STRUCTURAL BIOLOGY

High-resolution cryo-EM structures of respiratory complex I: Mechanism, assembly, and disease

Kristian Parey[1,2,3], Outi Haapanen[4], Vivek Sharma[4,5], Harald Köfeler[6], Thomas Züllig[6], Simone Prinz[2], Karin Siegmund[1,3], Ilka Wittig[7], Deryck J. Mills[2], Janet Vonck[2], Werner Kühlbrandt[2], Volker Zickermann[1,3]*

Respiratory complex I is a redox-driven proton pump, accounting for a large part of the electrochemical gradient that powers mitochondrial adenosine triphosphate synthesis. Complex I dysfunction is associated with severe human diseases. Assembly of the one-megadalton complex I in the inner mitochondrial membrane requires assembly factors and chaperones. We have determined the structure of complex I from the aerobic yeast *Yarrowia lipolytica* by electron cryo-microscopy at 3.2-Å resolution. A ubiquinone molecule was identified in the access path to the active site. The electron cryo-microscopy structure indicated an unusual lipid-protein arrangement at the junction of membrane and matrix arms that was confirmed by molecular simulations. The structure of a complex I mutant and an assembly intermediate provide detailed molecular insights into the cause of a hereditary complex I–linked disease and complex I assembly in the inner mitochondrial membrane.

INTRODUCTION

NADH:ubiquinone oxidoreductase (complex I), one of the largest membrane protein complexes known, plays a central role in aerobic energy metabolism (*1*). Redox-linked proton translocation by complex I generates ~40% of the proton motive force that drives adenosine triphosphate synthesis. Complex I dysfunction is associated with a wide range of neuromuscular and neurodegenerative diseases (*2*). Reverse electron transfer at complex I during ischemia/reperfusion releases toxic reactive oxygen species (ROS) (*3*). The physiologically important active-deactive (A/D) transition of complex I (*4*) is thought to counteract ROS-linked reperfusion injury. Structures of bacterial and mitochondrial complex I have been determined by x-ray crystallography or, more recently, by electron cryo-microscopy (cryo-EM) at increasing resolutions (*5*, *6*). Previously, we have reported the x-ray (*7*) and cryo-EM (*8*) structure of complex I from the aerobic yeast *Yarrowia lipolytica*, a model organism for eukaryotic complex I and complex I–linked diseases, at ~3.8- and 4.3-Å resolution. The arrangement of 14 central subunits in the L-shaped assembly, consisting of a membrane arm and a matrix arm, is conserved from bacteria to man. The matrix arm is divided into the N module for NADH (reduced form of nicotinamide adenine dinucleotide) oxidation and the Q module for ubiquinone (Q) reduction. The position of the Q reduction site ~20 Å above the membrane surface and its connection to the membrane bilayer by an access tunnel were unexpected. There is general agreement that Q reduction drives proton translocation, but the mechanism of long-range energy transfer from the Q reduction site to the proton pumping sites in the membrane arm remains unclear. We proposed that Q reduction generates the power stroke in the complex I catalytic cycle by a concerted rearrangement of three loops in the Q binding site and the interface region of the membrane and matrix arms (*7*, *8*). Recent experiments demonstrated that movement of the long loop connecting transmembrane helices (TMHs) 1 and 2 of subunit ND3 is required to drive proton translocation (*9*).

In mammalian complex I, the core of central subunits is surrounded by 31 accessory subunits. The function of most accessory subunits is unknown, but NDUFS4 was identified as critical for pathogenic mutations that cause Leigh syndrome, a serious neurological disorder (*10*). Various *Ndufs4* knockout mouse models were generated to study the physiology of complex I dysfunction but have failed to clarify the molecular basis of the disease (*11*). NDUFS4 stabilizes the N module (*12*) and is thought to play a role in complex I assembly linked with the assembly factor NDUFAF2. NDUFAF2 is a paralog of accessory subunit NDUFA12 (*13*). Its expression is regulated by cytokines and mutations in the NDUFAF2 gene that cause progressive encephalopathy (*13*). However, the structure of NDUFAF2 and how it interacts with complex I assembly intermediates was unknown until now. We now combine the power of cryo-EM, yeast genetics, and large-scale computer simulations to explore the structure and mechanism of complex I, the structural basis of NDUFS4-linked Leigh syndrome, and the role of assembly factor NDUFAF2 in the complex I assembly.

