3 (detecting COX1, COX2 and COX6A, respectively) (**Fig. 1B**). The analysis showed that residual COX1 was present in a faster migrating complex, termed the 'S2' intermediate (21), while COX2 and COX6A were not present in any observed intermediates (**Fig. 1B**). To further investigate the stability of the core mtDNA-encoded proteins, we utilized [³⁵S]-methionine pulse-chase labeling in the presence of anisomycin, an inhibitor of cytosolic translation (22). Control and COA7^{KO}-1 cells were labeled for 2 hours with [³⁵S]-Met before a chase without label and restoration of translation. Phosphorimage analysis SDS-PAGE showed that COX2 was highly labile in COA7^{KO} cells, with reduced labeling of this subunit at the zero-hour time-point and rapid loss over time, relative to the control cells (**Fig. 1C**). COX3 also displayed increased

turnover during the chase period in COA7^{KO} cells, compared with control mitochondria (**Fig. 1C**). In comparison, translation and stability of COX1 was similar in both lines (**Fig. 1C**). The levels of other mtDNA-encoded proteins did not appear to differ between control and COA7^{KO} mitochondria. These results point to a critical role for COA7 in complex IV assembly in steps that follow the formation of the COX1 module.

We next sought to understand the effect of COA7 loss on the mitoproteome using quantitative stable isotope labeling of amino acids in cell culture (SILAC) proteomics (23). In agreement with the pulse-chase analysis, a significant decrease in the abundance of COX2 and COX3 was observed (**Fig. 1D; Dataset S1**). In addition, the levels of a number of nuclear-encoded complex IV subunits (COX6A1, COX6B1, COX7A2, COX8A and NDUFA4), which are associated with the COX2 or COX3 modules or later complex IV assembly intermediates (4, 24, 25), were reduced in the absence of COA7 (**Fig. 1D; Dataset S1**). The visualization of the relative levels of complex IV subunits on the structure of complex IV using a topographical heatmap (PDB 5Z62) (26) shows that the COX1 module subunits (COX1, COX4 and COX5A) are largely unaffected by the loss of COA7, which is in stark contrast to the other mtDNA-encoded subunits COX2 and COX3, and their surrounding nuclear encoded proteins (**Fig. 1E**). Given that we observe a reduced COX1 signal by BN-PAGE in COA7^{KO} mitochondria compared to control mitochondria, it is possible that COX1-containing assemblies subsequent to the S2 subcomplex are prone to aggregation once extracted from the membrane environment. This suggests that the COX1 module is unable to mature and integrate with other COX modules, resulting in turnover of the COX2 and COX3 modules. Taken together, these results indicate that COA7 plays a critical role in the assembly of complex IV at a step that involves association of the COX2 and COX3 modules onto the COX1 module.

We performed affinity-enrichment mass spectrometry to identify stable COA7 interacting proteins in an unbiased manner. Mitochondria isolated from control and COA7^{KO} + COA7^{Flag} rescue cells were solubilized in digitonin, incubated with anti-Flag beads and the bound proteins, eluted with Flag peptide, were subjected to label free quantitative (LFQ) proteomics analysis. COA7^{Flag} failed to enrich any complex IV structural proteins or assembly factors (**Fig. 1F; Dataset S2**). COA7^{Flag} however, did enrich a number of IMS proteins, including apoptosis-inducing factor 1 (AIFM1) and CHCHD4, which are required for disulfide bond formation in the intermembrane space (27). In addition, COA7 was able to enrich

coproporphyrinogen oxidase (CPOX), an enzyme involved in heme biosynthesis (28), the mitochondrial protease LONP1 (29) and the disaggregase CLPB and the CLPB-interacting protein HAX1 (30) (**Fig. 1F; Dataset S2**). RCN2 an ER-luminal protein with potential localization in the mitochondrial IMS (31), was also identified in association with COA7^{Flag}. These results suggest that unlike many assembly factors, COA7 does not form stable interactions with components of the respiratory chain.

Figure 1. Loss of COA7 results in defects of respiratory chain enzymes. (A-B) Mitochondria isolated from control, COA7 knockout and rescued COA7 knockout cells (as indicated) were solubilized in 1% digitonin and subjected to BN-PAGE and analyzed using the antibodies shown. Coomassie brilliant blue (CBB) stained membranes are indicative of loading. (C) Cells were pulsed with [³⁵S]-Met for 2 hours and chased for the indicated times. Isolated mitochondria were analyzed by SDS-PAGE and autoradiography. SDHA was used as a loading control. (D) Volcano plot showing proteins with altered abundance in COA7^{KO} mitochondria. Subunits of the different OXPHOS complexes are color coded. Complex IV subunits are labeled. The horizontal line indicates $p = 0.05$, vertical lines indicate a fold change of 1.5. (E) Topographical heatmap of SILAC log₂(ratios) mapped onto the structure of complex IV (PDB 5Z62) (26). Proteins encoded by mtDNA are shown as surface renders, while nuclear encoded proteins shown as cartoon. Blue intensity is relative to the decrease in protein abundance, n.d. not detected. (F) Digitonin-solubilized mitochondria from control or COA7^{KO} + COA7^{Flag} cells were subjected to affinity enrichment using Flag agarose beads followed by label free quantitative proteomics. Significance was determined through a two-sided, two sample t-test using permutation-based FDR statistics. The curve represents the cutoff for FDR < 5% using an s0 parameter of 20.

Structural characterization of COA7

To uncover further functional information, we sought to determine the molecular structure of COA7. Recombinant glutathione S-transferase (GST) fusions with COA7, pathogenic mutant proteins ^{R39W}COA7, ^{Y137C}COA7, and ^{S149I}COA7 and additional variants (^{Y137S}COA7,

H119A COA7 and M156A COA7) were overexpressed and purified by affinity and size exclusion chromatography (SEC), yielding preparations of the GST-free proteins of approximate 90-95% purity as determined by SDS-PAGE, with intense bands at ~25 kDa, corresponding to the molecular weight of monomeric COA7 (**Fig. S2**). The analytical SEC elution profile of the resulting purified COA7 showed a broad elution peak with an elution volume corresponding to a molecular weight of ~55 kDa, indicating the possible presence of a dimeric species in solution (**Fig. S2**). The quaternary structure of COA7 was further investigated by analytical ultracentrifugation (AUC), using sedimentation velocity (SV) experiments. These data were fitted to a continuous size distribution model, which yielded sedimentation coefficients of 1.9 S and 3.5 S, demonstrating that the sample contained a mixture of monomeric and dimeric species (**Fig. S2, Table S2**), with the relative proportions of these species observed to vary with protein concentration, indicating concentration-dependent oligomerization behavior.

We determined the structure of COA7 by X-ray crystallography. The structure was solved and refined in space group $I4_1$ to 2.4 Å resolution, with a single molecule of COA7 per asymmetric unit (**Fig. 2A**). The final COA7 model includes residues Glu10-His218 and refinement of the model converged with residuals $R = 20.5\%$ and $R_{\text{free}} = 25.5\%$ (**Table S3**). COA7 is comprised of 11 α -helices, arranged as five helix-turn-helix (α/α) repeats (which vary in length from 30-36 residues) and an additional C-terminal helix (**Fig. 2A**). Together, these form an elongated, right-handed super-helix, with a concave and a convex face. The COA7 sequence includes 13 cysteine residues (at positions 24, 28, 37, 62, 71, 95, 100, 111, 142, 150, 172, 179 and 187), with the structure showing five intramolecular disulfide bonds between residues Cys28-Cys37, Cys62-Cys71, Cys100-Cys111, Cys142-Cys150 and Cys179-Cys185 (**Fig. 2B**), which bridge pairs of helices within each of the five α/α repeats. The ten Cys residues that participate in the intramolecular disulfide bonds are conserved in COA7 sequences across all metazoans however, the cysteine residues at positions 24, 95 and 172 (which are not observed to participate in disulfide bonds in this structure) are found in COA7 sequences from vertebrates only (**Fig. S1**) (10).

A search against the PDB (32) revealed that COA7 shares significant structural similarity with the *Helicobacter* cysteine-rich proteins B and C (HcpB, (PDB 1KLX) and HcpC (PDB 1OUV), respectively) (33, 34) from *Helicobacter pylori* (**Fig. S3**). Hcp proteins are

composed of α/α repeats and belong to the SLR protein family. The presence of disulfide bonds within the repeats for both the Hcps and COA7 is unusual for this protein family (13, 33). Indeed, to our knowledge, COA7 is the first structurally characterized human SLR protein possessing disulfide-bridged α/α repeats.

Figure 2. Structure of COA7. (A) Cartoon representation of the overall structure of COA7. Secondary structures are represented as cartoons with SEL1 repeats colored in gray, raspberry, cyan, teal and green. The transparent molecular surface is represented in gray (B) Cysteine residues (labeled) involved in intramolecular disulfide bonds are shown as yellow spheres, and free cysteine residues are shown as salmon spheres. (C) The molecular surface of COA7 is colored according to the electrostatic potentials (red, negatively charged; blue, positively charged; white, uncharged). The top images show the side view of the COA7 structure. The left and right bottom images show concave and convex views, respectively.

