nature communications

Article

Kingdom-specific lipid unsaturation calibrates sequence evolution in membrane arm subunits of eukaryotic respiratory complexes

Received: 7 June 2024

Accepted: 12 February 2025

Published online: 27 February 2025

Check for updates

Pooja Gupta $\mathbb{O}^{1,2}$, Sristi Chakroborty \mathbb{O}^1 , Arun K. Rathod $\mathbb{O}^{2,3}$, K. Ranjith Kumar¹, Shreya Bhat¹, Suparna Ghosh^{1,2}, Pallavi Rao T^{1,2}, Kameshwari Yele¹, Raman Bakthisaran¹, R. Nagaraj¹, Moutusi Manna $\mathbb{O}^{2,3}$ & Swasti Raychaudhuri $\mathbb{O}^{1,2}$

Sequence evolution of protein complexes (PCs) is constrained by proteinprotein interactions (PPIs). PPI-interfaces are predominantly conserved and hotspots for disease-related mutations. How do lipid-protein interactions (LPIs) constrain sequence evolution of membrane-PCs? We explore Respiratory Complexes (RCs) as a case study as these allow to compare sequence evolution in subunits exposed to both lipids in inner-mitochondrial membrane (IMM) and lipid-free aqueous matrix. We find that lipid-exposed surfaces of the IMM-subunits but not of the matrix subunits are populated with non-PPI disease-causing mutations signifying LPIs in stabilizing RCs. Further, IMMsubunits including their exposed surfaces show high intra-kingdom sequence conservation but remarkably diverge beyond. Molecular Dynamics simulation suggests contrasting LPIs of structurally superimposable but sequence-wise diverged IMM-exposed helices of Complex I (CI) subunit Ndufa1 from human and Arabidopsis depending on kingdom-specific unsaturation of cardiolipin fatty acyl chains. in cellulo assays consolidate inter-kingdom incompatibility of Ndufa1-helices due to the lipid-exposed amino acids. Plant-specific unsaturated fatty acids in human cells also trigger CI-instability. Taken together, we posit that altered LPIs calibrate sequence evolution at the IMM-arms of eukaryotic RCs.

Proteins interact, and often act as protein complexes (PCs)¹. PCs executing fundamental cellular functions are conserved from prokaryotes to eukaryotes. Structures of eukaryotic PCs with prokaryotic ancestry are subtly refined in different forms of life via alterations of assembly states and gain or loss of subunits^{1,2}. Sequence of the PC-subunits simultaneously evolve but rarely diverge at the protein-protein interaction (PPI) surfaces that define the core composition,

structure, and function of the complexes. Mutations of these conserved interactions are deleterious and associated with diseases³. In contrast, exposed PC-surfaces that do not engage into PPIs tend to be least conserved^{4,5}. Intriguingly, lipid-protein interactions (LPIs) are implicated in the structural organization of membrane-PCs⁶⁻¹⁰ but how LPIs impact sequence evolution of membrane-PCs is insufficiently addressed.

¹CSIR- Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad – 500007, India. ²Academy of Scientific and Innovative Research (AcSIR), Ghaziabad – 201002, India. ³CSIR- Central Salt and Marine Chemical Research Institute, Bhavnagar – 364002, Gujrat, India. 🖂 e-mail: rcswasti@ccmb.res.in

Four respiratory complexes (Complex I to IV; RCs) and ATP synthase (CV) are fundamental for oxidative phosphorylation (OxPhos) and ATP production in aerobic eukarvotes. Nuclear-encoded OxPhossubunits translocate from cytoplasm, assemble with mitochondriaencoded subunits at the inner-mitochondrial membrane (IMM) and mitochondrial matrix^{11,12}. All the RCs and CV contain matrix and IMMarms with subunits exposing surfaces to both hydrophilic matrix and hydrophobic lipids. Mutations accumulate much faster in mitogenome than nuclear-genome and mutation-rates can even vary between two closely related eukaryotic species13. Nuclear-encoded RCsubunits and CV co-evolve to compensate for the increased mutations in the mito-encoded counterparts at the IMM-arms resulting in interspecies incompatibility of nuclear and mito-encoded subunits^{14,15}. RCs and CV also adopt lineage-specific accessory subunits to accommodate exclusive functions, heterogeneous cristae morphologies, and IMM lipid-chemistry across eukaryotes^{16–19}. The dynamic PPIs and LPIs originated thereby at the matrix-arm and IMM-subunits of RCs and CV provide an opportunity to compare PPI and LPI-mediated sequence evolution.

Here, we map human OxPhos-deficiency mutations on RCsubunits on the high-resolution structures of the complexes. We find that while point mutations in matrix-arm subunits destabilize intra or inter-subunit PPIs, ~50% of the mutations in IMM-arms locate beyond PPI-interfaces. Rather, these non-PPI mutations populate the exposed surfaces of IMM-subunits, which display remarkable sequence divergence between kingdoms plants, Fungi, and metazoans²⁰. Coevolution of mito-encoded and nuclear-encoded subunits due to mito-genome instability fail to explain this phenomenon as the crosskingdom sequence divergence is consistent in lipid-embedded mitochondrial membrane PCs lacking mito-encoded subunits. Our results indicate that kingdom-specific unsaturation in cardiolipin is an important determinant of lipid-affinity of trans-IMM helices of RCsubunits to mediate stable assembly of the complexes. Together, we posit that while surfaces of matrix-arm subunits of RCs may have evolved neutrally, surfaces of IMM-subunits undergo adaptive coevolution along with the IMM-lipids across eukaryotes.

Results

Disease-related mutations populate membrane arm surfaces of respiratory complexes

Respiratory complex I (CI) is the first and largest complex in the respiratory chain^{12,21}. Human CI is composed of 44 subunits with 17 subunits in the matrix arm, 23 subunits embedded into the inner mitochondrial membrane (IMM), and 4 subunits exposed to the intermembrane space (IMS) (Fig. 1a). Fourteen of these subunits are designated as core subunits as these are conserved from bacteria to modern eukaryotes and distributed equally between the matrix and IMM-arms. Others are eukaryotic innovations and known as accessory subunits. OMIM database indicated 131 mutations in CI-subunits associated with OxPhos-deficiency diseases including Leigh, LHON, Dystonia and MELAS Syndrome (Fig. 1b, Supplementary Data S1a). Out of these, 43 mutations represented nucleotide-inversions, splicing defects, nonsense mutations, and frameshifts resulting in either truncation or extensive modulation of subunit sequences. Rest 88 missense mutations change single amino acids of which 6 showed positional-discrepancy as per assignment in the highly resolved cryoelectron microscopy structure of human CI22, 2 were redundant reports, and 1 mutation was at the mitochondria targeting signal of the nuclear encoded subunit Ndufa10 (Supplementary Data S2a). We excluded the discrepancies and mapped the remaining 79 point mutations on the solved structure of CI to identify the pivotal PPIs defining the PC.

Forty mutations were mapped on the matrix arm (Fig. 1c, Supplementary Data S2a). Seven of these involved buried amino acids (atoms found within 3Å² of neighboring space). Four buried amino

acids were found to be engaged in either inter-subunit or intra-subunit inter-chain H-bonding, salt bridge, or cation- π interactions (Fig. 1c. Supplementary Data S2a). Two mutations. A159D in Ndufs8 and P76O in Ndufb8, switches the amino acids from nonpolar to polar, V122M in Ndufs7 may result in steric hindrance with neighboring C171, L134, S149 and M120. Among the 33 matrix-arm mutations that involved amino acids with at least 1 atom exposed (no neighboring atoms within 3\AA^2 space), 25 were engaged in PPI. Seven mutations were mapped on the IMS-subunits and 5 of these participate in intra-subunit inter-chain PPI in the solved structure (Fig. 1c, Supplementary Data S2a). Thus, majority of the matrix arm and IMS mutations were likely to destabilize PPIs. In contrast, at the IMM-arm 10 out of 11 and 17 out of 21 mutations involved in buried and exposed amino acids, respectively, were not engaged in PPI (Fig. 1c, Supplementary Data S2a). Non-PPI mutations were also common at the exposed amino acids on the IMM-arms of other RCs and CV (Supplementary Fig. 1a, Supplementary Data S1b-e, S2b-e).

We focused on candidate matrix-subunit Ndufa2 and IMM-subunit Ndufa1 to investigate CI-instability in presence of PPI and non-PPI mutations. These nuclear encoded subunits are present across eukaryotes, well resolved in CI-structures, and interact with least number of neighboring subunits in the matrix and IMM-arm²³. Thus, these subunits contain both PPI-interfaces and extended non-PPI surfaces. Mammalian Ndufa2 is a 99 amino acid protein with 2 α -helix and 4 β-strands and interacts with Ndufs1. 4431.8 Å² out of 5628.2 Å² surface area of mammalian Ndufa2 is solvent-exposed²³. K45 at the 1st helix and E57 just outside the 2^{nd} β -strand were found to be mutated in human diseases²⁴. These amino acids were within H-bond distance with G376 and Y664 of neighboring subunit Ndufs1. Disease-mutations K45T and E57A destabilized these PPIs (Fig. 1d). N-terminal IMM-helix containing subunit Ndufa1 interacts with ND1 within the IMM and with Ndufa8, Ndufa13, and ND6 at the IMS²³. With an exposed surface of 2849.8 Å², Ndufa1 is the only structurally conserved nuclear-encoded IMM-subunit with both PPI and non-PPI disease-related mutations (Supplementary Data S2a).

Three mutations are reported in Ndufa1²⁵. R37S at the IMSexposed C-terminal of Ndufa1 was found to destabilize intra-subunit Hbonding with D51 (Fig. 1d). G32 is also IMS exposed but close to IMMsurface, does not participate in PPI, and was found to expose the arginine side chain after mutation (Fig. 1d). The other mutating residue G8 is positioned at the N-terminus of the trans-IMM helix, not involved in any PPI but like G32 exposes a +vely charged side-chain when mutated to arginine in disease (Fig. 1d).

Knockout of Ndufa1 and Ndufa2 in HEK293T cells structurally and functionally destabilize CI. Knockout cells are OxPhos-inefficient, and proliferate poorly²⁶. We improvised an *in cellulo* substitution strategy to assay CI-instabilities due to the mutations in Ndufa1 and Ndufa2 in steady-state OxPhos-efficient cells²⁷ (Fig. 1e). We overexpressed EGFPtagged Ndufa1 and Ndufa2. Ndufa1 was found to colocalize with mitochondria immunostained by outer membrane protein Tom20, however, Ndufa2 was distributed throughout the cells with weaker colocalization (Fig. 1f, Supplementary Fig. 1b, Supplementary Table 1). Interestingly, western blots for mitochondrial fractions indicated that both the subunits entered mitochondria despite the poor mitochondria colocalization of Ndufa2 (Fig. 1g). in-gel EGFP fluorescence assays, western blots for endogenous CI-subunit Ndufs1 and CII-subunit Sdha on native PAGE, and in-gel CI-activity assays suggested that EGFPtagged Ndufa1 and Ndufa2 both competitively replaced the endogenous subunits in CI-assemblies without affecting stability and function of the complex (Fig. 1h-i, Supplementary Fig. 1c-d). Diseaserelated mutants of Ndufa1 also localized to mitochondria (Fig. 1f-g, Supplementary Fig. 1b, Supplementary Table 1). Similar to wild type, Ndufa2-K45T was poorly colocalized with mitochondria immunostained by outer membrane protein Tom20 but Ndufa2-E57A was better colocalized (Fig. 1f, Supplementary Fig. 1b, Supplementary Table 1).



Both the mutants were detected in mitochondrial fraction by western blot (Fig. 1g). Endogenous CI-assemblies and CI-function were not perturbed in the disease-related mutant expressing cells but unlike the wild type subunits, the mutants of both the subunits could not efficiently substitute the endogenous counterparts in *in-gel* fluorescence assays despite no perturbations of their predicted tertiary structures (Fig. 1h–i, Supplementary Fig. 1e). Thus, similar to PPI-mutants of Ndufa1 (R37S) and Ndufa2 (K45T and E57A), non-PPI mutants of Ndufa1 (G8R and G32R) were outcompeted by the endogenous subunits due to their structural incompatibilities within *in cellulo* CI-assemblies. We posit that while steric hindrance can explain destabilization of the CI-structure due to non-PPI mutations at the buried hydrophobic core of the PC, altered interactions with the IMM-lipids might cause CI-instability due to mutations at the non-PPI residues at the surface.

Fig. 1 | CI-deficiency disease-mutants populate surfaces of IMM-arm subunits. a Schematic representing structural framework of metazoan CI. 44 subunits associate to form 4 functional modules namely NADH dehydrogenase module (Nmodule; pink), ubiquinone reduction module (Q-module; blue), proton pumping proximal and distal modules (P_P and P_D modules; purple and black). Curved arrow: reaction catalyzed by N-module; straight arrows: flow of protons (H⁺). NAD⁺ Nicotinamide adenine dinucleotide; NADH - Nicotinamide adenine dinucleotide hydrogen. Nuclear encoded and mitochondria encoded subunits prefixed as Nduf and Mt. b Pie-chart indicating different types of mutations in CI-deficiency diseases including MELAS, Leigh and LHON syndrome as obtained from OMIM database (omim.org). Details in Supplementary Data S1a. c Left: Cartoon explaining buried and exposed amino acids. Involvement in PPI or otherwise (non-PPI) indicated. Amino acids with no neighboring atom in 3 Å^2 space considered as surface exposed. PPIs include H-bonds, salt bridges, cation-π interactions. Right: Stacked-bars indicating distribution of 79 missense CI-deficiency mutations on CI-structure (PDB: 5XTD). Details in Supplementary Data S2a. d Stick representation of PPI and

IMM-arm subunits are sequence-conserved within eukaryotic kingdoms but diverged beyond

Next, we investigated the conservation of disease-related residues across eukaryotes. Eukaryotic tree of life (eTOL) was classically divided into kingdom-level taxa i.e. Protista, Plantae, Fungi and Animalia (Metazoa)²⁸. With the advent of sequencing technologies, diverse species including protists are increasingly sequenced. Accordingly, eukaryotes were divided into 5 supergroups²⁹. Kingdom plantae was included in the supergroup Archaeplastidia while kingdom Fungi and metazoa were subclades of supergroup Opisthokonta. With increasing rise in phylogenomics, reshuffling and newer supergroups have been added making eTOL of 13 supergroups³⁰. Structure, function and composition of the RCs have been extensively studied in plants, Fungi and metazoa^{22,31-37} whereas limited knowledge is available from other taxa. Further, CI of *Tetrahymena thermophila* (supergroup TSAR) contains 25 unique accessory subunits with no homologs in plants, Fungi and metazoa^{16,38}. Euglena gracilis and Trypanosoma brucei (supergroup Discoba) likely to contain 14 and 19 unique accessory subunits in CI with no homologs in plants, Fungi and metazoa^{39,40}. Taken together, we limited our study to plants, Fungi and metazoa because of well characterized canonical CI with well resolved crvo-EM structures and for the ease of understanding we refer these as 'Kingdoms'^{20,41-43}.