RESULTS AND DISCUSSION

Structure determination

Respiratory complex I from the aerobic yeast *Y. lipolytica* isolated in the detergent lauryl maltose neopentyl glycol (LMNG) was highly active (19.7 μmol NADH min^{-1} mg^{-1}). LMNG preserved activity much better than dodecyl maltoside (DDM), as reported for the ovine complex I (*14*). Single-particle cryo-EM yielded a 3.2-Å map (figs. S1 and S2, table S1, and movie S1), which enabled us to build a model containing 8057 of the 8300 residues, all cofactors, and 33 lipid molecules (Fig. 1). The model includes nearly all the regions that were unresolved or modeled as polyalanine in our previous cryo-EM structure (*8*). The high resolution permitted us to define the position of subunit NUUM unambiguously (fig. S3, E and F) and to identify

[1]Institute of Biochemistry II, University Hospital, Goethe University, Frankfurt am Main, Germany. [2]Department of Structural Biology, Max Planck Institute of Biophysics, Frankfurt am Main, Germany. [3]Centre for Biomolecular Magnetic Resonance, Institute for Biophysical Chemistry, Goethe University, Frankfurt am Main, Germany. [4]Department of Physics, University of Helsinki, Helsinki, Finland. [5]Institute of Biotechnology, University of Helsinki, Helsinki, Finland. [6]Core Facility Mass Spectrometry, Medical University of Graz, Graz, Austria. [7]Functional Proteomics, SFB815 Core Unit, Medical School, Goethe University, Frankfurt am Main, Germany. *Corresponding author. Email: zickermann@med.uni-frankfurt.de