The outer convex surface of the COA7 molecule is highly negatively charged due to the presence of a number of Glu and Asp residues (located within repeats 1, 2 and the linkers between repeats 3, 4 and 4, 5 and the C-terminal helix). In addition, the presence of lysine residues 49, 56, 59, 73, 82, 105 and 106 on the inner concave surface of the COA7 molecule, creates a positively charged surface in this region (**Fig. 2C**). The residues that comprise both of these charged surfaces are largely conserved in COA7 proteins (**Fig. S1**). Also on the concave surface is a hydrophobic cleft that results from clustering of several hydrophobic residues (Ile108, Ala109, Gly144, Gly145, Leu181 and Gly182) (**Fig. 2C**). Interestingly, the application of crystallographic symmetry operators reveals that in the crystal, this cleft is occupied by the N-terminal helix (residues Glu10-His30) of a neighboring molecule of COA7 (**Fig. S4**), such that the crystal is composed of an infinite network of these protein-protein interactions (**Fig. S4**). This interaction is mediated by hydrogen bonds and electrostatic interactions, with a total buried surface area of 1614 Å², representing ~7% of the surface areas of the two molecules combined (32). Given the limited extent of this interaction, it is unlikely to represent a persistent oligomeric structure in solution (35), but correlates with the concentration-dependent oligomerization behavior observed by AUC. However, this contact may indicate a molecular mechanism for an interaction between COA7 and protein partners or ligands in the mitochondrial IMS. Similar direct protein-protein interactions have been previously proposed for other TPR proteins, including HcpC (34), the Hsp70/Hsp90 organizing protein (Hop) (PDB 1ELR) (36, 37) and the peroxisomal targeting signal-1 PEX5 (PDB 1FCH) (38). In the case of the related HcpC structure, an interaction between the C-terminus of one molecule and the hydrophobic pocket of a neighboring molecule facilitates the formation of a similar, extended protein network in the crystal. This was proposed to represent a mechanism for protein-protein interactions in the cell (34).

COA7 pathogenic mutations

Pathogenic mutations in the *COA7* gene producing the mutant variants ^{D6G}COA7, ^{R39W}COA7, ^{Y137C}COA7 and ^{S149I}COA7 have been identified in patients exhibiting neurological symptoms of peripheral neuropathy, ataxia and leukoencephalopathy (18, 19). Analysis of the crystal structure of COA7 reveals that residue Arg39 is located on the C-terminal helix of repeat 1 and is exposed on the outer surface of the COA7 molecule (**Fig. 3A**). Residues Tyr137 and Ser149 are located in the ‘middle’ of the structure, on repeat 4, and participate in hydrogen bonding interactions with residues Glu101 and His112, respectively which are both found on repeat 3 (**Fig. 3A**). Residue Asp6 is at the N-terminus of the molecule and could not be resolved in the crystal structure.

Figure 3. COA7 pathogenic mutations. (A) The structure of COA7. Patient mutations are shown as pink sticks and hydrogen bonds between residues (labeled) shown as dashed lines. Inset: closer view of the interactions formed by the His112 and Y137 residues. (B) Mitochondria isolated from control, COA7^{KO} cells and COA7^{KO} cells expressing COA7^{Flag} and variants as indicated were subjected to SDS-PAGE and immunoblotting with antibodies as shown. ATP5A serves as a loading control. (C) Melting temperatures of the COA7 and variant proteins. These values were determined by differential scanning fluorimetry from purified protein samples.

In order to investigate the effects of the pathogenic mutations identified in patients, we generated doxycycline-inducible retroviral plasmids encoding COA7 mutant proteins for stable expression in COA7^{KO} cells. Expression of COA7 with a Flag epitope tag was able to rescue COX2 levels (although not to control levels) in the COA7^{KO} line (**Fig. 3B**). However,

upon expression of the patient derived COA7 variants (^{D6G}COA7, ^{R39W}COA7, ^{Y137C}COA7 and ^{S149I}COA7), the levels of COX2 were not able to be restored to the same extent, indicating that these variants were less active in complex IV assembly, relative to COA7 (**Fig. 3B**). Expression of the ^{Y137S}COA7 protein was also tested to determine if an amino acid substitution of similar geometry to the Y137C patient-identified mutation, but unable to form disulfide bridges, was functional. The Y137S mutant was not able to rescue the defect and showed a similar impact as the Y137C mutation (**Fig. 3B**).

Recombinant, purified forms of ^{R39W}COA7, ^{Y137C}COA7, ^{S149I}COA7 and ^{Y137S}COA7 were prepared and subjected to differential scanning fluorimetry (DSF, **Fig. 3C**) to measure their thermal stabilities (39). Reduced protein stabilities, relative to COA7 were observed for all variant proteins, consistent with the observed decreased expression in mitochondria (**Fig. 3B**). In particular, the melting temperatures of the ^{Y137C}COA7 and ^{Y137S}COA7 proteins were 16-19°C lower than COA7. The thermal transition for purified ^{S149I}COA7 could not be observed (**Fig. 3C**), indicating the protein was most likely misfolded. Analysis of the quaternary structure of ^{Y137C}COA7 by analytical ultracentrifugation showed evidence of concentration dependent dimerization (as observed for COA7), in addition to the presence of higher order oligomeric states indicative of aggregation. These observations may explain the reported aggregation and mislocalization of the ^{Y137C}COA7 protein in cells (10) (**Fig. S5, Table S2**).

A recent analysis of the COA7 and ^{Y137C}COA7 proteins *via* molecular modeling and simulation studies, proposed that the presence of residue Cys137, may lead to the disruption of the Cys100-Cys111 disulfide bond and the dynamic formation of alternative disulfide bonds between the introduced Cys137 and residues Cys100 and Cys111 (10). The structure reported here shows that the formation of an alternative Cys137-Cys111 disulfide bond would be possible without major structural rearrangement. However, the distance between the thiolate groups of the introduced Cys137 and residue Cys100 would be too long (3.9 Å) in the absence of a major structural reorganization (**Fig. S6**). Indeed, the Y137S mutation, which cannot form an alternative disulfide bond, resembled the Y137C mutation with respect to its melting temperature (**Fig. 3C**) and inability to rescue the COA7^{KO} phenotype (**Fig. 3B**). These findings lead us to conclude that a rearrangement of the disulfide bonding network of COA7 upon mutation is unlikely to be a factor in the pathogenic effects of this variant form of the protein.

The COA7 S149I pathogenic mutation is predicted to eliminate the hydrogen bond with residue His112 and the larger Ile sidechain may also disrupt the helical packing between repeats 3 and 4 (**Fig. 3A**). Given that ^{S149I}COA7 was likely misfolded and that both the Y137C and S149I mutations are located in the ‘centre’ of the crescent-shaped structure (**Fig. 3A**), this suggests that hydrogen bonds and other intramolecular interactions in this region may provide rigidity within the structure, most likely scaffolding one half of the assembly relative to the other. Loss of these interactions lead to instability of the proteins and decreased activities in terms of complex IV biogenesis. Finally, in the case of R39W variant, residue Arg39 is exposed on the outer surface of the COA7 molecule. The introduction of the R39W mutation is predicted to alter the charge in this region of the protein surface. The DSF analysis of the ^{R39W}COA7 variant protein showed it to be well folded (**Fig. 4C**). The pathogenesis of this mutation may therefore lie in altered interactions of this protein with as yet unidentified partners/substrates, associated with the change in the electrostatic structure of its surface.

COA7 binds heme with micromolar affinity

A key step in the biogenesis of complex IV is insertion of copper into COX2 from the IMS, which involves a suite of assembly factors (6). Loss of function mutations in these assembly factors blocks complex IV assembly at the level of COX2 maturation, similar to that observed here for COA7 (21). Furthermore, soluble high affinity Cu(I)-binding proteins (copper metallochaperones) such as ATOX1 (40), COX17 (41) and the recently identified Cu(I)-binding protein, COA6 (21) all show CysX_nCys sequence motifs, bind Cu(I) and function in copper delivery to protein partners within the cell. Due to the presence of multiple Cys residues within the COA7 sequence, we sought to determine whether COA7 binds copper (specifically Cu(I)) using a competition assay with the [Cu^IBcs₂]³⁻ complex (21, 42). Cu(I)-binding to the COA7 was not detected using this assay, however Cu(I) binding was observed for the recombinant ATOX1 protein, which was included as a positive control (**Fig. S7**).