Both the disease-causing PPI residues K45 and E57 on the CI matrix arm subunit Ndufa2 were conserved between metazoans, Fungi, and plants (Fig. 2a). Similarly, 20 out of the 29 residues that are involved in PPI-mutations on various subunits of CI matrix-arm were found to be conserved (Fig. 2b, Supplementary Data S2a and S2f). Nine disease-related but unconserved matrix-arm residues were also involved in PPI (Fig. 2b). Amino acids in the same positions in *A. thaliana* and *Y. lipolytica* CI-structures (PDB: 7ARB and 7071)^{31,32} were either absent or not resolved, non-interacting, or participating in different interactions suggesting subtle micro-arrangement of PPI in the CI matrix-arm in different kingdoms of eukaryotic life (Supplementary Fig. 2a).

PPI micro-arrangement was more prominent in the IMM-arm of CI as 3 out of 5 disease-related residues involved in PPI in human CI were unconserved (Fig. 2b, Supplementary Data S2a and S2f). Two of these were on the mito-encoded core subunits suggesting evolving pivotal interactions with neighboring nuclear-encoded subunits (Supplementary Data S2a). Intriguingly, divergence of disease-related PPIs was also observed in other RCs including CII, which lacks mito-genome encoded subunits (Supplementary Fig. 2b, Supplementary Data S2b-e and S2g-j). This suggested that evolution of PPIs at the IMM-arms of RCs might not be fully explained by increasing mutations rates in mito-encoded subunits.

Interestingly, 14 conserved non-PPI residues on the IMM-subunits were also mutated in CI-destabilizing diseases (Fig. 2b). These mutants

non-PPI amino acids in Ndufa2 and Ndufa1 as visualized by PyMol. Ndufa2 and Ndufa1: purple; interacting subunits: teal. Dashed black lines indicating PPI with bond length in Å. e Schematic representing workflow used for Figs. f-i. f Microscopy showing EGFP-tagged wild type (WT) and mutant subunits. Each image represents single Z-section of 0.3 µm. DAPI (blue): nucleus. Tom20 (red): mitochondrial marker. Scale bar: 10 µm. Colocalization coefficients for EGFP tagged subunits and mitochondria included in Supplementary Table 1. Experiments repeated thrice. g SDS-PAGE followed by immunoblots indicating WT and mutant proteins in mitochondrial fraction probed with EGFP antibody. Hsp60 as loading control. h in-gel EGFP fluorescence (IGF) assays of overexpressed wild type and mutant CI-subunits and western blots (WBs) of digitonin solubilised mitochondrial fractions of same probed for CI and CII using anti-Ndufs1 and anti-Sdha. SCsupercomplex: HC- holocomplex: SubC- subcomplex. Sdha*- HC: Sdha**- SubC. i ingel CI activity assays. Top: Reaction scheme. Bottom: BN-PAGE gels processed for CI activity. All uncropped gels and blots in Source Data. f-i: Data were verified in three independent experiments.

were found to expose charged or bulky side chains suggesting that physicochemical interactions between these conserved residues with IMM-microenvironment could be pivotal in stabilizing CI (Supplementary Fig. 2c, Supplementary Data S2a). Similarly, one IMS mutation G32R on the conserved glycine of Ndufa1 positioned in close proximity to the IMM-surface was found to expose +vely charged side chain of arginine (Figs. 1d and 2a). We also found 13 unconserved non-PPI mutations on various IMM-subunits (Fig. 2b). One of these, G8 at the trans-IMM helix of Ndufa1, was otherwise a buried residue in human CI but +vely charged side chain of arginine was exposed after mutation (Figs. 1d and 2a, c). In plants and Fungi, we observed a proline residue in the same position (Fig. 2a, c). Apart from this proline, multiple amino acids along the trans-IMM helix of Ndufa1 also diverged between plants, Fungi and metazoans in terms of charge and degree of hydrophobicity (Fig. 2c). Ndufa1 is neighbored by ND1 at the IMM and by Ndufa8, Ndufa13 and ND6 at IMS²². These neighboring subunits did not bury Ndufa1 entirely; rather the trans-IMM helix of Ndufa1 was surfaceexposed in metazoa, plant, or Fungi CI indicating a possibility of evolving interactions with IMM-lipids across eukaryotes (Fig. 2d).

Sequence divergence of Ndufa1 trans-IMM helix prompted us to analyze sequence evolution of other IMM-subunits of CI across eTOL. To do so, we have selected 134 representative organisms belonging to various clades of different taxonomic groups of eukaryotes i.e., plants, Fungi and metazoa as shown by Rebuead et al.44 (Supplementary Data S3). Sequences of matrix and IMS subunits were analyzed as control. Using blastp and human as query sequence, we identified sequence-homologs of all matrix-arm core subunits of human CI across all forms of eukaryotes (Fig. 2e, f, Supplementary Data S4a and S5a). Sequence identity with respect to (w.r.t) human was highest (~80%) within the metazoan homologs while plant and Fungi homologs showed ~60% sequence identity (Supplementary Fig. 2d). Sequence identity of 9 matrix-arm accessory, 5 IMM-arm core, and 3 P-module accessory subunits of CI showed high sequence identity of ~60% w.r.t human in metazoa (Fig. 2e, Supplementary Fig. 2d). In contrast, mean sequence identity of plants and Fungi homologs of the same subunits was reduced to ~40% w.r.t human (Fig. 2e, Supplementary Fig. 2e). No blast-hits were obtained for mitochondrial encoded core subunits ND4L and ND6 beyond metazoa when searched using human as query sequence.

Human CI contain 6 metazoan-specific accessory subunits distributed between N- and P-modules (Figs. 1a and 2e). Expectedly, when search against human query sequence, no homologs were identified for any of these beyond metazoans. Sequence homologs of 1 Q-module accessory subunit Ndufa7, 2 IMS, and 9 IMM-arm accessory subunits distributed between P_P and P_D -modules were also not identified among Fungi, plants, and unicellular eukaryotes by blastp using human as query sequence (Fig. 2e). We searched homologs of these subunits using sequences of Fungi *N. crassa* and plant *A. thaliana* that revealed



homologs with 60–80% sequence identity within the respective kingdoms (Fig. 2e, Supplementary Fig. 2e). Thus, among the IMM accessory subunits Ndufa13 was sequence conserved all across, 4 subunits were metazoan specific, and 9 subunits were uniquely sequence-conserved within eukaryotic kingdoms but extensively diverged beyond (Fig. 2e–f, Supplementary Data S4a and S5a). Interestingly, among Fungi, sequences for most of the subunits of *Y. lipolytica* were uniquely conserved with very few homologs even within the kingdom (Supplementary Data S4b and S5b) although the PDB structure (PDB: 7071) is nearly superimposable with human CI (PDB: 5XTD).

Similar to CI, IMM-subunits were sequenced-conserved within kingdoms and diverged beyond for all the RCs and CV (Supplementary Fig. 3, Supplementary Data S4c-f and S5c-f). Human CII is constituted by 4 core nuclear-encoded subunits³⁴ while plant CII contains

Fig. 2 | **IMM-arm subunits of CI are sequence-conserved within eukaryotic kingdoms. a** Multiple sequence alignment indicating conservation of amino acids of Ndufa2 and Ndufa1 mutated in human CI-deficiency diseases. Sequences of 13 representative organisms from metazoa, Fungi and plants aligned. Details in Supplementary Data S2f. **b** Stacked-bars representing conservation of 79 missense mutation involved in CI-deficiency diseases. Sequences of 13 representative organisms from metazoa, Fungi and plants indicated in (Fig. 2a) analyzed. Details in Supplementary Data S2a and S2f. **c** Conservation and hydrophobicity of trans-IMM helix of Ndufa1 represented. Color coded as per Goldman-Engelman-Steitz (GES) hydrophobicity scale indicating packing ability of trans-IMM helix into membrane. Black font: surface exposed amino acids; green font: buried amino acids. 1st and last amino acids indicated with number at top for each kingdom. *– identical;: – strongly similar;. – weakly-similar as per ClustalOmega convention. **d** Nudfa1 along with

8 subunits³³. Among the human subunits, Sdha and Sdhb are matrixexposed while Sdhc and Sdhd are IMM-embedded. Sdha and Sdhb were found to be sequence conserved across eTOL with mean identity of -70% w.r.t human. Sdhc was also found to be sequence conserved across eTOL but identity dropped to -52% w.r.t human. Homologs of Sdhd were found to be sequence-conserved within kingdom with mean identity of -60% w.r.t human, -41% w.r.t *Saccharomyces cerevisiae* and 63% w.r.t *Arabidopsis thaliana* in metazoa, Fungi and plants respectively (Supplementary Fig. 3).

Human CIII is an obligate dimer composed of 10 subunits³⁵. Three subunits are matrix-exposed, 1 is IMS exposed and rest 6 subunits are IMM-embedded out of which 3 are core subunits including the mitoencoded Mt-Cyb²². Matrix-arm and IMS-arm subunits were conserved across eTOL with mean identity of -55% w.r.t human. Core subunits of IMM arm were also conserved with -60% mean identity w.r.t human. Rest 3 nuclear-encoded IMM-subunits were well conserved within the eukaryotic kingdoms with mean identity of -60%, -50% and -65% in metazoa, Fungi and plants, respectively. Very few or no sequence-homologs were found beyond kingdoms for these 3 nuclear-encoded IMM-subunits (Supplementary Fig. 3).

Like CI, non-PPI mutations are also prominent on IMM-arm of CIV and were surface-exposed (Supplementary Fig. 1a) indicating the importance of lipid-protein interaction in stabilizing CIV. Three mitoencoded and 11 nuclear-encoded subunits constitute human CIV. Three nuclear-encoded accessory subunits are metazoan specific and IMM-embedded. Out of the rest 8 nuclear-encoded accessory subunits, Cox5B and opisthokont-specific Cox5A are matrix-exposed, Cox6B1 is IMS-subunit while others are IMM-embedded³⁶. Matrix-subunit Cox5B and IMS-subunit Cox6B1 were found to be sequence conserved across eTOL with mean identity of ~55% and 58% w.r.t human, respectively. Opisthokont-specific Cox5A was found to be sequence conserved in Fungi and metazoa with mean identity of ~63% w.r.t human. No homologs were detected for metazoan-specific subunits beyond metazoa. Five IMM-embedded accessory subunits were found to be sequence conserved within the kingdoms but poorly beyond. The mean identity for these subunits were ~60%, ~50% and ~65% within kingdom metazoa, Fungi, and plants (Supplementary Fig. 3). Thus, 80% of the accessory subunits of CIV are IMM-embedded and all of them showed kingdom-specific sequence conservation.

Similar to CI, human CV is composed of well-defined matrix and IMM arms³⁷. All F₁ subunits and 3 F_O subunits are matrix-exposed. Ten F_O subunits are IMM-embedded out of which 2 are mito-encoded, 2 are metazoan specific and 2 are Opisthokont-specific. All the F₁ subunits and IMM embedded core F_O subunits are sequence conserved across eukaryotes with mean identity of ~60% and 55% w.r.t human, respectively. IMM embedded metazoan specific subunits showed sequence conservation within metazoa with sequence identity of ~80% w.r.t human, respectively. Rest 4 nuclear-encoded IMM-embedded accessory subunits of human CV did not have sequence-homologs beyond metazoans while these were sequence conserved within Fungi and plants (Supplementary Fig. 3). Remarkably, 3 matrix-exposed subunits

neighboring subunits from CI-structures. *Homo sapiens* (PDB: 5XTD), *Yarrowia lipolytica* (PDB: 7O71), and *Arabidopsis thaliana* (PDB: 7ARB) representing metazoa, Fungi and plants. Red: surface exposed atoms and amino acids; yellow: partially exposed atoms and amino acids to surface; white: buried amino acids. Initial and end amino acids of IMM helix indicated by single letter code followed by position in sequence. Light gray: neighboring IMM subunit NDI; dark gray: neighboring IMS-subunit Ndufa8. **e** Heatmaps representing % sequence identity of CI-subunits searched using blastp against organisms mentioned. 134 organisms from eToL considered for analysis. Sequences of *Homo sapiens, Neurospora crassa and Arabidopsis thaliana* from metazoa, Fungi and plants used for blast search. Gray box indicate no blastp hits. M - matrix; I – IMM; I* – IMS. Details in Supplementary Data S4a. **f** Schematic representing mean % identity of CI-subunits calculated from Fig. 2e. Details in Supplementary Data S4a.