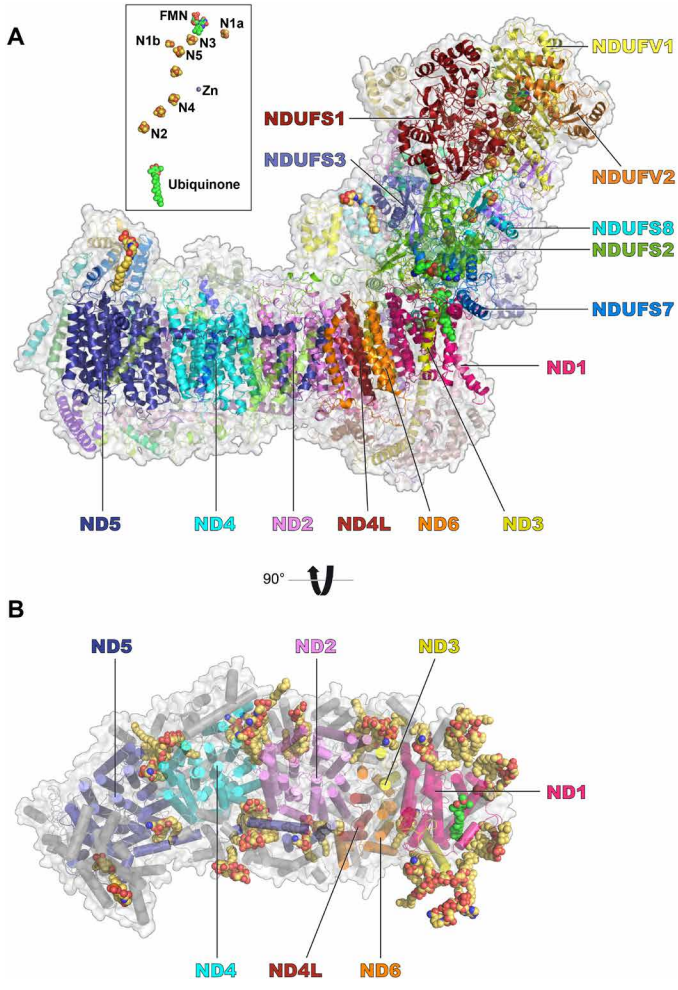

Fig. 1. Cryo-EM structure of respiratory complex I from *Y. lipolytica*. (**A**) Side view; inset shows FMN and FeS clusters [assignment of EPR signatures N1 to N5 according to (*72*)] and a Q molecule in the access pathway. (**B**) Top view with the matrix arm omitted for clarity and lipid molecules (yellow) and Q molecule (green) in sphere representation.

YALI1_D30003p as the last missing protein component of *Y. lipolytica* complex I (fig. S3C). The new NIGM subunit is an ortholog of NDUFB2 in mammalian complex I and resides at the distal tip of the membrane arm (fig. S3, D and E). Complex I of *Y. lipolytica*, thus, consists of 43 subunits. Subunits NUXM, NUUM, and ST1 clearly have no equivalent in the mammalian complex I (table S2). For ease of comparison with other complex I structures (see table S2), we have adopted the subunit nomenclature for human complex I wherever possible.

A second Q binding site

The preparation for cryo-EM contained 82.8 lipids and 1.9 Q9 molecules per complex I, as determined by quantitative mass spectrometry (fig. S4D and table S3). Q is reduced by terminal FeS cluster N2 that resides in the subunit NDUFS7 (*6*–*8*). The buried position of N2 requires a tunnel that provides access to Q from the hydrophobic bilayer core to the center of the matrix arm. The entrance to the extended access pathway is framed by TMHs 1 and 6 and the transverse helix α1 of subunit ND1. A clear density indicates a bound Q molecule in the tunnel (Fig. 2 and fig. S4C). At a center-to-center distance of 26.7 Å to FeS cluster N2, the head group is surrounded by a network of charged residues of ND1 and NDUFS7. Ile106^{NDUFS7} points to the center of the Q head group, providing hydrophobic interactions. In a recent study, exchange of this residue for alanine or phenylalanine decreased ubiquinone reductase activity (*15*). Several residues of TMHs 1 and 6 and of surface helix α1 of ND1 are in contact with the first four isoprenoid units of the Q tail. To explore the dynamics of protein Q interactions, we performed atomistic classical molecular dynamics (MD) simulations of the entire complex I structure in a membrane-solvent environment (fig. S5A). The Q binding pocket rapidly filled with water, and the Q head group engaged in tight hydrogen bonding interactions with three to four water molecules (fig. S5, B and C). In contrast, the isoprenoid units of Q interacted with nonpolar amino acid residues (fig. S6), in agreement with the cryo-EM structure (Fig. 2).

A tightly bound Q in complex I had been proposed earlier on the basis of biochemical data (*16*), and in our lower-resolution maps, we had identified a conspicuous nonprotein density in the Q tunnel (*8*). The strong density we observe in our present high-resolution map indicates firm binding and significant occupancy of the site by a Q molecule. The second Q molecule indicated by the near 2:1 stoichiometry was not identified in the map, most likely because it is distributed over a number of different low-occupancy sites. However, we cannot exclude that it is randomly distributed in the detergent micelle. The Q binding site in our cryo-EM structure matches a binding position in the extended Q tunnel predicted by molecular simulations (*17*, *18*). It is different from the Q reduction site near the FeS cluster N2 characterized in our previous work (*8*) and appears to represent a "stepping stone" for a substrate molecule on its way between the active site and the lipid bilayer (Fig. 2). While substrate exchange with the Q pool in the membrane must be possible on a short time scale, occupancy of the observed site indicates that an energy barrier restricts diffusion from the tunnel into the membrane. Considering the length of the access pathway and the restricted mobility in the tunnel, the height of the energy barrier (*17*) is a critical parameter for complex I turnover. As an alternative to a substrate molecule traveling the complete length of the access path in each reaction cycle, a bound Q molecule might act as a mobile carrier shuttling electrons between the Q reduction site near N2 and another site in the tunnel, from where it could transfer electrons to a Q molecule of the membrane Q pool at the protein-lipid interface (*19*). However, extended site-directed mutagenesis of the Q binding pocket and kinetics measured with different short-chain Q analogs suggest that added substrate molecules are directly reduced by cluster N2, which would argue against a shuttle mechanism involving a bound Q (*20*). Moreover, there are no robust structural or biochemical data supporting a Q binding site on the surface of ND1. Such a site would be required to enable electron transfer between the shuttling and substrate Q molecules. In any case, tight binding of a Q molecule at some distance from the Q reduction site is unexpected. The functional significance of the observed site in the catalytic cycle of complex I will be evaluated in future work.