Given these findings, along with our inability to detect an association between COA7 and other copper chaperones, we investigated whether COA7 might bind heme, the other cofactor required for complex IV biogenesis and activity. Complex IV binds heme *a* and heme *a*₃, which are essential for function. Heme *a* is important for the stability of COX1, while

insertion of heme a_3 , which has a farnesyl tail and sits between COX1 and COX2, is required for assembly and stabilization of COX2 (5, 6). We therefore tested whether COA7 might be involved in heme incorporation into complex IV by measuring whether the purified protein was able to bind heme in solution using ultra violet-visible (UV-Vis) spectrophotometry (43, 44). Upon titration of COA7 with hemin, difference spectra revealed a Soret band at 416 nm and a broad feature between 500-600 nm (**Fig. 4A**), which increased in intensity with the addition of increasing concentrations of hemin. Fitting of these titration data yielded $K_{D(\text{heme})} = 1.40 \pm 0.28 \mu\text{M}$ (**Fig. 4B**). This value is similar to those measured for other heme binding proteins. For instance, the HemQ (chlorite dismutase-like protein) from *Listeria monocytogenes*, which is involved in heme acquisition by the bacterium binds heme with $K_{D(\text{heme})} = 16.2 \mu\text{M}$ (45). The cytoplasmic heme binding protein Phus from *Pseudomonas aeruginosa* binds heme with $K_{D(\text{heme})} = 29 \mu\text{M}$ (46).

Figure 4. Heme binding to COA7. (A) Difference absorption spectra of COA7 (1 μM) with increasing concentrations of Fe³⁺-PPIX (from 0 to 10 μM) as indicated. (B) The binding curve was fitted to the data from (A) using an equation describing a single binding site ($Y = B_{\text{max}} * X / (K_D + X)$) with GraphPad Prism. (C) UV-Vis absorption spectra of COA7 in the presence of one molar equivalent of Fe³⁺-PPIX (black); Fe²⁺-PPIX (red), and Zn²⁺-PPIX (dotted, black).

Incubation of purified COA7 with one molar equivalent of hemin and reduction with sodium dithionite, yielded UV-Vis spectra indicative of heme-bound COA7 (**Fig. 4C**). These spectra were very similar to the those reported for the ferric and ferrous states, respectively of cytochrome *b*₅₆₂ from *E. coli* (47), which shows axial ligation of the heme iron by His and Met residues. Incubation of COA7 with Zn²⁺-PPIX also yielded a spectrum very similar to that of Zn²⁺-PPIX-loaded cytochrome *b*₅₆₂ (**Fig. 4C**) (47). The incubation of COA7 with excess heme, followed by removal of unbound cofactor, yielded a COA7-heme complex with a stoichiometry, as determined by inductively-coupled plasma mass spectrometry (ICP-MS), of 1.06 ± 0.03 iron atoms (heme cofactors) per protein molecule (**Fig. S8**). COA7 could not be co-purified with PPIX using the same approach (**Fig. S8**).

These experiments indicate specific heme binding to COA7, with spectral evidence for axial ligation to the iron through His and Met ligands. The only TPR protein characterized to bind heme is the HusA protein from *Porphyromonas gingivalis*, with a structure of four α/α repeats, which in contrast to COA7, exhibited iron-independent heme binding (that is, no direct coordination from the protein to the heme iron) (48). Aside from the presence of the TPRs, the HusA structure bears little similarity to that of COA7 in that it shows a relatively ‘closed’ structure compared with COA7 and does not include disulfide bonds. Examination of the structure of COA7 reveals a possible heme binding site located between repeats 3 and 4, with residues His119 and Met156 in comparable positions to those that act as iron ligands in cytochrome *b*₅₆₂ (47) (**Fig. S9A**). We therefore generated single and double mutant proteins (^{H119A}COA7, ^{M156A}COA7 and ^{H119A/M156A}COA7) to probe the contributions of these residues to heme binding.

Titration of the ^{H119A/M156A}COA7 variant with hemin, yielded no detectable binding (**Fig. S10A-C**) and the UV-Vis spectra of the variant protein following incubation with one equivalent of hemin were significantly different from that observed for COA7 (both in the presence and absence of sodium dithionite; **Fig. S11**). Finally, we incubated the heme-loaded COA7, ^{H119A}COA7 and ^{M156A}COA7 proteins with sodium cyanide and compared the UV-Vis spectra of the proteins measured before and after incubation (**Fig. S10D-F**). The spectrum of the heme-COA7 protein showed no change on incubation with CN⁻. In contrast, for both the heme-^{H119A}COA7 and heme-^{M156A}COA7 variants, shifts in the positions of the Soret bands

and spectral changes in the region between 550-650 nm were observed, indicating a change in iron coordination (therefore ligand substitution) on addition of CN⁻. These data taken together, indicate that heme binding to COA7 is mediated by axial ligation of residues His119 and Met156 to the heme iron. A superposition of the COA7 structure with that of cytochrome *b*₅₆₂ (PDB 256B; **Fig. S9B,C**) (49), shows that heme binding in this position, comparable to that seen in the structure of cytochrome *b*₅₆₂ (49), would be possible upon slight adjustments to the positioning of the surrounding helices. These data indicate that *in vitro* COA7 binds heme specifically through axial ligation of the central iron atom with His119 and Met156 protein ligands, which are conserved in the majority of COA7 sequences (**Fig. S1**).

Conclusion

Despite the importance of the assembly of complex IV in health and mitochondrial disease, we have an incomplete understanding of the molecular basis of the biogenesis of this enzyme owing to a lack of knowledge about the identities, structures and roles of all crucial assembly factors (6). The present study shows that deletion of COA7 resulted in a severe complex IV assembly defect. Crucially, the absence of COA7 causes specific defects in the maturation of the COX2 and COX3 modules, resulting in their degradation and therefore halting complex IV biogenesis. This suggests that the COX1 module may be unable to fully mature and integrate with the COX2 and COX3 modules.

The high-resolution crystal structure of the human COA7 protein shows the protein is composed of five SLRs, tethered by disulfide bonds. The overall fold of COA7 is similar to other TPRs and SLRs, proposed to function in protein-protein interactions due to their extensive solvent accessible surfaces (13, 15-17). Fascinatingly, recombinant COA7 was found to bind heme in solution *via* axial His/Met coordination to the heme iron, despite showing no structural similarity to well-characterized heme binding folds. A role for COA7 in chaperoning heme in the IMS for insertion into complex IV would be consistent with the observed ‘blockage’ of complex IV assembly at the COX1 module in the absence of the assembly factor.

Heme is a highly reactive and inherently toxic compound. As a cofactor, heme *b* is required for the activities of a plethora of hemeoproteins, present in virtually every subcellular compartment throughout the eukaryotic cell. It is also a precursor for the synthesis of other

heme types important for eukaryotic physiology, including hemes *a*, *c* and *o*. All of these processes require that heme is trafficked from its site of synthesis in the mitochondrial matrix to heme-dependent proteins throughout the cell (28). For complex IV biogenesis, heme *b* is converted to heme *a* through the sequential activities of the enzymes COX10 (which catalyzes the synthesis of heme *o*) and COX15, which have active sites facing the matrix and the IMS, respectively (28, 50). To date, no information is available on how heme *o* is transported across the IMM to the COX15 active site. Indeed, the aqueous insolubility and instability of hemes in general preclude the existence of ‘free’ (uncomplexed) forms of this cofactor, although low concentrations (<2.5 - 40 nM) of labile heme pools exist in different cellular compartments (51). Given the data presented here, it is tantalizing to suggest then that COA7 provides a chaperoning role for heme *o* and/or heme *a* in the IMS. If so, this would certainly provide a missing link for heme handling in mitochondria.

Online Methods

Cell culture and genome editing

Human epithelial kidney cells containing the SV40 T-antigen (HEK293T) were purchased from the ATCC and a previously established clonal line was used (52, 53). Cells were cultured in DMEM supplemented with 10% (v/v) fetal bovine serum (FBS), 1 × penicillin/streptomycin (Sigma-Aldrich; P4458), 1 × Glutamax (Life Technologies; 35050061) and 50 µg/mL uridine (Sigma-Aldrich; U3750). Cells were grown at 37°C with a humidified 5% CO₂ atmosphere. Galactose containing DMEM was prepared from glucose-free DMEM (Life Technologies; 11966-025) supplemented with 10% (v/v) dialysed FBS (GE Healthcare; SH30079.03), 25mM D-galactose (Sigma-Aldrich; G0750), 1 × penicillin/streptomycin, 1 × Glutamax, 1 × sodium pyruvate and 50 µg/mL uridine.

For SILAC culture, cells were grown as previously described (52). Briefly, cells were grown in DMEM without lysine/arginine, supplemented with 10% (v/v) dialysed FBS, 1 × penicillin/streptomycin, 1 × Glutamax, 1 × sodium pyruvate and 50 µg/mL uridine and either ‘heavy’ amino acids (¹³C₆¹⁵N₄-Arginine; Cambridge Isotope Labs Inc; CNLM-539-H-1, ¹³C₆¹⁵N₂-Lysine; Silantes; 211604102) or ‘light’ amino acids (arginine; Sigma-Aldrich; A5131, lysine; Sigma-Aldrich; L5626).