Atp5pb, pd, and Opisthokont-specific pf were similarly sequenceconserved within the eukaryotic kingdoms only i.e., within metazoa, Fungi and plants with identity of -55% w.r.t human, -44% w.r.t *Saccharomyces cerevisiae* and -60% w.r.t *Arabidopsis thaliana*, respectively. These 3 subunits constitute the F_O peripheral stalk as a connecting bridge between F₁ and F_O³⁷.

We also analyzed the subunit-sequences of mitochondrial outer membrane translocase Tom40 which switches between dimers and trimers⁴⁵. Seven proteins constitute the complex, which are either fully or partially OMM-embedded. All these proteins were sequenceconserved within the eukaryotic kingdoms. Sequence homologs were either not found by blastp search beyond the kingdoms or if found, sequence homology was below 30% (Supplementary Fig. 4, Supplementary Data S4g and S5g). Both CII and Tom40 complex are devoid of mito-encoded subunits but still showed consistent kingdomspecific sequence conservation. Thus, we hypothesized that in addition to the diverging mito-genome, sequence evolution of membranesubunits of mitochondrial PCs may be driven by membrane lipids.

Trans-IMM helices of Human and Arabidopsis Ndufa1 interact differently with IMM-lipids

To consolidate the hypothesis of lipid-driven sequence evolution of IMM-subunits of RCs, we performed amino acid wise conservation analysis for the lipid-exposed regions of 8 single-pass IMM-arm accessory subunits of CI each containing a 25-residue continuous stretch with more than 70% exposed amino acids (Supplementary Data S6a). These 8 accessory subunits were present in all 3 kingdoms of eukaryotic life and were well resolved in available cryo-EM structures^{22,31,32}. As control, we also analyzed similar matrix-exposed stretches of 8 matrix-arm subunits (Supplementary Data S6b). Figure 3a shows % identity for each amino acids of the matrix-exposed stretch of Ndufa2 and the exposed trans-IMM helix of Ndufa1 with mean identity of 40.92% and 45.64% w.r.t human across eukaryotes, respectively. Kingdom-specific sequence identity for matrix-subunit Ndufa2 was restricted to 48-56% within metazoa (w.r.t human) and plants (w.r.t Arabidopsis thaliana). In contrast, exposed trans-IMM helix of Ndufa1 showed 63.64% sequence identity within metazoa (w.r.t human) and 67.52% within plants (w.r.t Arabidopsis thaliana). Consistently, exposed surfaces of all the 8 IMM-arm subunits showed significant increase in mean sequence identities within kingdoms (Fig. 3b). This suggested that the exposed surfaces of matrix-arm CIsubunits could have evolved neutrally while exposed surfaces of IMMsubunits were under kingdom-specific selection pressure possibly due to interactions with IMM-lipids.

Cardiolipin (CL) is a unique lipid present only in mitochondria membranes in eukaryotes and a major membrane constituent in bacteria⁴⁶. CL has been implicated in the biogenesis of OMM import channel Tom40⁴⁷. Cryo-EM structures of CI, CIII, and CIV from various species have identified IMM-lipids including CL associated with the complexes^{22,31,32,36,48}. Simulation of CIII and CIV in model membranes identified CL to be the major interacting lipid^{49,50}. Interestingly, mol%



composition of IMM-lipids was found to be similar between plants and metazoa (Fig. 3c, Supplementary Data S7a) but the fatty acyl chain length and unsaturation were uniquely kingdom-specific (Figs. 3d, Supplementary Data S7b). CL is reported to be more unsaturated in plants (18:3 and 18:2, 55:38) than metazoa (18:2 and 18:1, 49:20) while 20:4 and 22:6 phosphatidylcholine (PC), phosphatidylethanolamine (PE), and phosphatidylinositol (PI) are reported in metazoa only^{18,51}

(Fig. 3c-d, Supplementary Data S7b). Also, sphingolipids, cholesterol, and phosphatidylserine (PS) are either not reported or present in very low abundance in plant mitochondria membranes⁵² (Fig. 3c and Supplementary Data S7a).

We simulated a model IMM mimicking metazoan lipid composition (Model 1, Fig. 3e). Variation of lipids between IMM-leaflets in metazoa was considered with PI, PE, and CL rich on matrix and PC rich

Fig. 3 Trans-IMM helix of Human and Arabidopsis Ndufa1 interacts differently
with IMM-lipids. a % identity of individual amino acids at structurally similar
exposed stretches of Ndufa2 and Ndufa1. Homo sapiens sequences used as refer-
ence for multiple sequence alignment across eukaryotes and within metazoa;
Arabidopsis thaliana sequences for plants. Details in Supplementary Data S6a-b. ${\bf b}\%$
identity of individual surface-exposed amino acids of matrix and IMM arm subunits
plotted. Whisker: 1 SD from mean. P-values: two-sided, two sample Student's t-test.
n = 8 subunits each from matrix, membrane arm. Details in Supplementary
Data S6a-b. c Bar-graph representing mol% of different lipids in IMM. Black dia-
mond: Data points and outleirs. Error bars: 1 SD from mean. $n = 9$ (metazoa) and 3
(plants). Details in Supplementary Data S7a. d Bubble graph showing mol% of dif-
ferent fatty acyl chains in IMM. $n = 4$ (metazoa) and 1 (plants). Details in Supple-
mentary Data S7b. c – d <i>n</i> : datasets obtained from literature. e Graph showing mol%
of various lipids used for simulating model 1, model 2. Details in Supplementary
Table 2. f Violin plot indicating area per lipid (Ų) calculated using final 500 ns of $1\mu s$

on IMS leaflets (Fig. 3e, Supplementary Data S7c). Low abundant IMMlipids such as sphingomyelin, sterols, lysophosphatidylcholine, lysophosphatidylethanolamine, phosphatidylglycerol, etc. were not included for simulation. Fatty acids 16:0, 18:0, 18:1, 18:2, and 20:4 were simulated as per their abundance in metazoan IMM. We also simulated a Model 2 where we introduced plant-specific unsaturation 18:3 at sn1 position of 1st phosphate of otherwise 18:2 CL, which resulted in 18:3-18:2, 18:2, and 16:0 CL in 54:40:6 ratio. Other lipids in Model 2 remained identical to Model 1 (Fig. 3e, Supplementary Data S8a). Total 256 lipid molecules were placed in each leaflet and all-atom Molecular Dynamics simulations were performed for 1 µs (Supplementary Table 2).

Time-averaged data calculated over the last 500 ns of microsecond-long all-atom simulation trajectories suggested that replacing just one acyl chain from 18:2 to 18:3 in CL in Model 2 resulted in alteration in membrane structural properties. Area per lipid (APL) in Model 2 was increased to 75.6 Å² from 73.8 Å² in Model 1 (Fig. 3f, Supplementary Fig. 5a). The mean bilayer thickness of Model 1 was 41.2 Å while that of Model 2 was 40.2 Å (Fig. 3g, Supplementary Fig. 5b). We then simulated trans-IMM helices of human (Hs) and Arabidopsis (At) Ndufa1 (Hs-A1 and At-A1) within the IMM-models. IMM-exposed helix of Hs-A1 was constituted by N-terminal M1-H27 while the corresponding At-A1 helix was from V4-H30 (Fig. 2c). Y25-H27 in Hs-A1 was not included for simulation as the stretch was found to be identical as Y28-H30 in At-A1 (Fig. 2c). Y27 in At-A1 helix corresponded to A24 in Hs-A1. Y27 was found to be involved in sulfur- π interaction with M98 of ND1 in At-A1 (PDB: 7ARB) but A24 was non-interactive as per the human CI (PDB: 5XTD) (Supplementary Fig. 5c). We excluded At-A1(Y27) and Hs-A1(A24) for simulation. Thus, 1st 23 amino acids of Hs-A1 (M1-T23) and the corresponding (V4-Q26) in At-A1 were not involved in PPIs, were sequence diverged between kingdoms, and therefore considered for simulation. We placed single trans-IMM helices of each in the model membranes. Both the trans-IMM helices lied within the membrane normal in both model membranes indicating that they were accommodated within the membrane hydrophobic core rather than interacting with the head groups of the lipids (Supplementary Fig. 5d). The mean APL remained unaltered when Hs-A1 trans-IMM helix (M1-T23) was placed in Model 1 while that slightly decreased to 73.3 Å² with At-A1 (V4-Q26) (Fig. 3f, Supplementary Fig. 5a). The mean bilayer thickness of Model 1 was increased to 41.6 Å for the Hs-A1 trans-IMM helix and 42.03 Å for At-A1 (Fig. 3g, Supplementary Fig. 5b). Conversely, mean APL remained unaltered when At-A1 trans-IMM helix was placed within Model 2 while Hs-A1 trans-IMM helix marginally increased the APL to 76.1 Å² (Fig. 3f, Supplementary Fig. 5a). The mean bilayer thickness of Model 2 remained unaltered when Hs-A1 trans-IMM helix was placed while At-A1 trans-IMM helix increased the mean bilayer thickness to 40.7 Å (Fig. 3g, Supplementary Fig. 5b).

Number of contacts between the peptides and the acyl chains in the model membranes are calculated considering contacts between the non-hydrogen atoms of these groups within a cut-off distance of 5 simulation time. **g** Violin plot indicating bilayer thickness (Å) calculated using final 500 ns of 1 μ s simulation time. **f**-**g** w/o- without trans-IMM helix (TIH); At- At-A1 TIH; Hs- Hs-A1 TIH. Circles: Mean; Error bars: 1 SD from mean. Details in Source Data. **h** Graph showing time-averaged total acyl chain contacts of different lipid species of matrix and IMS leaflets with Hs-A1 and At-A1 TIHs i.e., Hs-A1 (M1-M12 residues), At-A1 (V4-II5 residues) and Hs-A1 (G13-T23 residues), At-A1 (G16-Q26 residues). Total contacts averaged over final -500 ns of 1 μ s simulation time (mean). Error bar: 1 SD from mean. Details in Supplementary Data S8b and Source Data. **i** Schematic summarizing interactions of acyl chains of different lipids with Hs-A1 and At-A1 trans-IMM helix. Data obtained from Fig. 3h and supplementary fig. 5e. 1 lipid molecule- 50-110 contacts; 2 lipids- 110-150 contacts; 3 lipids- 150 and above contacts. **e**-**i** Model 1 (metazoan mimic): PC, PE, CL, PS, and PI composition of metazoan IMM; Model 2 (plant mimic): PC, PE, PS, and PI from metazoan IMM with plant CL18:3.

Å⁵³. We considered N-terminal 1-12 amino acids of the trans-IMM helices as matrix leaflet interacting while the rest was considered to interact with IMS-leaflet lipids. Figure 3h and Supplementary Fig. 5e indicates total contacts of the acyl chains with trans-IMM helices averaged over the final 500 ns of simulation time. At-A1 (V4-I15) in Model 1 established maximum interactions with PE (260.7 contacts) of which 160.5 and 63.7 interactions with 20:4 and 18:0 PE at the matrix leaflet, respectively. Matrix leaflet interactions with CL was found to be 44.3 amongst which 41.5 interactions were with unsaturated CL18:2. On the other hand, Hs-A1 (M1-M12) established maximum contacts with CL (196.8 contacts) amongst which 181.2 was with unsaturated 18:2. Further, interaction of Hs-A1(M1-M12) with 20:4 PE at the matrix leaflet was only 61.4 in contrast to 160.5 of At-A1(V4-I15) suggesting that both the trans-IMM helices mostly interacted with unsaturated acyl chains at the matrix leaflet of Model 1 but the interacting lipids differed (Fig. 3h-i, Supplementary Fig. 5e, Supplementary Data S8b).

In Model 2, At-A1 (V4-I15) established 146 interactions with CL at the matrix leaflet with 75.6 and 65.3 contacts with 18:3-18:2 and 18:2 CL. 20:4 PE interactions were reduced to 42.8 contacts compared to Model 1. Interestingly, Hs-A1 (M1-M12) gained interaction with PE (142.6) in Model 2 as compared to Model 1 (109.1) and CL interactions in Model 2 were reduced to 132.5 as compared to 196.8 in Model 1. However, even in Model 2, Hs-A1 (M1-M12) preferred 18:2 CL (86.4) as compared to 18:3-18:2 CL (45.3). Thus, At-A1 (V4-I15) interacted more with unsaturated PE in Model 1 in absence of 18:3-18:2 CL at the matrix leaflet whereas the Hs-A1 (M1-M12) interacted more with PE as compared to CL in presence of 18:3-18:2 CL in Model 2 (Fig. 3h, Supplementary Fig. 5e, Supplementary Data S8b). At the matrix leaflet both the trans-IMM helices interacted with PC as well, however, Hs-A1 (M1-M12) interacted with PC more in Model 1 (118.4 contacts) as compared to Model 2 (81.9 contacts). At-A1 (V4-I15) established similar number of contacts with PC in both the models (Fig. 3h, Supplementary Fig. 5e, Supplementary Data S8b).

Both the trans-IMM helices interacted with PC, CL and PE at the IMS leaflet. At-A1 (G16-Q26) interacted with unsaturated PC (18:2 and 20:4) and PE 20:4 (51.7, 99.1 and 57.5 contacts, respectively) at the IMS leaflet of Model 1. In contrast, the same trans-IMM helix interacted with saturated 16:0 PC, 18:0 PC, 18:0 PE and unsaturated 18:2 CL (41.5, 64.6, 58.8 and 32.5 contacts, respectively) in Model 2. On the other hand, Hs-A1 (G13-T23) interacted with 16:0 PC, 18:0 PC, 18:0 PC, 20:4 PC, 18:1 CL and 18:2 CL (47.2, 33.3, 34.3, 38.3, 44.5 contacts, respectively) at the IMS leaflet of Model 1 in contrast to Model 2 where it gained interactions with unsaturated PC 18:1 (38.8), 18:2 (57.1), and 20:4 (58.4), respectively (Fig. 3h, Supplementary Fig. 5e, Supplementary Data S8b). Interaction of both the trans-IMM helices with PI and PS were minimal in both the models with exception where At-A1 (G16-Q26) made 61.5 contacts with PS in Model 2.