Gene editing was performed using the pSp-CAS9(BB)-2A-GFP (PX458) CRISPR/Cas9 construct (a gift from F. Zhang; Addgene plasmid 48138) (20). The guide RNA was designed to target the 3’ end of the first exon with the sequence 5’-CTGCTACCACGAGAAGGACC-3’. Indel sequencing of the genomic locus where the CRISPR/Cas9 was targeted demonstrated that COA7^{KO}-1 contained the following mutations c.[93_99del];[101_102insC] leading to p.[E31Dfs*25; D35Gfs*17]. Analysis of COA7^{KO}-2 demonstrated this cell line contained the mutations c.[100_101insTAATATCTTTGTGTTTACAGTCAAATTAATTC AATTATCTCTCTAACAGCCTTGTATCGTATATGCAAATATGAAGGAATCATGGGAAATAGGCCCTCACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGC; 101_102ins AATATCTTTGTGTTTACAGTCAAATTAATTC AATTATCTCTCTAACAGCCTTGTATCGTATATGCAAATATGAAGGAATCATGGGAA

ATAGGCCCTCACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAA
GGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAA
AATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAG
GCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTAC
CGGATACCTGTCCGCC] leading to p.[P34Lfs*9;D35Ifs*8].

Transfection and stable cell line generation

Cells were transfected using Lipofectamine LTX (Life Technologies; 15338100) according to manufacturer's instructions. Stable, inducible expression was achieved by transfecting cells with pLVX-TetOne-Puro (Clontech; 631849) encoding the wildtype COA7 cDNA. For mutational analysis in COA7^{KO} cells, wildtype or mutant COA7^{Flag} was cloned into pRetroX-TetOne-Puro (Clontech; 634309), downstream of a GFP-T2A sequence to facilitate sorting of COA7-expressing cells (54). The COA7 cDNA was amplified from a HEK293T cDNA library derived from the clonal HEK293T cells. Mutant COA7 cDNA was generated using mutagenic primers to generate PCR fragments that incorporate the desired mutation, which were then subsequently assembled using Gibson Assembly (NEB). All plasmids were sequence verified by Micromon (Monash University).

Lentiviral supernatants were prepared by co-transfecting HEK293T cells with pLVX-TetOne-Puro plasmid encoding the cDNA of interest together with pVSV-G and pSPAX2 helper vectors. Retroviral supernatants were prepared by co-transfecting HEK293T cells with pRetroX-TetOne-Puro plasmid encoding the cDNA of interest together with pVSV-G and Gag-Pol helper vectors. Supernatants were collected after 48h and filtered through a 0.45µm low protein binding filter (Merck). COA7^{KO}-1 cells were then infected with the lenti/retroviral supernatant with the addition of 8 µg/mL polybrene for stable and inducible expression of COA7^{Flag}. Transduced cells were selected by the addition of 2 µg/mL puromycin until control cells died. Cells transduced with pRetroX-TetOne-Puro COA7^{Flag} variants were selected by fluorescence-activated cell sorting (FACS), after expression of GFP.

mtDNA subunit labeling

Radiolabeling of mtDNA-encoded proteins was performed using [³⁵S]-Methionine/Cysteine (Perkin Elmer) in the presence of 10 µg/mL Anisomycin (Sigma-Aldrich; A9789) to block cytosolic translation as previously described (22). Labeling was performed for 2 hours,

followed by a chase with regular DMEM for 0, 3 and 24 hours. Isolated mitochondria were then analyzed by Tris-Tricine SDS-PAGE and phosphorimaging digital radiography (GE Healthcare).

Mitochondrial isolation, whole cell lysate preparation, gel electrophoresis and western blotting

Mitochondria were isolated by differential centrifugation as previously described (53, 55). Protein concentration was determined by bicinchoninic acid assay (BCA; Thermo Fisher Scientific; 23223). Mitochondria were used immediately or aliquoted and frozen at -80°C until required.

Tris-tricine SDS-PAGE and BN-PAGE were performed as previously described (56-58). For tris-tricine SDS-PAGE, mitochondria were solubilized in $1\times$ LDS Sample Buffer (Thermo Fisher Scientific; NP0008) supplemented with 100 mM DTT and run on 10% acrylamide gels. Radiolabeling samples were analyzed using continuous 10-16% Tris-tricine SDS-PAGE gels. For BN-PAGE, mitochondria were solubilized in detergent buffer (20 mM Bis-Tris pH 7.0, 50 mM NaCl, 10% (w/v) glycerol) containing 1% digitonin for 10 minutes on ice, followed by centrifugation at 16,000 g for 10 minutes. The clarified supernatant containing mitochondrial complexes was then transferred to a clean tube containing one-tenth of the volume of 10x BN-PAGE loading dye (5% (w/v) Coomassie blue G250, 500 mM ϵ -amino n-caproic acid, 100 mM Bis-Tris pH 7.0) before loading onto the gel. Following electrophoresis, transfer onto PVDF membrane (Merck; IPVH00010) was performed using a Novex Semi-Dry Blotter (Thermo Fisher) or an Invitrogen Power Blotter System (ThermoFisher) according to manufacturer's instructions.

Primary antibodies used in this study were α -COA7 (25361-1-AP; Proteintech), α -NDUFA9 (in-house), α -SDHA (ab14715; Abcam), α -UQCRC1 (459140; ThermoFisher), α -COX1 (459600; ThermoFisher), α -COX2 (A-6404; ThermoFisher), α -COX4 (ab110261; Abcam), α -COX6A (11460-1-AP; Proteintech) and α -ATP5A (ab14748; Abcam). Anti-mouse (Sigma-Aldrich; A9044) or anti-rabbit (Sigma-Aldrich; A0545) horseradish peroxidase-conjugated secondary antibodies were used at a dilution of 1:10,000. Clarity western ECL chemiluminescent substrate (BioRad; 1705061) was used for detection on a BioRad ChemiDoc XRS+ imaging system according to the manufacturer's instructions.

SILAC proteomics

SILAC proteomics was performed as described previously (52). Briefly, equal amounts of differently labeled ‘heavy’ and ‘light’ cells were mixed and mitochondria were subsequently isolated. Analysis consisted of two light control/heavy COA7^{KO} pairs and a label switched heavy control/light COA7^{KO} sample. For sample preparation, 100 µg of isolated mitochondria were solubilized in 1% (w/v) SDC, 100 mM Tris-Cl pH 8.1, 40 mM chloroacetamide and 10 mM TCEP prior to vortexing and heating for 5 minutes at 99°C with 1500 rpm shaking. Samples were then sonicated for 15 minutes in a room temperature water bath prior to the addition of 1 µg trypsin (Promega; V5113) and incubation overnight at 37°C. The supernatant was then transferred to 3x 14G 3MTMEmporeTM SDB-RPS stage tips (59). Ethyl acetate (99%) and 1% TFA was added to the tip before centrifugation at 3000 g at room temperature as described (59). Stage tips were washed first with 99% ethyl acetate and 1% TFA and then with ethyl acetate supplemented with 0.2% TFA. Samples were eluted in 80% acetonitrile (ACN) and 1% NH₄OH and acidified to a final concentration of 1% TFA prior to drying in a SpeedVac.

Affinity enrichment mass spectrometry

Affinity enrichment was performed as previously described (60). Mitochondria from control or COA7^{KO} + COA7^{Flag} expressing cells were isolated and solubilized in 1% (w/v) digitonin solubilization buffer (20mM Bis-Tris (pH 7.0), 50 mM NaCl, 10% (v/v) glycerol). Following clarification of lysates by centrifugation (16,000 g, 10 mins, 4°C), supernatants were applied to Flag affinity gel (Sigma) and incubated for 2 hours at 4°C with gentle rotation. Affinity gel was washed with 0.1% digitonin in solubilization buffer. Proteins were eluted using 150 µg/mL Flag peptide in solubilization buffer with 0.1% (w/v) digitonin. Elutions were mixed with 5 volumes of ice-cold acetone and incubated at -20°C overnight. Precipitated proteins were then pelleted by centrifugation (16,000 g, 10 mins, 4°C) and subsequently solubilized in 50 mM ammonium bicarbonate and 8M urea. This was then sonicated in a water bath at room temperature for 15 min before addition of Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) to a final concentration of 5 mM and chloroacetamide (Sigma Aldrich) to a final concentration of 50 mM and the sample was incubated at 37 °C for 30 mins while shaking. The sample was then treated with ammonium bicarbonate to dilute the urea to a final concentration of 2 M. Trypsin (1 µg; Promega) was added to the sample before incubation overnight at 37°C. After this, the sample was acidified with 10% trifluoroacetic acid (TFA) to a final concentration of 1%. Stage tips were generated with two plugs of 3 MTM EmporeTM

SDB-XC Extraction Disks (Fisher Scientific, Pittsburgh, PA) that were activated with 100% acetonitrile (ACN) via centrifugation. All spins were performed at 1800 g. The tips were washed with 0.1% TFA, 2% ACN three times. The sample was added to the stage tip and eluted with 80% ACN and 0.1% TFA. The eluates were subsequently dried down using a SpeedVac.