In summary, structurally superimposable trans-IMM helices of Ndufa1 from two different kingdoms of eukaryotic life showed

contrasting but complementary interactions with fatty acyl chains of IMM-lipids depending on the 18:2 or 18:3 chemistry of CL (Fig. 3i).

Lipid-exposed amino acids prevent exchange of Hs-A1 trans-IMM helix with At-A1 trans-IMM helix in cell culture

MD simulations depicted that At-A1 helix preferred 18:3 CL over 18:2. Mammalian cells lack 18:3 CL. Therefore, we hypothesized that At-A1 helix might not stably interact with IMM-lipids in HEK293T cells and thereby might be incompatible to integrate into CI. To validate, we performed in cellulo assays using EGFP-tagged constructs of Ndufa1 containing either Hs or At helix. We replaced the non-PPI, trans-IMM helix part of Hs-A1 (M1-T23) with 4th-26th amino acids of At-A1 where V4 was mutated to methionine (Chimera-A1; Ch-A1). As per AlphaFold, the structures are superimposable and most of the PPIs between Hs-A1 and its interacting partners ND1 and Ndufa8 remained unaltered (Fig. 4a-b, Supplementary Fig. 6a). A gain of interaction between N21 of Ch-A1 and T256 of human ND1 was also observed (Supplementary Fig. 6a). We prepared stable inducible lines for Hs-A1, At-A1, and Ch-A1. All these proteins colocalized with Tom20 stained mitochondria (Fig. 4c, Supplementary Fig. 6b). Intriguingly, western blots of mitochondrial fractions revealed excess load of At-A1 and Ch-A1 as highly abundant cleaved products ~3-kDa lighter in mass than the full-length proteins (Fig. 4d). We also prepared mitoplasts devoid of the mitochondrial outer membrane protein Tom40 but containing IMM and matrix proteins Ndufa13 and Hsp60, respectively (Supplementary Fig. 6c). A prominent band for Hs-A1 G8R mutant was observed in the mitoplast fraction (Fig. 4e) indicating successful import of the mutant into IMM while it was unable to integrate into CI (Fig. 1h). Interestingly, fulllength At-A1 band was almost absent and Ch-A1 band was comparatively less intense in the mitoplast fractions than Hs-A1 and Hs-A1G8R (Fig. 4e). The cleaved products of At-A1 and Ch-A1 were still observed in the mitoplasts indicating that the breakdown of the full-length proteins could be an IMM-associated event. We created a construct deleting the trans-IMM helix in human Ndufa1 (Δ 27Hs-A1). The cleaved product of Ch-A1 run in the same position with Δ 27Hs-A1 in SDS-PAGE suggesting that cleavage might have happened after the trans-IMM helix (Fig. 4f). Local chemistry driven cleavage of peptide bonds after N and Q have been reported for multiple proteins⁵⁴⁻⁵⁶. At-A1 trans-IMM helix present in Ch-A1 contains both N and Q at the C-terminus (Fig. 2c). The cleaved band of Ch-A1 disappeared after introducing L in place of N21 (Fig. 4g). Interestingly, replacing Q23 with Hs-A1 analogous T23 did not prevent the cleavage of Ch-A1 (Fig. 4g). These results indicated that N containing trans-IMM helices of At-A1 and Ch-A1 interact inappropriately with human IMM and therefore unstable.

Next, we tested if At-A1 trans-IMM helix was compatible to human CI. in-gel EGFP fluorescence assays using digitonin-solubilized mitochondrial extracts from Hs-A1, At-A1, and Ch-A1 cells indicated successful incorporation of Hs-A1 into CI-assemblies while both At-A1 and Ch-A1 did not integrate (Fig. 4h). However, two bands at ~450 and ~180 kDa was observed in At-A1 lane indicating the formation of aberrant subassemblies (Fig. 4h). Endogenous RC-assemblies were not destabilized in these cells as evident by the blots with Ndufs1 and Sdha antibodies (Fig. 4h, Supplementary Fig. 6d). CI activity was also not perturbed as per in-gel CI activity (Supplementary Fig. 6d). Interestingly, we also found mismatch of five amino acids between the trans-IMM helices of human and mouse Ndufa1 (Fig. 4i). Therefore, we switched the first 23 amino acids of Hs-A1 with the mouse sequence (Ch-A1^{Mm-switch}). The chimeric product successfully integrated into human CI indicating that intra-kingdom amino acid mismatches did not perturb LPIs of Ndufa1 (Fig. 4i, Supplementary Fig. 6e). This is probably due to the consistency of physicochemical properties of the mismatched amino acids.

Detailed analysis of PDB: 5XTD indicated that side chains of 6 out of 23 amino acids constituting Hs-A1 trans-IMM helix contained no atoms within 3 Å² of neighboring space i.e., side chains of these amino acids were exposed to IMM-lipids (Figs. 2d and 5a). Among these, matrix leaflet residues F3 and P7 were conserved between human and mouse helix (Fig. 4i). As per simulation, F3 of Hs-A1 interacted with both PC and CL tails while P7 was more interactive with CL fatty acyl chains in model IMM 1. Interestingly, corresponding L6 and L10 of At-A1 interacted with PE chains (Fig. 5b). IMS-leaflet nonpolar L21 was physicochemically conserved as V21 in mouse but found to be polar as N24 in At-A1 (Fig. 4i). However, A22 in Hs-A1 remained physicochemically conserved as S22 and S25 in mouse and Arabidopsis helices, respectively. Simulation results indicated that Hs-A1 L21 and A22 interacted with PC and CL while corresponding N24 and S25 of At-A1 interacted with mostly PC and PE chains (Fig. 5b).

Interestingly, when we replaced L21 and A22 with corresponding N and S from At-A1 trans-IMM helix, the resulting protein Hs-A1^{LA/NS} was cleaved and imported less into mitochondria as evident in the western blot of the mitochondrial fraction (Fig. 5c). However, it was efficiently integrated into human CI as evident by in-gel EGFP fluorescence (Fig. 5d, Supplementary Fig. 6f). However, we did not observe any cleavage of Hs-A1^{FP/LL} where F3 and P7 were replaced with corresponding Leucine residues from At-A1 but the protein was not observed in productive CI-assemblies (Fig. 5c, e). Instead, we observed a smear as well as a prominent but aberrant subcomplex at ~300 kDa and faint subcomplexes at ~600 kDa and 1048 kDa (Fig. 5c, e, Supplementary Fig. 6g). Similarly, switching all the side chain exposed amino acids in Hs-A1 with the analogous At-A1 amino acids (Hs-A1^{Ex-switch}) resulted in absolute disappearance of SC and HC bands in the in-gel fluorescence assay while two bands at ~320 and ~180 kDa indicated formation of aberrant subassemblies (Fig. 5f-g, Supplementary Fig. 6h).

Plant-specific unsaturated lipids trigger CI instability in human cells

Finally, we tested if human CI could withstand engineered changes in unsaturation of fatty acyl chains in IMM-lipids. Recently, Oemer et al. demonstrated modulation of molar composition and unsaturation of various lipids including CL in mammalian cells treated with different fatty acids^{57,58}. Increased incorporation of 18:2 fatty acyl chains in CL has also been linked to efficient OxPhos⁵⁸ while increase in 18:3 fatty acid perturbs mitochondrial respiration^{59,60}. We treated Hs-A1 cells with BSA conjugated 18:2 linoleic acid and 18:3 alpha-linolenic acid and performed mass spectrometry of the mitochondrial CLs (Fig. 6a). CL species 64:4, 66:1, 72:4, 66:5, 70:5 with 16:1, 18:1 and 18:2 fatty acids were rich in BSA-treated cells (Fig. 6a) suggesting their natural abundance in HEK293T mitochondria. Treatment with linoleic acid resulted in replacement of at least one of the fatty acyl chains with 18:2 resulting in CL 70:6, 70:7, 72:7, 72:8 etc. CL 72:8 with all four 18:2 fatty acyl chains were highly abundant in these cells (Fig. 6a). Alpha-linolenic acid treatment resulted in multiple unique CLs with 18:3 fatty acyl chains that were absent in both BSA and linoleic acid treated cells while CLs with 16:1, 18:1 and 18:2 fatty acyl chains were prominently reduced (Fig. 6a). Microscopy indicated fragmented mitochondria correlating with the degree of unsaturation of CLs (Supplementary Fig. 7a). Simultaneously, in-gel EGFP fluorescence assays revealed efficient entry of Hs-A1 in CI-HC and SCs in BSA-treated and 18:2 treated cells but only partial entry was observed in 18:3 treated cells (Fig. 6b). Interestingly, the cleaved product of Ch-A1 was reduced when cells were treated with 18:2 and 18:3 fatty acids suggesting slightly better stability of the trans-IMM helix in presence of unsaturated lipids (Fig. 6c). However, Ch-A1 was never integrated with CI-assemblies as probed by in-gel EGFP fluorescence assay (Supplementary Fig. 7c) correlating with perturbation of endogenous assemblies of CI in 18:3 fatty acid treated cells as probed by Ndufs1 western blots and in-gel Clactivity assay (Figs. 6d-e and Supplementary Fig. 7c). Together, these results suggested that the chemistry of the exposed amino acids in the



trans-IMM helix of Ndufa1 with the surrounding saturated or unsaturated lipids determined the integration of the subunit in CI assemblies in different forms of eukaryotic life (Fig. 7).

Discussion

Intra- and inter-protein interactions are established determinants of protein sequence evolution. Local chemistry with other interacting

macromolecules also impose sequence divergence in proteins. Rapid evolution of mt-rRNA due to mito-genome instability is correlated with sequence-divergence of mito-ribosome and respiratory complex (RC) subunits¹³. Alteration of lipid chemistry also calibrate lipid-protein interactions⁶⁻¹⁰ but lipid-influenced sequence divergence of proteins and protein complexes across eukaryotic evolution is not investigated extensively. Rather, it is proposed that the sequences of membrane Fig. 4 | Incompatibility of A. thaliana trans-IMM helix in human cells. a Cartoon showing Ndufa1 and ND1 of human (PDB: 5XTD), A. thaliana (PDB: 7ARB) and Chimera (AlphaFold modelling) as visualized by PyMol. Chimera-Ndufa1 (Ch-A1) A. thaliana Ndufa1 (At-A1) trans-IMM helix in Human Ndufa1 (Hs-A1). b Superimposed images of Hs-A1, At-A1, and Ch-A1. Hs-ND1: light gray; At-ND1: dark gray. RMSD indicated out of 388 aligned amino acids. (a-b) Black dotted lines: trans-IMM regions. c Microscopy of cells expressing EGFP-tagged Hs-A1, At-A1, and Ch-A1. Single Z-section of 0.3 µm. DAPI: nucleus. Tom20: mitochondrial marker. Scale bar: 10 µm. Colocalization coefficients in Supplementary Table 1. d Immunoblots of EGFP-tagged Hs-A1, At-A1, and Ch-A1 expressing cells probed with EGFPantibody in mitochondrial fraction. Hsp60 as loading control. e Immunoblots indicating Ndufal-EGFP in mitochondrial fractions and mitoplast fractions of Hs-A1. At-A1. Ch-A1 and A1(G8R) expressing cells probed with EGFP-antibody. Hsp60 as loading control. Tom40 - Outer mitochondrial membrane marker. f Immunoblot indicating Ndufa1-EGFP in total cell lysates of Hs-A1, Ch-A1, and Δ27Hs-A1-EGFP expressing cells, probed with EGFP-antibody. Δ27Hs-A1-EGFP was treated with

proteins, especially the transmembrane regions, are highly constrained by membrane lipids and therefore diverge less than other cellular proteins^{61,62}. Our results indicate that membrane-proteins of mitochondrial complexes are highly sequence conserved within eukaryotic kingdoms but undergo divergent evolution across eukaryotes correlating with the evolving chemistry of surrounding lipids. In case of RCs, we posit that pivotal PPIs are evolutionarily conserved for the structural maintenance of both membrane-embedded and extramembrane arms of these complexes. Conversely, divergence of lipidprotein chemistry at the IMM-arm is required to preserve the nearly super-imposable structural designs for optimized OxPhos across eukaryotes. We further hypothesize that higher mutation rates in the mito-genome provide an indirect sampling advantage to protein sequences of IMM-subunits in adapting divergent lipid environments. These hypotheses are in line with previous postulates which suggest that lipid biosynthesis is probably a later step after finalizing the conformations of fundamental protein-machines as exemplified by structural conservation of ATP synthase in two distinctly divergent lipid environments of bacteria and archaea⁶³.

Mutations affecting fundamental interactions destabilize protein complexes and may result in diseases. CI-deficiency disease mutations are present on both the matrix and IMM-arms. Missense PPI-mutations at the matrix arm result in reduced CI-assembly and activity leading to early infantile onset diseases (Supplementary Data S9a). Disease-related LPI-mutations are reported largely in mito-encoded IMM-subunits of CI with only three in nuclear-encoded Ndufs8, Ndufa1 and Ndufc2 (Supplementary Data S9a). Disease onset for LPI-mutations in mito-encoded CI-subunits ranges from infantile to adulthood suggesting that heteroplasmy of these mutations may partially safeguard disease manifestations. Ndufa1 and Ndufc2 mutations always result in early-onset diseases suggesting comparatively severe phenotypic manifestation of the altered LPIs (Supplementary Data S9a). Interestingly, LPI-mutations are not reported for nuclear-encoded IMM-subunits of CIV and may indicate lethal phenotypes further substantiating stringency of LPIs in stabilizing RC-structures (Supplementary Data S9b).

Our elaboration of lipid-protein chemistry driven sequence evolution in transmembrane proteins is based on 18:2/18:3 chemistry of mitochondria specific lipid cardiolipin (CL). CL-homeostasis has been implicated in several degenerative conditions including Barth syndrome, cancer, neurodegenerative diseases, non-alcoholic fatty liver disease, and heart failure⁶⁴⁻⁶⁶. Unsaturated CL18:3 is so far reported in plants while 18:2 unsaturation is prevalent in all tissue types of metazoans. However, variation of molar composition of CL 18:2 is reported between tissues within the same organism and even within the same tissue at different metabolic states (Supplementary Data S7a-c)^{18,51}. Our data of instability of human RCs due to further unsaturated CL 18:3 suggest that amino acid composition of trans-IMM helices of RC-subunits are more sensitive to degree of CL-unsaturation 2.5 μM MG132 for 24 h. β-actin as loading control. * Cleaved bands of A1-EGFP and full-length Δ27Hs-A1-EGFP. **g** Immunoblot indicating cleaved product of Ndufa1-EGFP in total cell lysates of Ch-A1, Ch-A1-N21L, and Ch-A1-Q23T expressing cells probed with EGFP-antibody. β-actin as loading control. **d-e**, **g** * Cleaved band of A1-EGFP. **h** *in-gel* EGFP fluorescence (IGF) and western blot (WB) of digitonin solubilised mitochondrial fractions on native-PAGE from cells shown in **4 d**. WB sequentially probed with anti-Ndufs1 (Supplementary Fig. 6d) and anti-Sdha. **i Top**: Pair-wise sequence alignment of trans-IMM helix of Hs-A1 and Mm-A1. *- identical;: - strongly similar; as per Clustal Omega convention. 1st and 23rd amino acids represented with number at top. Red font: mismatched amino acids in Mm-A1. **Bottom**: IGF and WB of digitonin solubilised mitochondrial fractions on native-PAGE from Hs-A1-EGFP and Ch -A1^{Mm-switch}-EGFP. WB sequentially probed with anti-Ndufs1 (Supplementary Fig. 6e) and anti-Sdha. **h-i** SC - supercomplex; HC - holocomplex; SubC - subcomplex. Sdha*- HC; Sdha**- SubC. All uncropped gels and blots in Source Data. **c-i** Data verified in three independent experiments.

than molar composition. However, the diversity of molar composition of 18:2 CL and other lipids may be implicated in the heterogeneous assembly of RCs in different supercomplexes among tissues^{51,67,68} and may partially explain the tissue-specific defects in functional assembly of RCs in human OxPhos-deficiency diseases⁶⁹. The balance of saturated and unsaturated phospholipids regulates curvature, fluidity, protein stability, and crowding in all cellular membranes^{6,70}. Plant lipid metabolism has been evolved to favor polyunsaturated fatty acids (PUFA) which provide membrane plasticity to cope up with diverse habitats⁷¹. PUFA is also prevalent in aqueous metazoans to maintain membrane fluidity⁷². Accordingly, investigating evolution of lipid unsaturation in non-mitochondrial membranes and correlated sequence divergence in membrane-embedded proteins would be interesting.

Our results suggest the trans-IMM domain of Arabidopsis Ndufa1 is highly unstable and is cleaved after C-terminal asparagine (N) at IMM in human cells. IMM-targeting of proteins is mediated by dedicated import complexes TIM23 and TIM22. These complexes are not waterfilled import channels rather their concave surfaces bind and simultaneously expose the clients to IMM-lipids⁷³⁻⁷⁵. IMM-sorting proteins like Mgr2 or Oxa1 may transiently safeguard the clients from the IMMlipids after which the clients are either laterally released or conservatively sorted into IMM^{76,77}. In both cases, it is imperative that composition of IMM-lipids may directly influence the successful insertion of the clients into IMM. Lipid-assisted insertion mechanism has already been proposed for tail-anchored or signal-anchored proteins at the OMM^{78,79}. Our results with Ndufa1 supports similar lipidassisted mechanism for IMM-proteins. The mechanism of cleavage of the trans-IMM domain of Arabidopsis Ndufa1 in human cells is currently unknown. Membrane-resident serine proteases are known to cleave bonds between N and S⁸⁰. The C-terminal of At-A1 helix contain consecutive N and S and thus may expose a cleavage sequence for human IMM-proteases. Further, there could be impaired binding of the At-A1 helix with the human IMM-sorting proteins Mgr2 or Oxa1 or equivalent import-partners forcing a cleavage. Surface area also influences stable insertion of transmembrane helices in saturated or unsaturated lipid environments⁸¹. It is also conceivable that the lipidprotein incompatibility may pose significant thermodynamic challenge so that the peptide bond at N is rather cleaved to avoid the energy-expenditure to insert At-A1 helix into human IMM. Addressing these questions will elucidate some thermodynamic puzzles of lipidprotein chemistry.

Methods

Analysis of disease-related amino acids in RC-subunits

Online Mendelian Inheritance in Man (OMIM) database (http://omim. org/) was searched for known disease-related mutations in Respiratory Complex (RC) genes. Missense mutations associated with RC-



Fig. 5 | **Exchange of lipid-exposed amino acids of Hs-A1 trans-IMM helix with At-A1 trans-IMM helix triggers** *in-cellulo* incompatibility. **a** Cartoons showing Ndufa1 and ND1 when visualized by PyMol. **Left:** Human Ndufa1 (blue) and ND1 of CI (PDB: 5XTD). **Middle:** Human ND1 and Hs-A1^{Exswitch} (blue-green) where exposed amino acids of Hs-A1 replaced with corresponding amino acids of At-A1 (AlphaFold modeling). **Right:** Superimposed image. Hs-A1: blue; Hs-A1^{Exswitch}: teal; Hs-ND1: light gray. RMSD indicated out of 388 aligned amino acids. Template modeling score > 0.9. **b** Bar-graph showing total contacts of individual amino acids of Hs-A1 and At-A1 trans-IMM helices with the acyl chains of different lipid species in Model 1 (metazoan mimic) simulation experiment. Total contacts averaged over final -500 ns of 1 μs simulation time (mean). Error bar: 1SD from mean. Model 1 (metazoan mimic): PC, PE, CL, PS, and PI composition of metazoan IMM. Details in Source Data. **c** Immunoblot indicating Ndufa1-EGFP in mitochondrial fractions of Hs-A1, Ch-A1, Hs-A1^{EP/LL}, Hs-A1^{LA/NS} probed with EGFP antibody. Hsp60 as loading control. * Cleaved band of A1-EGFP. **d** Native-PAGE processed for *in-gel* EGFP fluorescence (IGF) and western blot (WB) using digitonin solubilised mitochondrial fractions from Hs-A1, Ch-A1, and Hs-A1^{LANS}. WB sequentially probed with Ndufs1 (Supplementary Fig. 6f) and Sdha antibodies. **e** Native-PAGE processed for *in-gel* EGFP fluorescence (IGF) and western blot (WB) using digitonin solubilised mitochondrial fractions from Hs-A1, Ch-A1, and Hs-A1^{FP/LL}. WB sequentially probed with Ndufs1 (Supplementary Fig. 6g) and Sdha antibodies. **f** Immunoblot indicating Ndufa1-EGFP in mitochondrial fractions of Hs-A1, Ch-A1, and Hs-A1^{Exwitch} expressing cells with EGFP antibody. *Cleaved band of A1-EGFP. **g** Native-PAGE processed for *in-gel* EGFP fluorescence (IGF) and western blot (WB) using digitonin solubilised mitochondrial fractions from Hs-A1, Ch-A1, and Hs-A1^{Exwitch}. WB sequentially probed with Ndufs1 (Supplementary Fig. 6h) and Sdha antibodies. All uncropped gels and blots in Source Data. **d**-**g** SC- supercomplex; HC- holocomplex; SubC- sub-complex. Sdha*- HC; Sdha**- SubC. **c**-**g** Data were verified in three independent experiments.



deficiency diseases including Leber hereditary optic neuropathy (LHON) syndrome, Leigh syndrome and mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes (MELAS) were considered. Mutations were mapped on available structures of human-RCs (CI – PDB: 5XTD; CII – PDB: 8GS8; CIII – PDB: 5XTE; CIV – PDB: 5Z62; CV – PDB: 8H9V) using PyMol and bond lengths were calculated. Only the side chain of mutating residue and not the backbone atoms were considered for PPI analysis. Cut-offs for hydrogen bond – 3.6 Å; salt bridges – 3.9 Å, and cation- π interactions – 5.4 Å. Amino acids with no interactions within these ranges were considered as non-interacting. Surface exposed residues were analyzed by commonly used PyMol plug-in FindSurfaceResidues⁸² (https://pymolwiki.org/index.php/FindSurfaceResidues). Cut-off for surface exposed residues was increased from default of 2.5 Å² to 3 Å². Red in Fig. 2d denotes

Fig. 6 | **Supplementing 18:3 fatty acid triggers CI-instability in HEK293T cells. a Top:** Schematic representing workflow. **Bottom:** Bar-graph showing abundance of different CL-species in mitochondria of Hs-AI-EGFP cells treated with BSA, linoleic and linolenic acid. *– CLs with multiple isomeric species; 68:2| (16:0)₂(18:1)₂ / (16:0)₁(16:1)₁(18:0)₁(18:1)₁; 68:5| (16:1)₂(18:1)₁(18:2)₁/ (16:0)₁(16:1)₂(18:2)₂; 72:5|(18:1)₃(18:2)₁/ (18:0)₁(18:1)₁(18:2)₂; 68:6| (16:1)₂(18:1)₁(18:0)₁(18:1)₂; 72:7|(16:0)₁(16:1)₁(18:0)₁(18:1)₁(18:1)₂(18:1)₁/ (16:0)₁(18:0)₁(18:3)₂; 70:7|(16:1)₁(18:1)₁(18:3)₂; 72:10| (16:1)₁(18:1)₂(18:3)₃/ (16:1)₂(18:2)₁(18:3)₃/ (18:0)₁(18:1)₁(18:3)₂; 72:10| (18:2)₂(18:3)₃/ (18:1)₁(18:3)₃; 68:7| (16:1)₂(18:3)₁/ (16:0)₁(16:1)₁(18:3)₂. Diamond: Data points and outliers. CLs with at least 3 data points have only been plotted. *n* = 6 biological replicates. Error bar: 1 SE from mean. Calculations in Source Data. **b. Left:** Schematic representing BSA-conjugated lipid supplementation assays. **Right:** Native-PAGE processed for *in-gel* fluorescence (IGF). Densitometry of SCs provided. Diamond: Data points. *n* = 3 biological replicates. Error bar: 1 SD from mean. **c** Immunoblot showing cleaved product of Ch-A1-



Fig. 7 | **Schematic summary of** *in cellulo* **substitution assays. Top left:** Human Ndufa1-EGFP (Hs-A1; blue) stably integrate into CI in HEK293T cells. **Bottom left:** Chimera Ndufa1-EGFP (Ch-A1) – Hs-A1 trans-IMM helix exchanged with At-A1, Hs-A1^{FP/LL} – Hs-A1 with lipid exposed amino acids F3 and P7 exchanged with At-A1 amino acid L3 and L7, Hs-A1^{Exswitch} – Hs-A1 with lipid exposed amino acids exchanged with At-A1 could not stably integrate into CI holocomplex and supercomplexes. **Top right:** treating cells with 18:2 fatty acid (Linoleic acid) does not perturb CI. **Bottom right:** treating cells with 18:3 fatty acid (alpha-linolenic acid) destabilizes CI.

all the atoms of side-chains being exposed, yellow denotes at least 1 atom of side-chain being exposed, and white denotes completely buried. Conservation analysis of mutating residues was performed using 13 representative organisms from 3 different kingdoms of Eukaryotes i.e., plants, Fungi and metazoa. Organisms are as follows:

- 1. Metazoa Homo sapiens, Bos taurus, Mus musculus, Danio rerio, Drosophila melanogaster, Caenorhabditis elegans.
- 2. Fungi Neurospora crassa, Yarrowia lipolytica, Aspergillus nidulans.
- 3. Plants Arabidopsis thaliana, Zea mays, Hibiscus syriacus, Gracilariopsis chorda.

Protein sequences of RC-subunits for these organisms were retrieved from NCBI and multiple sequence alignment (MSA) was performed using Clustal Omega from UniPro uGene⁸³.

Structure prediction and alignment

Structures of RC-deficiency disease-mutants of Ndufa2 and Ndufa1 were predicted using RoseTTAFold⁸⁴ (https://robetta.bakerlab.org/). No prediction templates were used. Structures of Ndufa1 chimeras/ mutants with or without interacting partners ND1 and Ndufa8 were predicted using AlphaFold2-multimer⁸⁵ (https://colab.research.google. com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb) where parameters were set as default. Predicted structures were aligned to available PDB structure (PDB: 5XTD). RMSDs were calculated using

EGFP upon supplementation with control BSA or BSA conjugated 18:2 and 18:3 fatty acids. Hsp60 as loading control. *Cleaved bands. Densitometry of cleaved bands normalized by Hsp60 provided. Diamond: Data points. n = 4 biological replicates. Error bar: 1 SD from mean. **d** Western blots (WB) using anti-Ndufs1, anti-Sdha and other mitochondria-specific (Glud1+Glud2, Hsp60) antibodies. WB sequentially probed with Ndufs1, Glud1+Glud2, Hsp60 and Sdha. <<< Ndufs1 SC; << Ndufs1 HC; < Ndufs1 SubC; \$Glud1+Glud2; #Hsp60; *Sdha HC; **Sdha SubC. Diamond: Data points. Densitometry analysis for Ndufs1-SCs normalized to Glud1+Glud2 from 3 biological replicates provided. Error bar: 1 SD from mean. **c-d** arb. units: Arbitrary unit. Native-PAGE processed for *in-gel* Cl-activity assay. Densitometry of activity bands (bracketed) provided. n = 3 biological replicates. Error bar: 1 SD from mean. (**b** & **e**) SC- supercomplex; HC- holocomplex; SubC- subcomplex. All uncropped gels and blots in Source Data. **b-e** P values from two-sided, paired Student's t-test with 95% confidence interval without any corrections.

online tool Pairwise Structure Alignment (https://www.rcsb.org/ alignment) using jCE alignment method⁸⁶ with default parameters.

blastp search, data filtration and heatmap preparation

Sequences of RC-subunits and subunits of Tom40 complex from organisms such as Homo sapiens, Neurospora crassa, Saccharomyces cerevisiae and Arabidopsis thaliana were retrieved from NCBI. blastp⁸⁷ was performed using these sequences where maximum search target was set as 5000 using non-redundant database, BLO-SUM62 matrix and rest of the parameters were default. Search results were filtered using an in-house python script where sequences corresponding to 134 representative organisms from Rebeaud et al.44 were filtered using e-value cut-off of 0.01 and minimum query coverage of 30% after removing all the duplicate entries. These 134 organisms belong to different clades covering all the supergroups and different taxanomic groups of eTOL (Supplementary Data S3). Percent identity w.r.t to query sequence was used to study conservation of subunits. Percent identity was plotted as heatmap using Morpheus (https://software.broadinstitute.org/morpheus/). Saccharomyces cerevisiae sequences were used for blastp searches for Fungi RCs and Tom40 complex. Neurospora crassa sequences were used for Fungi CI as Saccharomyces cerevisiae lacks canonical CI. Accession number of hits obtained by blastp search are mentioned in Supplementary Data S5a-g.