Mass-spectrometry and Data Analysis

Peptides from the above procedures were reconstituted in 2% ACN (v/v), 0.1% TFA (v/v) and transferred to autosampler vials for analysis by online nano-HPLC/electrospray ionization-MS/MS on either a ThermoFisher Scientific Orbitrap Q Exactive Plus (SILAC) or Orbitrap Elite Hybrid Ion-Trap (AE-MS) instrument as follows:

For SILAC mitochondria, data were acquired on an Orbitrap Q Exactive Plus instrument connected to an Ultimate 3000 HPLC. Peptides were loaded onto a trap column (Dionex-C18 trap column 75 $\mu\text{m} \times 2\text{ cm}$, 3 μm , particle size, 100 $\text{\AA}$ pore size; ThermoFisher Scientific) at 5 $\mu\text{l/min}$ before switching the pre-column in line with the analytical column (Dionex-C18 analytical column 75 $\mu\text{m} \times 50\text{ cm}$, 2 μm particle size, 100 $\text{\AA}$ pore size; ThermoFisher Scientific). The separation of peptides was performed at 300 nl/min using a 125 min non-linear ACN gradient of buffer A [0.1% formic acid, 2% ACN, 5% DMSO] and buffer B [0.1% formic acid in ACN, 5% DMSO]. The mass spectrometer was operated in positive-ionization mode with spray voltage set at 1.9 kV and source temperature at 275 $^{\circ}\text{C}$. Lockmass of 401.92272 from DMSO was used. Data were collected using the Data Dependent Acquisition using m/z 375–1400 m/z scan range at 70,000 m/z resolution with AGC target of 3e6. HCD for MS/MS was performed on the top 15 most intense ions with charges ≥ 2 . Fragment ion spectra were acquired in Orbitrap at 17,500 m/z resolution, AGC target of 5e4 and maximum injection time of 50ms. Dynamic exclusion with of 30 seconds was applied for repeated precursors.

AE-MS was acquired on an Orbitrap Elite Hybrid Ion-Trap instrument connected to an Ultimate 3000 HPLC (ThermoFisher Scientific). Chromatography was performed as above but using a 90 min non-linear ACN gradient. Data was collected in Data Dependent Acquisition (DDA) mode using m/z 300–1650 as MS scan range, rCID for MS/MS of the 20 most intense ions. Lockmass of 401.92272 from DMSO was used. Other instrument parameters were: MS scan at 100,000 resolution, maximum injection time 150 ms, AGC

target 1E6, CID at 30% energy for a maximum injection time of 150 ms with AGC target of 5000. Dynamic exclusion with of 30 seconds was applied for repeated precursors.

Raw files were analyzed using the MaxQuant platform (61) version 1.6.16.0 searching against the UniProt human database containing 42,434 reviewed, canonical, isoforms entries (June 2019) and a database containing 246 common contaminants. For label-free (LFQ) AE-MS experiments, default search parameters were used with “Label free quantitation” set to “LFQ” and “Match between runs” enabled. In all cases, Trypsin/P cleavage specificity (cleaves after lysine or arginine, even when proline is present) was used with a maximum of 2 missed cleavages. Oxidation of methionine and N-terminal acetylation were specified as variable modifications. Carbamidomethylation of cysteine was set as a fixed modification. A search tolerance of 4.5 ppm was used for MS1 and 20 ppm for MS2 matching. False discovery rates (FDR) were determined through the target-decoy approach set to 1% for both peptides and proteins.

Using the Perseus platform (62) version 1.6.5.0, or 1.6.14.0 or 1.6.15.0, proteins group LFQ intensities were log2 transformed. Values listed as being “Only identified by site,” “Reverse,” or “Contaminants” were removed from the data set. Mitochondrial annotations were imported by matching with the Mitocarta2.0 data set (63) by gene name and/or ENSG identifier. Experimental groups were assigned to each set of triplicates and the number of valid values for row group calculated. For each experiment (containing a control and an enrichment group) rows having less than 3 valid values in the enrichment group were removed and the missing values in the relevant control group imputed to values consistent with the limit of detection. Proteins identified from less than 2 unique peptides were excluded from further analysis. A modified two-sided t test based on permutation-based FDR statistics (62) was performed between the two groups. The negative logarithmic p values were plotted against the differences between the log2 means for the two groups. The significance threshold used for these experiments is noted in the relevant figure legend and supplementary tables. Non-mitochondrial proteins were removed based on their absence from MitoCarta 2.0 (63), before plotting of the data. For SILAC experiments, default search parameters were used with multiplicity set to 2 (Lys8, Arg10) and “Match between runs” enabled. Using the Perseus platform (62) version 1.6.5.0, proteins group normalized H/L ratios were log2 transformed. Label switched samples (L/H) were inverted to KO/Control orientation before this step. Values listed as being “Only identified by site,” “Reverse,” or “Contaminants” were removed from the data set. Mitochondrial annotations were imported by matching with the

Mitocarta2.0 data set (63) by gene name and/or ENSG identifier. Experimental groups were assigned to each set of triplicates rows with < 2 valid values for each group removed. Non-mitochondrial proteins were removed based on the absence of a Mitocarta2.0 annotation, and a one sample Student's two-sided t test was performed within each group. The negative logarithmic p values were plotted against the differences between the mean ratios for each group. A significance threshold ($p < 0.05$) was used for all experiments. For the topographical heatmap, log₂-transformed median SILAC ratios were mapped onto the structure of complex IV (PDB 5Z62).

Protein Overexpression, Purification, and Characterization

The R39W, Y137C, Y137S, S149I, H119A, M156A and H119A/M156A mutations were introduced into the ORF encoding the COA7 protein in the pGEX-6p-1-COA7 plasmid using a Q5[®] Site-Directed Mutagenesis kit (New England, BioLabs), according to the manufacturer's instructions. Forward and reverse primers for all mutant proteins were designed using NEBaseChanger online tool (<http://nebasechanger.neb.com/>) and are given in **Table S1**. The DNA sequence encoding the variant was amplified *via* PCR followed by ligation and template DNA removal using KLD enzyme mix (containing a kinase, a ligase and DpnI, New England BioLabs[®]).

The pGEX-6p-1 plasmids containing DNA sequence encoding COA7, ^{R39W}COA7, ^{Y137C}COA7, ^{Y137S}COA7, ^{S149I}COA7, ^{H119A}COA7, ^{M156A}COA7 and ^{H119AM156A}COA7, were individually transformed into *Escherichia coli* strain SHuffle[®] T7 (New England, BioLabs). Cultures were grown at 30°C in Luria Broth (LB) supplemented with ampicillin (100 µg/mL), chloramphenicol (35 µg/mL) and streptomycin (50 µg/mL) to an OD₆₀₀ of 0.8, induced with isopropyl β-D-1-thiogalactopyranoside (IPTG, 0.2 mM) and harvested after 16 hours (with shaking) at 16°C.

Purified COA7, ^{R39W}COA7, ^{Y137C}COA7, ^{Y137S}COA7, ^{S149I}COA7, ^{H119A}COA7, ^{M156A}COA7 and ^{H119AM156A}COA7 proteins were prepared by affinity and size exclusion chromatography. Frozen cell pellets of the GST-fusion proteins were thawed at room temperature and resuspended in PBS (phosphate-buffered saline pH 7.4). Cells were disrupted by passage through a TS series bench top cell disruptor (Constant Systems Ltd) at 35 kpsi. Cell debris were removed by centrifugation (Beckman JLA-25.50, 30000 g, 20 min, 4°C) and the soluble

fraction was incubated with glutathione sepharoseTM 4B resin (GE Healthcare) equilibrated with PBS. The GST tag was cleaved with PreScission Protease followed by size-exclusion chromatography (SEC; HiLoad 16/600 Superdex 75 pg, GE Healthcare; 20 mM NaHepes pH 7.2, 50 mM NaCl). Cleavage of the N-terminal GST tag introduced five additional residues (GPLGS) to the N-terminus of all proteins. The purified proteins were concentrated to 20 mg/mL before storage at -80°C.

Sedimentation velocity analysis

COA7 and ^{Y137C}COA7 samples were analyzed using an XL-I analytical ultracentrifuge (Beckman Coulter, Fullerton, CA) equipped with an AnTi-60 rotor. Protein samples were loaded in the sample compartment of double-sector epon centrepieces, with buffer (20 mM NaHepes pH 7.5, 50 mM NaCl) in the reference compartment. Radial absorbance data was acquired at 20°C using a rotor speed of 50,000 rpm and a wavelength of 280 nm, with radial increments of 0.003 cm in continuous scanning mode. The sedimenting boundaries were fitted to a model that describes the sedimentation of a distribution of sedimentation coefficients with no assumption of heterogeneity ($c(s)$) using the program SEDFIT (64). Data were fitted using a regularization parameter of $p = 0.95$, floating frictional ratios, and 100 sedimentation coefficient increments in the range of 0–15 S.