Violin plot was prepared using blastp filtered results. For a group of subunits, where blast search was restricted to one kingdom, an arbitrary value of 10 was imputed into other kingdoms, to indicate kingdom-specific sequence conservation.

Mean percent identity calculation of surface-exposed stretches

Continuous stretches of 25 amino acids from representative Clsubunits of H. sapiens (PDB: 5XTD) and corresponding stretches from A. thaliana (PDB: 7ARB) with more than 70% amino acids exposed to surface were identified using FindSurfaceResidues plugin of PyMol. Amino acid conservation at these exposed stretches was performed by MSA using Clustal Omega from UniPro uGene⁸³ on the filtered sequences obtained by blastp across eToL. Maximum number of iteration for MSA was set at 100 and rest of the parameters were default. Alignment ProfileGrid measuring % conservation was generated which indicates the % frequency of all 20 amino acids at respective positions⁸⁸. Sequences of surfaceexposed stretches of H. sapiens and A. thaliana were manually compared with the ProfileGrid and % frequencies corresponding to H. sapeins and A. thaliana sequences were selected. Mean of 25 residues was calculated and was termed as mean % identity. Eight nuclear encoded subunits, each from matrix and IMM arm were used for analysis. Two-sample Student's t-test was used for significance calculation.

Simulation

Model 1 and Model 2 IMMs along with trans-IMM helices of Ndufa1 were prepared using CHARMM-GUI server⁸⁹⁻⁹¹. The trans-IMM helices were placed along the bilayer normal (Z-axis) and at the center of the bilayer plane (X-Y plane). Pure IMMs and IMMs with trans-IMM helices were solvated with water molecules and neutralized ions. TIP3P model was used to describe water molecules⁹². KCl was added to neutralize the charge and maintain the physiological salt concentration of 150 mM. Energy minimization was carried out on pure membranes for 5000 steps using the steepest descent method to remove any atomic clashes followed by equilibration with position restraints of lipid headgroups. For model IMMs with trans-IMM helices, energy minimization was carried out on membranes for ~100 ns, 5000 steps with a short 1 ns equilibration with harmonic restrains on trans-IMM helices backbone atoms. Subsequently all restraints were released slowly and the systems were subjected to production run of 1000 ns. First 500 ns was considered as equilibration period and last 500 ns used for trajectory analyzes.

Simulations were conducted using the GROMACS tool (version 2021)⁹³ and all-atom CHARMM36 force field parameters^{94,95}. TIP3P model was used to describe water molecules⁹². The equation of motions was integrated at a timestep of 2 fs, using an isothermalisobaric (NPT) ensemble. To constrain the covalent bond lengths to hydrogen atoms, a linear constraint solver (LINCS) algorithm⁹⁶ was employed. A Nose-Hoover thermostat was used with a coupling constant of 1.0 ps to maintain the temperature of the systems at 303.15 K⁹⁷. The temperatures of the lipid bilayer and solvent (water and ions) were controlled independently. For models with trans-IMM helices, helix was considered as a separate temperature group. The Parrinello-Rahman barostat was used to control the system's pressure at 1 bar with a 5.0 ps coupling constant⁹⁸. With a cut-off of 12 Å, the long-range electrostatics interactions are considered via the particle-mesh Ewald (PME) method⁹⁹. The system's periodic boundary conditions are applied to all three directions within the Ewald summation method. The Lennard-Iones potential was used to compute the van der Waals interactions with a cutoff of 12 Å by employing a force-based switching function (at 10 Å). The analyzes were performed using standard gmx tools. VMD was used for trajectory visualization¹⁰⁰.

5 Å cut-off was used to calculate trans-IMM helix-fatty acyl chain contacts including the only heavy atoms. Number of contacts of all the acyl chains of same lipid species (for example16:0, 18:0, 18:2, 18:3, and 20:4 for PC) were summed together as followed.

Total contacts at matrix leaflet with PC =
$$\sum_{i=1}^{12} ((16:0)_i + (18:0)_i + (18:1)_i + (18:2)_i + (20:4)_i)$$

Similarly, number of contacts with other lipids in both leaflets were calculated. Helix-lipid interactions shown in the schematic (Fig. 3i) was prepared as per total contacts shown in Fig. 3h. Number of lipid molecules per lipid species were placed depending on the total contact that lipid species establishes with the trans-IMM helix. Less than 50 contacts were ignored; 1 lipid molecule was placed for 50-110 contacts; 2 lipid molecules for 110-150 contacts; 3 lipid molecules for 150and above contacts. Lipid of a particular lipid species that makes maximum contacts with the trans-IMM helix was only use for representation. Rest were ignored.

Expression constructs

EGFP was PCR-amplified from pCIneo-FlucDM-EGFP¹⁰¹ and subcloned in pcDNA4/TO using the restriction enzymes XhoI and XbaI (NEB). Ndufa1 and Ndufa2 was amplified from HEK293T cDNA and cloned into pcDNA4/TO-EGFP using KpnI and XhoI (NEB). Gene sequences of At-A1, Ch-A1 and Hs-A1^{Ex-switch} were codon optimized for human, synthesized, and sub-cloned in pcDNA4/TO-EGFP. Variants of Ndufa2 and Ndufa1 were prepared by site directed mutagenesis using primers in Supplementary Table 3. Ch^{Mm-switch} was prepared by overlap extension PCR, where 1st 23 amino acids were amplified from Neuro2A cDNA and 24th to 70th amino acid was amplified from pcDNA4/TO-Hs-A1-EGFP using primers in Supplementary Table 3. These two fragments were annealed, extended and cloned in pcDNA4/TO-EGFP using Kpn1 and Xho1. 28th to 70th amino acids from Hs-A1 was amplified and inserted before EGFP in pcDNA4/TO-EGFP using Kpn1 and Xho1 to prepare Δ 27-Hs-A1-EGFP. Primers are indicated in Supplementary Table 3.

Cell culture and transfection

HEK293T-6TR (stable overexpression of tetracycline repressor, pcDNA6/TR)¹⁰² cells were maintained in Dulbecco's Modified Eagle's Medium (Gibco) supplemented with 10% fetal bovine serum (FBS) (Gibco) and 90 U/ml penicillin (Sigma-Aldrich) 50 µg/ml streptomycin (Sigma-Aldrich) at 37 °C and 5% CO₂. Transfection of cells was performed with Lipofectamine 3000 reagent (Invitrogen) as per manufacturer's protocol. For stable cell line preparation, pcDNA4/TO system was used (Life Technologies). HEK293T-6TR cells was transfected with various pcDNA4/TO constructs and selected with 200 µg/ml Zeocin (Invitrogen). Surviving colonies after 2-3 weeks were tested for induction by adding 1µg/ml Doxycycline (MP Biomedicals) by fluorescence microscopy and western blot. Positive clones were frozen and/or maintained for further experiment.

For western blot and mitochondria isolation, cells were plated and allowed to adhere for 12 h. 1 μ g/ml Doxycycline induction was given for 36 h and cells were pelleted, washed with PBS and stored in -80°C until further used. 150 mm×3 dishes was used for mitochondria isolation. 60 mm×1 dish was used for total cell lysate.

BSA-conjugated lipid supplementation

HEK293T cells stably expressing Hs-A1-EGFP and Ch-A1-EGFP were treated with BSA-conjugated linoleic acid and alpha-linolenic acid as described by Omer et al with slight modifications⁵⁷. Briefly, 10 mM solutions of fatty acids (Sigma) were prepared by dissolving in absolute ethanol. Meanwhile. 10% fatty acid-free BSA (Sigma) was prepared in 1X PBS (pH 7.4) and incubated at 55 °C for 30 mins with intermittent vortexing. The solution was cooled to 25 °C and filter sterilized. The 10 mM fatty acid solution was diluted 1:20 in 10% BSA (final fatty acid concentration: 500 µM) and incubated at 55 °C for 30 mins with intermittent vortexing. The BSA-conjugated fatty acid stock solution was further diluted 1:10 in incomplete DMEM (final concentration: 50 µM of fatty acid in 1% BSA) and supplemented to cells. Control cells were supplemented with BSA-conjugated ethanol. The cells were induced with 1µg/ml doxycycline and incubated at 37 °C and 5% CO₂ for 72 h. Prior to lipid supplementation, cells were plated in DMEM and 10% FBS and allowed to adhere for at least 16 h.

Immunofluorescence microscopy

Cells were grown on 18 mm glass coverslips in 12 well plates (Nunc) and fixed using 4% paraformaldehyde. 0.1% Triton X-100 in PBS was used to permeabilize the cells prior to immunostaining. Next, the coverslips were incubated with 5% BSA prepared in PBS with 0.05% Tween-20 (PBST) for 30 min at RT with constant shaking. 1:1000 dilution of Tom20 primary antibody (Santa Cruz Biotechnology) in 0.1% BSA with PBST was added and coverslips were incubated at 4 °C overnight followed by PBST wash. Next day, 1:500 dilution of Alexa Fluor Plus 647 Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody (Invitrogen) was added to the coverslips and incubated for 1 hr at room temperature. Coverslips were mounted on slides using 10 µl Antifade Mounting Medium (cat: H-1000, Vectashield) after DAPI staining. Imaging was done using OLYMPUS FV3000 and Leica TCS SP8 microscope. Colocalization coefficients were calculated using Coloc2 plug-in in ImageJ using default parameters (Supplementary Table 1).

Mitochondria-enriched fraction preparation was performed according to Rawat et al. with slight modification¹². Briefly, ~20 million cells were homogenized at 4 °C in a Dounce homogenizer with 600 μ l of buffer A (83 mM sucrose, 10 mM MOPS pH 7.2). After adding an equal volume of buffer B (250 mM sucrose, 30 mM MOPS pH 7.2), nuclei and unbroken cells were removed by centrifugation at 1000 g for 5 min at 4 °C. The supernatant was further centrifuged at 12,000 g for 5 min and washed twice with buffer C (320 mM sucrose, 1 mM EDTA, 10 mM Tris-HCl pH 7.4). The pellet was resuspended in 1X NativePAGE sample buffer (Invitrogen) to obtain mitochondria-enriched fraction followed by storing at -80°C until further use.

Mitoplast preparation protocol was adopted from Hussain M et al. with some modifications¹⁰³. 200 μ g of mitochondria-enriched fraction was suspended in hypotonic buffer (20 mM KCl and 10 mM HEPES, pH 7.2) and was treated with trypsin (final concentration 20 μ g/ml) (Thermo Scientific) in ice for 15 min. Next, 1X PIC (cocktail of protease inhibitors) (Roche) was added and the fraction was passed through sucrose cushion (0.8 M sucrose in 10 mM HEPES, pH 7.2) at 12000 g for 10 min. Pellet fraction was resuspended in hypotonic buffer, 0.1% digitonin, and incubated in ice for 20 min and centrifuged at 12000 g. Same steps from resuspension of pellet to passing through sucrose cushion were repeated once with 10 μ g/ml trypsin for 5 min to obtain the mitoplasts as pellet.

Cell pellets, mitochondria-enriched pellets, and mitoplast pellets were lysed in SDS-Lysis buffer (0.2 M Tris–HCl, pH 6.8, 4% SDS, 4% 2-mercaptoethanol, 40% glycerol), vortexed. Cell pellets were heated at 95 °C for 15 min whereas mitochondria-enriched pellets and mitoplast pellets were heated at 55 °C for 10 min followed by centrifugation at 12,000 g for 15 min to prepare total cell lysate, mitochondrial fraction, and mitoplast fraction, respectively. Protein estimation was carried out by Amido Black Protein assay, fractions were separated by SDS-PAGE, and transferred onto 0.2 μ m PVDF membrane (Roche) for 90 min at 300 mA using Mini-Trans Blot cell system (Bio-Rad). Membranes were probed with primary and secondary antibodies (Supplementary Table 4) and imaged using Chemidoc MP imaging system, Bio-Rad.