Differential scanning fluorimetry

The melting temperatures (T_m) of the COA7, ^{R39W}COA7, ^{Y137C}COA7, ^{Y137S}COA7 and ^{S149I}COA7 proteins were measured by differential scanning fluorimetry with purified proteins in 20 mM NaHepes pH 7.2, 50 mM NaCl (1.0 mg/mL). Differential scanning fluorimetry experiments were performed at the CSIRO Collaborative Crystallisation Centre (<http://www.csiro.au/C3>), Melbourne, Australia.

Copper binding experiments

Measurements of the copper binding affinities of the COA7 and ATOX1 proteins were performed as previously described (21). Briefly, purified proteins (in 20 mM Tris-Mes pH 8.0) were titrated at various concentrations (1–30 μM) into solutions containing buffer (20 mM Tris-Mes pH 7.0), CuSO₄ (20 μM), Bcs (200 μM) and NH₂OH (1 mM). The exchange of Cu(I) from the [Cu^IBcs₂]³⁻ complex to the proteins was monitored by measuring the absorbance of the resulting solutions at 483 nm. The data were analyzed by plotting

[Cu^IBcs₂]³⁻ (as determined from the absorbance values at 483 nm) versus protein:copper ratios, and the data fitted using the equation previously described (21, 42).

Fe³⁺-PPIX and PPIX titrations for the COA7 and ^{H119A}COA7 and ^{M156A}COA7 proteins

Fe³⁺-PPIX and PPIX binding assays were performed for the COA7 and ^{H119A}COA7 proteins by difference spectroscopy at room temperature and aerobically using a double beam UV-Vis spectrophotometer (Cary 5000 UV-Vis-NIR spectrophotometer) as previously described (43, 44). A baseline was selected by scanning samples from 280 to 750 nm, with buffer (20 mM NaHepes pH 7.2, 50 mM NaCl) in the reference cuvette and 1.0 μM of purified protein (in 20 mM NaHepes pH 7.2, 50 mM NaCl) in the sample cuvette. After the baseline was set, increasing concentrations of Fe³⁺-PPIX (0-10 μM) or PPIX (0-200 μM) were added to both the sample and the reference cuvettes, therefore the readings only reflected the binding of Fe³⁺-PPIX or PPIX to the proteins and not the absorbance of the free ligands. Note that only the titrations with Fe³⁺-PPIX yielded spectra that indicated ligand binding (**Fig. 5 and S10**) and that the PPIX titration was carried out for the COA7 protein only. The data were analyzed by plotting absorbance at 416 nm versus hemin concentration and the data were fitted using equation describing a single binding site model ($Y = \frac{B_{max} * X}{K_d + X}$) using GraphPad Prism software.

Fe³⁺-PPIX and PPIX loading of COA7, ^{H119A}COA7 and ^{M156A}COA7

Purified protein samples were incubated for 1 hour with Fe³⁺-PPIX or PPIX (molar ratio 1:2 protein:ligand) in 20 mM NaHepes pH 7.2, 50 mM NaCl. In order to remove the excess Fe³⁺-PPIX or PPIX from the mixture, the incubated protein sample was applied to a SEC column (10 mL, Bio-Gel P-4 Gel, Bio-Rad). The presence of bound Fe³⁺-PPIX or PPIX in the eluted fractions was detected by UV-Vis spectrophotometry (between 250-700 nm; Cary 300 UV-Vis Absorbance spectrophotometer) and those protein fractions containing bound hemin (with absorption at 416-420 nm) were pooled and concentrated by centrifugal ultrafiltration.

The UV-Vis spectra of Fe²⁺-PPIX-COA7 were recorded (65) following incubation of purified COA7 with one molar equivalent of Fe³⁺-PPIX, and the subsequent addition of 1.0 mM sodium dithionite to reduce all Fe³⁺ to Fe²⁺. UV-Vis spectra of Fe³⁺-PPIX loaded COA7, ^{H119A}COA7 and ^{M156A}COA7 samples were carried out in presence of sodium cyanide (1.0 mM) (66).

789

790 **Metal analysis of the Fe³⁺-PPIX-COA7 complex**

791 The Fe:COA7 stoichiometry of the Fe³⁺-PPIX-COA7 complex was determined by ICP-MS.
 792 The Fe³⁺-PPIX-COA7 protein (at various concentrations: 10-100 ppb in buffer: 20 mM
 793 NaHepes, pH 7.2, 50 mM NaCl) was digested in 200 µL of 35% HNO₃ (Merck Suprapur,
 794 Australia), diluted to a final volume of 1.0 mL using MilliQ H₂O and heated at 96°C for 15
 795 min. Insoluble material was removed by centrifugation at 18,000 × g for 25 min. Technical
 796 triplicate measurements of samples and controls were analyzed on an Agilent 8900 ICP-
 797 MS/MS with a MicroMist nebulizer (Glass Expansion, Australia). Torch positioning, sample
 798 depth adjustment and lens optimization were set according to manufacturer recommendations
 799 while other instrumental parameters were optimized during a batch-specific user tune.
 800 Helium collision gas flow rate of 5 mL/min to minimize polyatomic interferences. Samples
 801 were introduced via an integrated automation system (IAS) autosampler (Agilent) using a
 802 peristaltic pump. The division of the slope of the standard curve and a plot of Fe
 803 concentration versus protein concentration was calculated to yield a value for the Fe:COA7
 804 protein sample.

805

806 **Protein Crystallization and Data Collection**

807 Initial high-throughput crystallization experiments using COA7 (20 mg/mL) were performed
 808 at the CSIRO Collaborative Crystallisation Centre (<http://www.csiro.au/C3>), Melbourne,
 809 Australia using c3 salty screens (c3_5 and c3_6) by sitting drop vapor diffusion in 96-well
 810 plates. Diffraction-quality, plate-like crystals of COA7 grew overnight to a maximum size of
 811 250 x 30 x 20 µ at 20°C by hanging-drop vapor diffusion with drops containing equal
 812 volumes (1 µL) of COA7 (20 mg/mL in 20 mM NaHepes pH 7.2, 50 mM NaCl) and
 813 crystallization solution (0.23 M sodium Mes pH 6.72, 2.35 M ammonium sulfate, 8.2% (v/v)
 814 ethanol and 4.5% (v/v) pentaerythritol ethoxylate (3/4 EO/OH) equilibrated against 500 µL
 815 reservoir solution. COA7 crystals were cryoprotected in reservoir solution containing 25%
 816 (v/v) glycerol before flash-cooling in liquid nitrogen. Diffraction data were recorded on an
 817 EIGER X 16M detector at the Australian Synchrotron on beamline MX2 at a wavelength of
 818 0.954 Å. All data were collected at 100 K and were processed and scaled with XDS (67) and
 819 Aimless (68). Data collection statistics are detailed in **Table S3**.

820

821 **Structure Solution and Refinement**

The crystal structure of COA7 was solved by molecular replacement using the program PHASER (69) from the CCP4 suite (70). Models including residues 33-104 and 110-216 of the HcpC crystal structure (PDB 1OUV) (34) were used as the search models after removal of all water molecules. The structure was refined using REFMAC5 (71), with manual model building and the addition of water molecules carried out in COOT (72). The refinement of the model converged with residuals $R = 0.20$ and $R_{\text{free}} = 0.25$ (**Table S3**) and showed excellent geometry as determined by MOLPROBITY (73) (**Table S3**).

Data Availability

The mass spectrometry proteomics data will be deposited in the ProteomeXchange Consortium via the PRIDE (74) partner repository upon publication. The atomic coordinates and structure factors have been deposited in the Protein Data Bank (<https://www.rcsb.org>; PDB 7MQZ).

Acknowledgements

This study was funded by the Australian Research Council (DP140102746 and FT180100397 to MJM), the National Health and Medical Research Council (GNT1165217 to MTR and MJM; GNT1140906 and 1140851 to DAS) and an Australian Government Research Training Program Scholarship to SM. Part of this study was carried out using the MX2 beamline at the Australian Synchrotron, which is part of ANSTO, and made use of the ACRF detector. We thank the beamline staff for their enthusiastic and professional support. We thank Dr. Yee-Foong Mok at the Melbourne Protein characterization facility, The Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne for performing the AUC experiments and data analyzes. We thank the Bio21 Mass Spectrometry and Proteomics Facility for the provision of instrumentation, training, and technical support. We thank Daniel Machell for Pymol and coding assistance and Dr Katie Ganio and Prof. Christopher McDevitt for the ICP-MS analysis. Finally, we thank Mrs Mahnaz Dideh Var for her extraordinary support with the structural biology aspects of this study.