BN-PAGE, in-gel activity and native western blotting

Protein content in mitochondria-enriched fractions were estimated using Bradford Assay reagent (Sigma-Aldrich) followed by solubilization as previously described¹². Briefly, 12 gram digitonin (Sigma-Aldrich) per gram of mitochondria-enriched fraction was added and incubated for 10 min in ice. Supernatant was collected after 30 min centrifugation at 13,000 g. Protein concentration in supernatant (mitochondrial-soluble fraction) was estimated by Bradford Assay. Next, 6 gram of 5% G-250 NativePAGE Sample additive (Invitrogen) was added per gram of mitochondrial-soluble fraction. Glycerol was added at final concentration of 10% and 20-30 µg of sample was loaded onto NativePAGE 3-12% Bis-Tris Protein Gels (Invitrogen). Native PAGE was run using anode buffer (50 mM Bis-Tris HCl, pH 7.0) and blue cathode buffer (50 mM Tricine, 15 mM Bis-Tris HCl, pH 7.0, 0.02% Coomassie Brilliant Blue G-250) at 100 V until samples entered the separating gel and then at 150 V until the dye front reached 1/3rd of the entire gel. At this point, blue cathode buffer was replaced with dye-free cathode buffer for the remainder of the run at 250 V until the dye front reached the end of the gel.

in-gel CI-Activity—the gel was incubated overnight at room temperature in freshly prepared CI NADH dehydrogenase substrate solution (5 mM Tris-HCl pH 7.4, 0.1 mg/ml NADH and 0.7 mg/ml Nitrotetrazolium Blue chloride). Appearance of purple-bands is indicative of CI activity in the gel. The reaction was stopped with 10% acetic acid, washed with water and imaged. Native western blotting–Western blotting after BN–PAGE was carried out similar to previously reported¹². Briefly, the gel was equilibrated in soaking buffer (48 mM Tris-HCl, 39 mM glycine, 0.0037% SDS and 20% methanol) for 30 min and transferred onto 0.45 µm PVDF membrane (Roche) overnight at 30 V, followed by 300 mA for 90 min, using the Mini-Trans Blot cell system (Bio-Rad). The membrane was stained with 0.1% Coomassie Brilliant Blue R-250 dye, 50% methanol and 7% acetic acid and destained with destaining solution 1 (50% methanol and 7% acetic acid) followed by destaining solution 2 (90% methanol and 10% acetic acid) and final destaining using 100% methanol. The membrane was washed with Tris-buffered saline containing Tween 20 (TBST) followed by blocking with 5% BSA in TBST for overnight and probed with appropriate primary and secondary antibodies (Supplementary Table 4).

Lipid isolation and lipidomics by LC-MS/MS

150 µg of mitochondria isolated from BSA, linoleic and alpha-linolenic acid treated Hs-A1-EGFP cells were resuspended in 250 µl of PBS. Lipids were isolated by modified Folch method¹⁰⁴. Briefly, 750 µl of 2:1 (v/v) chloroform (CHCl₃): methanol (MeOH) containing CL 56:0 (14:0)₄ as an internal standard (IS) was added to the mitochondrial fraction and vortexed thoroughly. Phase separation was achieved by centrifugation at 2000 g for 5 mins. The organic phase was collected in a fresh glass vial. Re-extraction was performed from the aqueous phase by adding 500 µl of chloroform, 10% (v/v) formic acid, followed by vortex and centrifuge. The organic phase was pooled and dried under the stream of nitrogen gas. Isolated mitochondria and dried lipids were stored in -80 until further use.

Dried lipids were resuspended in 60 µl of 2:1 (v/v) dichloromethane (CH_2Cl_2): MeOH on the day of analysis. 25 µl of sample was analyzed using Shimadzu UHPLC LC40D coupled with Sciex QTRAP 6500+ mass spectrometer. Lipid extracts were separated on C18 reverse phase column (Phenomenex, 5 µm, 50 ×4.6 mm) using solvent A as 95:5 (v/v) H₂O:MeOH + 0.1% Formic acid + 10 mM ammonium formate and solvent B as 60:35:5 (v/v/v) Isopropanol:MeOH:H₂O + 0.1% Formic acid + 10 mM ammonium formate. LC gradient was set for 65 min: 5% B for 5 min with 0.5 ml/min flow, a linear gradient of B from 5-100% over 25 min with 0.3 ml/min flow. 100% B from 25-55 min with 0.3 ml/min flow, 100% B from 55-60 min with 0.5 ml/min flow and reequilibration with 5% B for 5 min. The column was maintained at 55 °C. Lipids were quantified using scheduled multiple reaction monitoring (MRM) method in high mass and negative ion mode using collisioninduced dissociation (CID). 96 transitions from 45 CL species including IS (Supplementary Table 5) were monitored. Source and gas settings: curtain gas pressure 35 A.U.; Collision Gas (CAD) High; Ionspray Voltage -4500V; temperature 550 °C; Ion source gas 1&2 55; declustering potential -80V; collision energy -120V; entrance potential -10V; exit potential -13V; MRM detection window of 45 s for IS and 180 s for other CLs; minimum dwell time100ms. Multiple blanks were run between samples.

Lipidomics data analysis

The area under curve (AUC) was calculated using Sciex OS, Analytics version 3.3.1.43 software with a minimum peak width of 7 points while the rest of the parameters were default. A minimum 3 transitions per CL species were monitored (Supplementary Table 5) where one transition – -152.9 corresponded to the head group (Glycerolphosphate- H_2O) and the other two transitions belonged to fatty acid chains. -152.9 was defined as a qualifier. CL species lacking qualifier peaks were not considered for analysis. Further, 2 transitions of fatty acids were considered as quantifiers, and CL species where both the transitions were identified and had AUC greater than or equal to the order of qualifier were only considered for analysis. Each spectra was subtracted with its respective blank. AUC of each CL species was normalized and %

Abundance was calculated using the formula

$$AUC_N = \frac{AUC \text{ of individual } CL}{AUC \text{ of } IS}$$
(1)

% Abundance of indivdual CL species =
$$\frac{AUC_N}{\sum AUC_N} X100$$
 (2)

One transition per species was used to plot a graph.

Statistics & Reproducibility

All the bar graphs, scatter plots, and violin plots were prepared using OriginPro 2021 (OriginLab Corporation, Northampton, MA, USA). Statistical analysis was performed in OriginPro 2021 using a two-sided Student's t-test with 95% confidence interval without any corrections. Stacked bar graphs were prepared using GraphPad Prism 7.0 (Graph-Pad Software, Boston, Massachusetts USA). The bubble graph was prepared using the online tool Visual Paradigm (https://online.visualparadigm.com/). Details are provided in the figure legends. All the experiments except bio-informatics and MD-simulation were performed at least thrice. A number of biological replicates used for each statistical analysis are provided in the figure legend. Schematics were prepared using Adobe Illustrator CC 2019.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Lipidomics data generated during this study have been deposited in Metabolomics Workbench ST003660 [https://doi.org/10.21228/ M86N9Q]¹⁰⁵. Raw data of MD-simulation and uncropped gel and blots are provided as Source data with this paper (https://doi.org/10. 6084/m9.figshare.27050524)¹⁰⁶. Source data are provided with this paper.

Code availability

Python script used for blastp data filtration is available on Code Ocean (https://doi.org/10.24433/CO.7212579.v1)¹⁰⁷.

References

- 1. Greco, T. M. & Cristea, I. M. The biochemical evolution of protein complexes. *Trends Biochem Sci.* **41**, 4–6 (2016).
- Wan, C. et al. Panorama of ancient metazoan macromolecular complexes. *Nature* 525, 339–344 (2015).
- Yates, C. M. & Sternberg, M. J. The effects of non-synonymous single nucleotide polymorphisms (nsSNPs) on protein-protein interactions. J. Mol. Biol. 425, 3949–3963 (2013).
- Lewis, A. C., Saeed, R. & Deane, C. M. Predicting protein-protein interactions in the context of protein evolution. *Mol. Biosyst.* 6, 55–64 (2010).
- Marsh, J. A. & Teichmann, S. A. Structure, dynamics, assembly, and evolution of protein complexes. *Annu Rev. Biochem* 84, 551–575 (2014).
- Romanauska, A. & Kohler, A. Lipid saturation controls nuclear envelope function. *Nat. Cell Biol.* 25, 1290–1302 (2023).
- Machin, J. M., Kalli, A. C., Ranson, N. A. & Radford, S. E. Proteinlipid charge interactions control the folding of outer membrane proteins into asymmetric membranes. *Nat. Chem.* 15, 1754–1764 (2023).
- Raghupathy, R. et al. Transbilayer lipid interactions mediate nanoclustering of lipid-anchored proteins. *Cell* 161, 581–594 (2015).
- 9. Huang, X. et al. Sequences flanking the transmembrane segments facilitate mitochondrial localization and membrane fusion by mitofusin. *Proc. Natl Acad. Sci. USA* **114**, E9863–E9872 (2017).

- Pogozheva, I. D., Mosberg, H. I. & Lomize, A. L. Life at the border: adaptation of proteins to anisotropic membrane environment. *Protein Sci.* 23, 1165–1196 (2014).
- 11. Fernandez-Vizarra, E. & Zeviani, M. Mitochondrial complex III Rieske Fe-S protein processing and assembly. *Cell Cycle* **17**, 681–687 (2018).
- 12. Rawat S., Ghosh S., Mondal D., Anusha V., Raychaudhuri S. Increased supraorganization of respiratory complexes is a dynamic multistep remodelling in response to proteostasis stress. *J. Cell Sci.* **133**, jcs248492 (2020).
- 13. Hill, G. E. Mitonuclear Compensatory Coevolution. *Trends Genet* **36**, 403–414 (2020).
- Yan, Z., Ye, G. & Werren, J. H. Evolutionary rate correlation between mitochondrial-encoded and mitochondria-associated nuclear-encoded proteins in insects. *Mol. Biol. Evol.* 36, 1022–1036 (2019).
- 15. Mishmar, D. et al. Adaptive selection of mitochondrial complex I subunits during primate radiation. *Gene* **378**, 11–18 (2006).
- Zhou, L., Maldonado, M., Padavannil, A., Guo, F. & Letts, J. A. Structures of Tetrahymena's respiratory chain reveal the diversity of eukaryotic core metabolism. *Science* **376**, 831–839 (2022).
- Soufari, H., Parrot, C., Kuhn, L., Waltz, F. & Hashem, Y. Specific features and assembly of the plant mitochondrial complex I revealed by cryo-EM. *Nat. Commun.* 11, 5195 (2020).
- Oemer, G. et al. Molecular structural diversity of mitochondrial cardiolipins. Proc. Natl Acad. Sci. USA 115, 4158–4163 (2018).
- Panek, T., Elias, M., Vancova, M., Lukes, J. & Hashimi, H. Returning to the fold for lessons in mitochondrial crista diversity and evolution. *Curr. Biol.* **30**, R575–R588 (2020).
- Schoch C. L., et al. Ncbi Taxonomy: A Comprehensive Update On Curation, Resources And Tools. Database (Oxford) 2020, (2020).
- Rawat, S., Anusha, V., Jha, M., Sreedurgalakshmi, K. & Raychaudhuri, S. Aggregation of Respiratory Complex Subunits Marks the Onset of Proteotoxicity in Proteasome Inhibited Cells. J. Mol. Biol. 431, 996–1015 (2019).
- Guo, R., Zong, S., Wu, M., Gu, J. & Yang, M. Architecture of Human Mitochondrial Respiratory Megacomplex I2III2IV2. *Cell* **170**, 1247–1257 e1212 (2018).
- 23. Fiedorczuk, K. et al. Atomic structure of the entire mammalian mitochondrial complex I. *Nature* **538**, 406–410 (2016).
- 24. Perrier, S. et al. Recessive mutations in NDUFA2 cause mitochondrial leukoencephalopathy. *Clin. Genet* **93**, 396–400 (2018).
- Fernandez-Moreira, D. et al. X-linked NDUFA1 gene mutations associated with mitochondrial encephalomyopathy. *Ann. Neurol.* 61, 73–83 (2007).
- Stroud, D. A. et al. Accessory subunits are integral for assembly and function of human mitochondrial complex I. *Nature* 538, 123–126 (2016).
- Dieteren, C. E. et al. Subunit-specific incorporation efficiency and kinetics in mitochondrial complex I homeostasis. *J. Biol. Chem.* 287, 41851–41860 (2012).
- Whittaker, R. H. New concepts of kingdoms or organisms. evolutionary relations are better represented by new classifications than by the traditional two kingdoms. *Science* **163**, 150–160 (1969).
- 29. Keeling, P. J. et al. The tree of eukaryotes. *Trends Ecol. Evol.* **20**, 670–676 (2005).
- Burki, F., Roger, A. J., Brown, M. W. & Simpson, A. G. B. The new tree of eukaryotes. *Trends Ecol. Evol.* 35, 43–55 (2020).
- Klusch, N., Senkler, J., Yildiz, O., Kuhlbrandt, W. & Braun, H. P. A ferredoxin bridge connects the two arms of plant mitochondrial complex I. *Plant Cell* 33, 2072–2091 (2021).
- 32. Parey, K. et al. High-resolution structure and dynamics of mitochondrial complex I-Insights into the proton pumping mechanism. *Sci. Adv.* **7**, eabj3221 (2021).

- 33. Schikowsky, C., Senkler, J. & Braun, H. P. SDH6 and SDH7 contribute to anchoring succinate dehydrogenase to the inner mitochondrial membrane in arabidopsis thaliana. *Plant Physiol.* **173**, 1094–1108 (2017).
- 34. Briere, J. J. et al. Succinate dehydrogenase deficiency in human. *Cell Mol. Life Sci.* **62**, 2317–2324 (2005).
- Stephan, K. & Ott, M. Timing of dimerization of the bc(1) complex during mitochondrial respiratory chain assembly. *Biochim Biophys. Acta Bioenerg.* **1861**, 148177 (2020).
- 36. Zong, S. et al. Structure of the intact 14-subunit human cytochrome c oxidase. *Cell Res* **28**, 1026–1034 (2018).
- 37. Lai, Y. et al. Structure of the human ATP synthase. *Mol. Cell* **83**, 2137–2147.e2134 (2023).
- Han, F. et al. Structures of Tetrahymena thermophila respiratory megacomplexes on the tubular mitochondrial cristae. *Nat. Commun.* 14, 2542 (2023).
- Miranda-Astudillo, H. V., Yadav, K. N. S., Boekema, E. J. & Cardol, P. Supramolecular associations between atypical oxidative phosphorylation complexes of Euglena gracilis. *J. Bioenerg. Biomembr.* 53, 351–363 (2021).
- 40. Perez, E. et al. The mitochondrial respiratory chain of the secondary green alga Euglena gracilis shares many additional subunits with parasitic Trypanosomatidae. *Mitochondrion* **19**, 338–349 (2014).
- 41. Clark, J. W. et al. Evolution of phenotypic disparity in the plant kingdom. *Nat. Plants* **9**, 1618–1626 (2023).
- 42. Guijarro-Clarke, C., Holland, P. W. H. & Paps, J. Widespread patterns of gene loss in the evolution of the animal kingdom. *Nat. Ecol. Evol.* **4**, 519–523 (2020).
- 43. Li, Y. et al. A genome-scale phylogeny of the kingdom Fungi. *Curr. Biol.* **31**, 1653–1665 e1655 (2021).
- 44. Rebeaud M. E., Mallik S., Goloubinoff P., Tawfik D. S. On the evolution of chaperones and cochaperones and the expansion of proteomes across the tree of life. *Proc. Natl. Acad. Sci. USA* **118**, e2020885118 (2021).
- 45. Shiota, T. et al. Molecular architecture of the active mitochondrial protein gate. *Science* **349**, 1544–1548 (2015).
- 46. Carranza, G. et al. Cardiolipin plays an essential role in the formation of intracellular membranes in Escherichia coli. *Biochim Biophys. Acta Biomembr.* **1859**, 1124–1132 (2017).
- Gebert, N. et al. Mitochondrial cardiolipin involved in outermembrane protein biogenesis: implications for Barth syndrome. *Curr. Biol.* 19, 2133–2139 (2009).
- Grba, D. N., Chung, I., Bridges, H. R., Agip, A. A. & Hirst, J. Investigation of hydrated channels and proton pathways in a highresolution cryo-EM structure of mammalian complex I. Sci. Adv. 9, eadi1359 (2023).
- Arnarez, C., Mazat, J. P., Elezgaray, J., Marrink, S. J. & Periole, X. Evidence for cardiolipin binding sites on the membrane-exposed surface of the cytochrome bc1. J. Am. Chem. Soc. 135, 3112–3120 (2013).
- 50. Malkamaki, A. & Sharma, V. Atomistic insights into cardiolipin binding sites of cytochrome c oxidase. *Biochim Biophys. Acta Bioenerg.* **1860**, 224–232 (2019).
- Oemer, G. et al. Phospholipid acyl chain diversity controls the tissue-specific assembly of mitochondrial cardiolipins. *Cell Rep.* 30, 4281–4291 e4284 (2020).
- Liu, Y. T., Senkler, J., Herrfurth, C., Braun, H. P. & Feussner, I. Defining the lipidome of Arabidopsis leaf mitochondria: specific lipid complement and biosynthesis capacity. *Plant Physiol.* **191**, 2185–2203 (2023).
- 53. Manna, M. & Mukhopadhyay, C. Cholesterol driven alteration of the conformation and dynamics of phospholamban in model membranes. *Phys. Chem. Chem. Phys.* **13**, 20188–20198 (2011).

- 54. Friedrich, M. G., Wang, Z., Schey, K. L. & Truscott, R. J. W. Spontaneous cleavage at Glu and Gln Residues in long-lived proteins. *ACS Chem. Biol.* **16**, 2244–2254 (2021).
- 55. Kato, K., Nakayoshi, T., Ishikawa, Y., Kurimoto, E. & Oda, A. Computational analysis of the mechanism of nonenzymatic peptide bond cleavage at the C-terminal side of an asparagine residue. *ACS Omega* **6**, 30078–30084 (2021).
- Capasso, S., Mazzarella, L., Sorrentino, G., Balboni, G. & Kirby, A. J. Kinetics and mechanism of the cleavage of the peptide bond next to asparagine. *Peptides* 17, 1075–1077 (1996).
- 57. Oemer, G. et al. The lipid environment modulates cardiolipin and phospholipid constitution in wild type and tafazzin-deficient cells. *J. Inherit. Metab. Dis.* **45**, 38–50 (2022).
- 58. Oemer, G. et al. Fatty acyl availability modulates cardiolipin composition and alters mitochondrial function in HeLa cells. *J. Lipid Res* **62**, 100111 (2021).
- 59. Gao, Y. et al. Mitochondrial folate metabolism-mediated alphalinolenic acid exhaustion masks liver fibrosis resolution. *J. Biol. Chem.* **299**, 104909 (2023).
- 60. Monaco, C. M. F. et al. alpha-linolenic acid supplementation prevents exercise-induced improvements in white adipose tissue mitochondrial bioenergetics and whole-body glucose homeostasis in obese Zucker rats. *Diabetologia* **61**, 433–444 (2018).
- Oberai, A., Joh, N. H., Pettit, F. K. & Bowie, J. U. Structural imperatives impose diverse evolutionary constraints on helical membrane proteins. *Proc. Natl Acad. Sci. USA* **106**, 17747–17750 (2009).
- 62. Corin, K. & Bowie, J. U. How physical forces drive the process of helical membrane protein folding. *EMBO Rep.* **23**, e53025 (2022).
- 63. Gould, S. B. Membranes and evolution. *Curr. Biol.* **28**, R381–R385 (2018).
- Falabella, M., Vernon, H. J., Hanna, M. G., Claypool, S. M. & Pitceathly, R. D. S. Cardiolipin, mitochondria, and neurological disease. *Trends Endocrinol. Metab.* 32, 224–237 (2021).
- Kiebish, M. A., Han, X., Cheng, H., Chuang, J. H. & Seyfried, T. N. Cardiolipin and electron transport chain abnormalities in mouse brain tumor mitochondria: lipidomic evidence supporting the Warburg theory of cancer. J. Lipid Res. 49, 2545–2556 (2008).
- Petrosillo, G. et al. Mitochondrial dysfunction in rat with nonalcoholic fatty liver Involvement of complex I, reactive oxygen species and cardiolipin. *Biochim Biophys. Acta* 1767, 1260–1267 (2007).
- 67. Cogliati, S. et al. Mechanism of super-assembly of respiratory complexes III and IV. *Nature* **539**, 579–582 (2016).
- 68. Milenkovic, D. et al. Preserved respiratory chain capacity and physiology in mice with profoundly reduced levels of mitochondrial respirasomes. *Cell Metab.* **35**, 1799–1813 e1797 (2023).
- Gorman, G. S. et al. Mitochondrial diseases. *Nat. Rev. Dis. Prim.* 2, 16080 (2016).
- Harayama T., Antonny B. Beyond fluidity: the role of lipid unsaturation in membrane function. *Cold Spring Harb. Perspect. Biol.* 15, a041409 (2023).
- Upchurch, R. G. Fatty acid unsaturation, mobilization, and regulation in the response of plants to stress. *Biotechnol. Lett.* 30, 967–977 (2008).
- 72. Fiorini, R. et al. Lipid-protein interactions in mitochondrial membranes from bivalve mollusks: molecular strategies in different species. *Comp. Biochem Physiol. B Biochem Mol. Biol.* **227**, 12–20 (2019).
- 73. Qi, L. et al. Cryo-EM structure of the human mitochondrial translocase TIM22 complex. *Cell Res* **31**, 369–372 (2021).
- 74. Fielden, L. F. et al. Central role of Tim17 in mitochondrial presequence protein translocation. *Nature* **621**, 627–634 (2023).
- 75. Sim, S. I., Chen, Y., Lynch, D. L., Gumbart, J. C. & Park, E. Structural basis of mitochondrial protein import by the TIM23 complex. *Nature* **621**, 620–626 (2023).

- Wiedemann, N. & Pfanner, N. Mitochondrial machineries for protein import and assembly. *Annu Rev. Biochem* 86, 685–714 (2017).
- Busch, J. D., Fielden, L. F., Pfanner, N. & Wiedemann, N. Mitochondrial protein transport: versatility of translocases and mechanisms. *Mol. Cell* 83, 890–910 (2023).
- Kemper, C. et al. Integration of tail-anchored proteins into the mitochondrial outer membrane does not require any known import components. *J. Cell Sci.* **121**, 1990–1998 (2008).
- Krumpe, K. et al. Ergosterol content specifies targeting of tailanchored proteins to mitochondrial outer membranes. *Mol. Biol. Cell* 23, 3927–3935 (2012).
- Teixeira, L. M. C., Coimbra, J. T. S., Ramos, M. J. & Fernandes, P. A. Transmembrane protease serine 2 proteolytic cleavage of the SARS-CoV-2 Spike protein: a mechanistic quantum mechanics/molecular mechanics study to inspire the design of new drugs to fight the COVID-19 pandemic. J. Chem. Inf. Model 62, 2510–2521 (2022).
- Lorent, J. H. et al. Structural determinants and functional consequences of protein affinity for membrane rafts. *Nat. Commun.* 8, 1219 (2017).
- Lo, R. S. et al. The functional impact of 1,570 individual amino acid substitutions in human OTC. Am. J. Hum. Genet 110, 863–879 (2023).
- Okonechnikov, K., Golosova, O. & Fursov, M. team U. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* 28, 1166–1167 (2012).
- Baek, M. et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).
- 85. Mirdita, M. et al. ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).
- Shindyalov, I. N. & Bourne, P. E. Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng.* **11**, 739–747 (1998).
- 87. Camacho, C. et al. BLAST+: architecture and applications. *BMC Bioinforma*. **10**, 421 (2009).
- Roca, A. I., Almada, A. E. & Abajian, A. C. ProfileGrids as a new visual representation of large multiple sequence alignments: a case study of the RecA protein family. *BMC Bioinforma*. 9, 554 (2008).
- Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. J. Comput Chem. 29, 1859–1865 (2008).
- Jo, S., Kim, T. & Im, W. Automated builder and database of protein/ membrane complexes for molecular dynamics simulations. *PLoS One* 2, e880 (2007).
- Lee, J. et al. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. J. Chem. Theory Comput 12, 405–413 (2016).
- Jorgensen WLC; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. J Chem. Phys. 79, 926–935 (1983).
- Van Der Spoel, D. et al. GROMACS: fast, flexible, and free. J. Comput Chem. 26, 1701–1718 (2005).
- Klauda, J. B., Monje, V., Kim, T. & Im, W. Improving the CHARMM force field for polyunsaturated fatty acid chains. *J. Phys. Chem. B* 116, 9424–9431 (2012).
- Klauda, J. B. et al. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. J. Phys. Chem. B 114, 7830–7843 (2010).
- Hess, B., Bekker, H., Berendsen, H. J. & Fraaije, J. G. LINCS: A linear constraint solver for molecular simulations. *J. computational Chem.* 12, 1463–1472 (1997).
- 97. S. N. A unified formulation of the constant temperature molecular dynamics methods. **81**, 511–519 (1984).
- Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: a new molecular dynamics method. J. Appl. Phys. 52, 7182–7190 (1981).

- Darden, T. Y. D. & Pedersen, L. Particle mesh Ewald: An N- log (N) method for Ewald sums in large systems. J. Chem. Phys. 98, 10089–10092 (1993).
- 100. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. J. Mol. Graph **14**, 27–38 (1996). 33-38.
- 101. Gupta, R. et al. Firefly luciferase mutants as sensors of proteome stress. *Nat. Methods* **8**, 879–884 (2011).
- 102. Mansuri S., et al. Widespread Nuclear Lamina injuries defeat proteostatic purposes of alpha-Synuclein amyloid Inclusions. *J. Cell Sci.*, **137**, jcs261935 (2024).
- 103. Hussain, M. et al. MITOL-dependent ubiquitylation negatively regulates the entry of PolgammaA into mitochondria. *PLoS Biol.* 19, e3001139 (2021).
- Kelkar, D. S. et al. A chemical-genetic screen identifies ABHD12 as an oxidized-phosphatidylserine lipase. *Nat. Chem. Biol.* 15, 169–178 (2019).
- Gupta, P. et al. Kingdom-specific lipid unsaturation calibrates sequence evolution in membrane arm subunits of eukaryotic respiratory complexes. *metabolomicsworkbench*, https://doi.org/ 10.21228/M86N9Q (2025).
- 106. Gupta, P. et al. Kingdom-specific lipid unsaturation calibrates sequence evolution in membrane arm subunits of eukaryotic respiratory complexes. *figshare*, https://doi.org/10.6084/m9. figshare.27050524 (2025).
- Gupta, P. et al. Kingdom-specific lipid unsaturation calibrates sequence evolution in membrane arm subunits of eukaryotic respiratory complexes. codeocean, https://doi.org/10.24433/CO. 7212579.v1 (2025).

Acknowledgements

Drs. Shrish Tiwari, Mandar V. Deshmukh, Biswajit Pal, Hirak Chakraborty, Pooja Joshi for discussion, Debadutta Patra for python script, Dr. Pradeep Kumar for PyMol analysis, CSIR-4PI for computation space. Funding from CSIR grants MLP0143, MLP0159, and HCP47 is acknowledged.

Author contributions

P.G. and S.R. conceived and designed the project. P.G., S.C, and S.B. performed bioinformatics. M.M. and A.K.R performed MD simulations. P.G., S.C, S.G., and P.R.T. performed molecular and cell biology experiments. P.G., K.R.K., K.Y, and R.B. executed lipid analysis. R.N. provided reagents. P.G., S.C, R.N., and S.R. analyzed results. S.R. supervised the work, arranged funding, and provided reagents. P.G. and S.R. drafted the manuscript. All authors read and edited manuscript.

Competing interests

The authors declares no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-57295-7.

Correspondence and requests for materials should be addressed to Swasti Raychaudhuri.

Peer review information *Nature Communications* thanks Steven Claypool, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http:// creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2025