References

1. E. Fernandez-Vizarra, V. Tiranti, M. Zeviani, Assembly of the oxidative phosphorylation system in humans: what we have learned by studying its defects. *Biochim Biophys Acta* **1793**, 200-211 (2009).
2. T. Tsukihara, H. Aoyama, E. Yamashita, T. Tomizaki, H. Yamaguchi, K. Shinzawa-Itoh, R. Nakashima, R. Yaono, S. Yoshikawa, The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 Å. *Science* **272**, 1136-1144 (1996).
3. D. U. Mick, T. D. Fox, P. Rehling, Inventory control: cytochrome c oxidase assembly regulates mitochondrial translation. *Nat Rev Mol Cell Biol* **12**, 14-20 (2011).
4. E. Balsa, R. Marco, E. Perales-Clemente, R. Szklarczyk, E. Calvo, M. O. Landazuri, J. A. Enriquez, NDUFA4 is a subunit of complex IV of the mammalian electron transport chain. *Cell Metab* **16**, 378-386 (2012).
5. A. Signes, E. Fernandez-Vizarra, Assembly of mammalian oxidative phosphorylation complexes I-V and supercomplexes. *Essays Biochem* **62**, 255-270 (2018).
6. A. Timon-Gomez, E. Nyvltova, L. A. Abriata, A. J. Vila, J. Hosler, A. Barrientos, Mitochondrial cytochrome c oxidase biogenesis: Recent developments. *Semin Cell Dev Biol* **76**, 163-178 (2018).
7. J. A. Letts, K. Fiedorczuk, L. A. Sazanov, The architecture of respiratory supercomplexes. *Nature* **537**, 644-648 (2016).
8. Y. Li, M. D'Aurelio, J. H. Deng, J. S. Park, G. Manfredi, P. Hu, J. Lu, Y. Bai, An assembled complex IV maintains the stability and activity of complex I in mammalian mitochondria. *J Biol Chem* **282**, 17557-17562 (2007).
9. F. Diaz, H. Fukui, S. Garcia, C. T. Moraes, Cytochrome c oxidase is required for the assembly/stability of respiratory complex I in mouse fibroblasts. *Mol Cell Biol* **26**, 4872-4881 (2006).
10. K. Mohanraj, M. Wasilewski, C. Beninca, D. Cysewski, J. Poznanski, P. Sakowska, Z. Bugajska, M. Deckers, S. Dennerlein, E. Fernandez-Vizarra, P. Rehling, M. Dadlez, M. Zeviani, A. Chacinska, Inhibition of proteasome rescues a pathogenic variant of respiratory chain assembly factor COA7. *EMBO Mol Med* **11**, e9561 (2019).
11. V. Kozjak-Pavlovic, F. Prell, B. Thiede, M. Gotz, D. Wosiek, C. Ott, T. Rudel, Clorf163/RESA1 is a novel mitochondrial intermembrane space protein connected to respiratory chain assembly. *J Mol Biol* **426**, 908-920 (2014).

12. M. A. Andrade, C. Perez-Iratxeta, C. P. Ponting, Protein repeats: structures, functions, and evolution. *J Struct Biol* **134**, 117-131 (2001).
13. P. R. Mittl, W. Schneider-Brachert, Sell-like repeat proteins in signal transduction. *Cell Signal* **19**, 20-31 (2007).
14. L. Cervený, A. Strasková, V. Danková, A. Hartlová, M. Cecková, F. Staud, J. Stulík, Tetratricopeptide repeat motifs in the world of bacterial pathogens: role in virulence mechanisms. *Infect Immun* **81**, 629-635 (2013).
15. G. L. Blatch, M. Lassle, The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays* **21**, 932-939 (1999).
16. A. V. Kajava, Review: proteins with repeated sequence--structural prediction and modeling. *J Struct Biol* **134**, 132-144 (2001).
17. L. D. D'Andrea, L. Regan, TPR proteins: the versatile helix. *Trends Biochem Sci* **28**, 655-662 (2003).
18. A. Martinez Lyons, A. Ardisson, A. Reyes, A. J. Robinson, I. Moroni, D. Ghezzi, E. Fernandez-Vizarra, M. Zeviani, COA7 (C1orf163/RESA1) mutations associated with mitochondrial leukoencephalopathy and cytochrome c oxidase deficiency. *J Med Genet* **53**, 846-849 (2016).
19. Y. Higuchi *et al.*, Mutations in COA7 cause spinocerebellar ataxia with axonal neuropathy. *Brain* **141**, 1622-1636 (2018).
20. F. A. Ran, P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott, F. Zhang, Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**, 2281-2308 (2013).
21. D. A. Stroud, M. J. Maher, C. Lindau, F. N. Vogtle, A. E. Frazier, E. Surgenor, H. Mountford, A. P. Singh, M. Bonas, S. Oeljeklaus, B. Warscheid, C. Meisinger, D. R. Thorburn, M. T. Ryan, COA6 is a mitochondrial complex IV assembly factor critical for biogenesis of mtDNA-encoded COX2. *Hum Mol Genet* **24**, 5404-5415 (2015).
22. L. E. Formosa, A. Hofer, C. Tischner, T. Wenz, M. T. Ryan, Translation and Assembly of Radiolabeled Mitochondrial DNA-Encoded Protein Subunits from Cultured Cells and Isolated Mitochondria. *Methods Mol Biol* **1351**, 115-129 (2016).
23. S. E. Ong, B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey, M. Mann, Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* **1**, 376-386 (2002).
24. R. D. Pitceathly, S. Rahman, Y. Wedatilake, J. M. Polke, S. Cirak, A. R. Foley, A. Sailer, M. E. Hurles, J. Stalker, I. Hargreaves, C. E. Woodward, M. G. Sweeney, F. Muntoni, H. Houlden, J. W. Taanman, M. G. Hanna, U. K. Consortium, NDUF4

- 920 mutations underlie dysfunction of a cytochrome c oxidase subunit linked to human
921 neurological disease. *Cell Rep* **3**, 1795-1805 (2013).
- 922 25. D. H. Hock, B. Reljic, C. S. Ang, L. Muellner-Wong, H. S. Mountford, A. G.
923 Compton, M. T. Ryan, D. R. Thorburn, D. A. Stroud, HIGD2A is Required for
924 Assembly of the COX3 Module of Human Mitochondrial Complex IV. *Mol Cell*
925 *Proteomics* **19**, 1145-1160 (2020).
- 926 26. S. Zong, M. Wu, J. Gu, T. Liu, R. Guo, M. Yang, Structure of the intact 14-subunit
927 human cytochrome c oxidase. *Cell Res* **28**, 1026-1034 (2018).
- 928 27. E. Hangen *et al.*, Interaction between AIF and CHCHD4 Regulates Respiratory Chain
929 Biogenesis. *Mol Cell* **58**, 1001-1014 (2015).
- 930 28. S. A. Swenson, C. M. Moore, J. R. Marcero, A. E. Medlock, A. R. Reddi, O.
931 Khalimonchuk, From Synthesis to Utilization: The Ins and Outs of Mitochondrial
932 Heme. *Cells* **9**, 579 (2020).
- 933 29. S. Deshwal, K. U. Fiedler, T. Langer, Mitochondrial Proteases: Multifaceted
934 Regulators of Mitochondrial Plasticity. *Annu Rev Biochem* **89**, 501-528 (2020).
- 935 30. R. R. Cupo, J. Shorter, Skd3 (human ClpB) is a potent mitochondrial protein
936 disaggregase that is inactivated by 3-methylglutaconic aciduria-linked mutations.
937 *Elife* **9**, e55279 (2020).
- 938 31. V. Hung, P. Zou, H. W. Rhee, N. D. Udeshi, V. Cracan, T. Svinkina, S. A. Carr, V. K.
939 Mootha, A. Y. Ting, Proteomic mapping of the human mitochondrial intermembrane
940 space in live cells via ratiometric APEX tagging. *Mol Cell* **55**, 332-341 (2014).
- 941 32. E. Krissinel, K. Henrick, Secondary-structure matching (SSM), a new tool for fast
942 protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr*
943 **60**, 2256-2268 (2004).
- 944 33. L. Luthy, M. G. Grutter, P. R. Mittl, The crystal structure of Helicobacter pylori
945 cysteine-rich protein B reveals a novel fold for a penicillin-binding protein. *J Biol*
946 *Chem* **277**, 10187-10193 (2002).
- 947 34. L. Luthy, M. G. Grutter, P. R. Mittl, The crystal structure of Helicobacter cysteine-
948 rich protein C at 2.0 Å resolution: similar peptide-binding sites in TPR and SEL1-like
949 repeat proteins. *J Mol Biol* **340**, 829-841 (2004).
- 950 35. S. Jones, J. M. Thornton, Protein-protein interactions: a review of protein dimer
951 structures. *Prog Biophys Mol Biol* **63**, 31-65 (1995).

36. C. Scheufler, A. Brinker, G. Bourenkov, S. Pegoraro, L. Moroder, H. Bartunik, F. U. Hartl, I. Moarefi, Structure of TPR domain-peptide complexes: critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. *Cell* **101**, 199-210 (2000).
37. A. Brinker, C. Scheufler, F. Von Der Mulbe, B. Fleckenstein, C. Herrmann, G. Jung, I. Moarefi, F. U. Hartl, Ligand discrimination by TPR domains. Relevance and selectivity of EEVD-recognition in Hsp70 x Hop x Hsp90 complexes. *J Biol Chem* **277**, 19265-19275 (2002).
38. G. J. Gatto, Jr., B. V. Geisbrecht, S. J. Gould, J. M. Berg, Peroxisomal targeting signal-1 recognition by the TPR domains of human PEX5. *Nat Struct Biol* **7**, 1091-1095 (2000).
39. S. A. Seabrook, J. Newman, High-throughput thermal scanning for protein stability: making a good technique more robust. *ACS Comb Sci* **15**, 387-392 (2013).
40. Z. Xiao, J. Brose, S. Schimo, S. M. Ackland, S. La Fontaine, A. G. Wedd, Unification of the copper(I) binding affinities of the metallo-chaperones Atx1, Atox1, and related proteins: detection probes and affinity standards. *J Biol Chem* **286**, 11047-11055 (2011).
41. L. Banci, I. Bertini, S. Ciofi-Baffoni, A. Janicka, M. Martinelli, H. Kozlowski, P. Palumaa, A structural-dynamical characterization of human Cox17. *J Biol Chem* **283**, 7912-7920 (2008).
42. J. Brose, S. La Fontaine, A. G. Wedd, Z. Xiao, Redox sulfur chemistry of the copper chaperone Atox1 is regulated by the enzyme glutaredoxin 1, the reduction potential of the glutathione couple GSSG/2GSH and the availability of Cu(I). *Metallomics* **6**, 793-808 (2014).
43. N. Gupta, S. W. Ragsdale, Thiol-disulfide redox dependence of heme binding and heme ligand switching in nuclear hormone receptor rev-erb{beta}. *J Biol Chem* **286**, 4392-4403 (2011).
44. I. C. Soto, F. Fontanesi, R. S. Myers, P. Hamel, A. Barrientos, A heme-sensing mechanism in the translational regulation of mitochondrial cytochrome c oxidase biogenesis. *Cell Metab* **16**, 801-813 (2012).
45. S. Hofbauer, A. Hagmuller, I. Schaffner, G. Mlynek, M. Krutzler, G. Stadlmayr, K. F. Pirker, C. Obinger, H. Daims, K. Djinić-Carugo, P. G. Furtmüller, Structure and heme-binding properties of HemQ (chlorite dismutase-like protein) from *Listeria monocytogenes*. *Arch Biochem Biophys* **574**, 36-48 (2015).

46. D. J. Deredge, W. Huang, C. Hui, H. Matsumura, Z. Yue, P. Moenne-Loccoz, J. Shen, P. L. Wintrode, A. Wilks, Ligand-induced allostery in the interaction of the *Pseudomonas aeruginosa* heme binding protein with heme oxygenase. *Proc Natl Acad Sci U S A* **114**, 3421-3426 (2017).
47. E. A. Della Pia, Q. Chi, M. Elliott, J. E. Macdonald, J. Ulstrup, D. D. Jones, Redox tuning of cytochrome b562 through facile metal porphyrin substitution. *Chem Commun (Camb)* **48**, 10624-10626 (2012).
48. J. L. Gao, A. H. Kwan, A. Yammine, X. Zhou, J. Trehwella, B. M. Hugrass, D. A. T. Collins, J. Horne, P. Ye, D. Harty, K. A. Nguyen, D. A. Gell, N. Hunter, Structural properties of a haemophore facilitate targeted elimination of the pathogen *Porphyromonas gingivalis*. *Nat Commun* **9**, 4097 (2018).
49. F. Lederer, A. Glatigny, P. H. Bethge, H. D. Bellamy, F. S. Matthew, Improvement of the 2.5 Å resolution model of cytochrome b562 by redetermining the primary structure and using molecular graphics. *J Mol Biol* **148**, 427-448 (1981).
50. H. J. Kim, O. Khalimonchuk, P. M. Smith, D. R. Winge, Structure, function, and assembly of heme centers in mitochondrial respiratory complexes. *Biochim Biophys Acta* **1823**, 1604-1616 (2012).
51. D. A. Hanna, R. M. Harvey, O. Martinez-Guzman, X. Yuan, B. Chandrasekharan, G. Raju, F. W. Outten, I. Hamza, A. R. Reddi, Heme dynamics and trafficking factors revealed by genetically encoded fluorescent heme sensors. *Proc Natl Acad Sci U S A* **113**, 7539-7544 (2016).
52. D. A. Stroud, E. E. Surgenor, L. E. Formosa, B. Reljic, A. E. Frazier, M. G. Dibley, L. D. Osellame, T. Stait, T. H. Beilharz, D. R. Thorburn, A. Salim, M. T. Ryan, Accessory subunits are integral for assembly and function of human mitochondrial complex I. *Nature* **538**, 123-126 (2016).
53. L. E. Formosa, L. Muellner-Wong, B. Reljic, A. J. Sharpe, T. D. Jackson, T. H. Beilharz, D. Stojanovski, M. Lazarou, D. A. Stroud, M. T. Ryan, Dissecting the Roles of Mitochondrial Complex I Intermediate Assembly Complex Factors in the Biogenesis of Complex I. *Cell Rep* **31**, 107541 (2020).
54. Z. Liu, O. Chen, J. B. J. Wall, M. Zheng, Y. Zhou, L. Wang, H. R. Vaseghi, L. Qian, J. Liu, Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Sci Rep* **7**, 2193 (2017).
55. A. J. Johnston, J. Hoogenraad, D. A. Dougan, K. N. Truscott, M. Yano, M. Mori, N. J. Hoogenraad, M. T. Ryan, Insertion and assembly of human tom7 into the

1019 preprotein translocase complex of the outer mitochondrial membrane. *J Biol Chem*
1020 **277**, 42197-42204 (2002).

1021 56. M. McKenzie, M. Lazarou, D. R. Thorburn, M. T. Ryan, Analysis of mitochondrial
1022 subunit assembly into respiratory chain complexes using Blue Native polyacrylamide
1023 gel electrophoresis. *Anal Biochem* **364**, 128-137 (2007).

1024 57. H. Schagger, G. von Jagow, Tricine-sodium dodecyl sulfate-polyacrylamide gel
1025 electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal*
1026 *Biochem* **166**, 368-379 (1987).

1027 58. I. Wittig, H. P. Braun, H. Schagger, Blue native PAGE. *Nat Protoc* **1**, 418-428
1028 (2006).

1029 59. N. A. Kulak, G. Pichler, I. Paron, N. Nagaraj, M. Mann, Minimal, encapsulated
1030 proteomic-sample processing applied to copy-number estimation in eukaryotic cells.
1031 *Nat Methods* **11**, 319-324 (2014).

1032 60. L. E. Formosa, B. Reljic, A. J. Sharpe, D. H. Hock, L. Muellner-Wong, D. A. Stroud,
1033 M. T. Ryan, Optic atrophy-associated TMEM126A is an assembly factor for the
1034 ND4-module of mitochondrial complex I. *Proc Natl Acad Sci U S A* **118** (2021).

1035 61. S. Tyanova, T. Temu, J. Cox, The MaxQuant computational platform for mass
1036 spectrometry-based shotgun proteomics. *Nat Protoc* **11**, 2301-2319 (2016).

1037 62. S. Tyanova, T. Temu, P. Sinitcyn, A. Carlson, M. Y. Hein, T. Geiger, M. Mann, J.
1038 Cox, The Perseus computational platform for comprehensive analysis of (prote)omics
1039 data. *Nat Methods* **13**, 731-740 (2016).

1040 63. S. E. Calvo, K. R. Clauser, V. K. Mootha, MitoCarta2.0: an updated inventory of
1041 mammalian mitochondrial proteins. *Nucleic Acids Res* **44**, D1251-1257 (2016).

1042 64. P. Schuck, Size-distribution analysis of macromolecules by sedimentation velocity
1043 ultracentrifugation and lamm equation modeling. *Biophys J* **78**, 1606-1619 (2000).

1044 65. I. Barr, F. Guo, Pyridine Hemochromagen Assay for Determining the Concentration
1045 of Heme in Purified Protein Solutions. *Bio Protoc* **5**, e1594 (2015).

1046 66. E. A. Berry, B. L. Trumpower, Simultaneous determination of hemes a, b, and c from
1047 pyridine hemochrome spectra. *Anal Biochem* **161**, 1-15 (1987).

1048 67. W. Kabsch, XDS. *Acta Crystallogr* **D66**, 125-132 (2010).

1049 68. P. R. Evans, G. N. Murshudov, How good are my data and what is the resolution?
1050 *Acta Crystallogr* **D69**, 1204-1214 (2013).

69. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, Phaser crystallographic software. *Journal of Applied Crystallography* **40**, 658-674 (2007).
70. M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, Overview of the CCP4 suite and current developments. *Acta Crystallogr* **D67**, 235-242 (2011).
71. G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr* **D67**, 355-367 (2011).
72. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr* **D66**, 486-501 (2010).
73. V. B. Chen, W. B. Arendall, 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr* **D66**, 12-21 (2010).
74. Y. Perez-Riverol *et al.*, The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* **47**, D442-D450 (2019).