Complex I lipids

We found clear density for 8 phosphatidylcholine, 15 phosphatidylethanolamine, 6 cardiolipin (CL), 4 phosphatidylinositol, and 2 detergent molecules that were built into the map (Figs. 1 and 3A and fig. S2). A comparison with lipid binding sites in mouse (*21*) and ovine (*14*) complex I is shown in fig. S7. Matrix arm subunits NDUFA9,

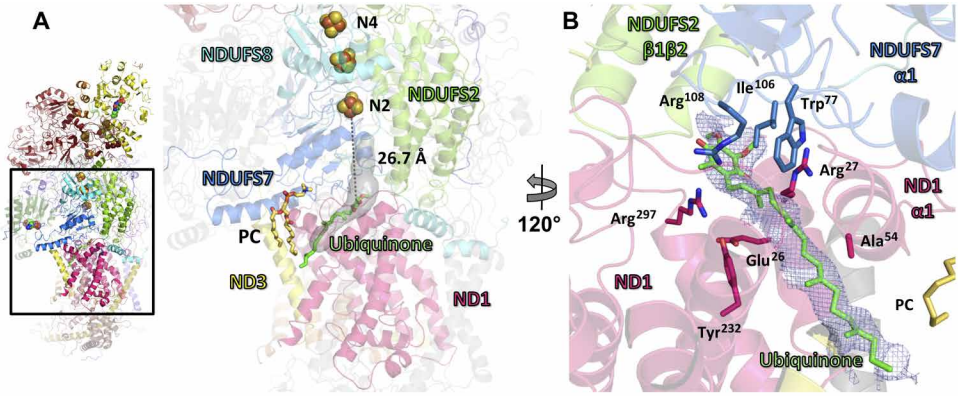

Fig. 2. Q molecule in the access pathway. (**A**) Slice of membrane arm/matrix arm interface viewed from the back of complex I, a cavity (gray) opens in ND1 and permits access of Q from the bilayer to the Q reduction site near cluster N2. (**B**) Selected residues near the Q binding site (density map drawn at 1.0 σ contour level in blue mesh; Q molecule, green sticks) (see fig. S4).

NDUFS7, and NDUFA12 together bind the head groups of six lipid molecules. The head groups in this lipid cluster form a plateau that is out of plane with lipids in the central part of the membrane arm, indicating a local deformation of the planar bilayer geometry (Fig. 3A). The N-terminal helix α1 of subunit NDUFS7 at the center of the plateau is positioned directly above the entrance to the Q tunnel (fig. S8A), suggesting that modulation of the membrane surface might facilitate the access of substrate Q from the membrane. Two lipids flanking this helix were also observed in the structure of complex I from mouse (fig. S7) (*21*). The matrix arm/membrane arm interface is lined by eight amphipathic helices of central and accessory subunits (fig. S8B). A CL binding site between amphipathic helices from NDUFS8, NDUFA12, and NDUFA7 is consistently observed in mouse, ovine, and yeast complex I (*14*, *21*). Overall, lipid tails in the interface region are in a notably tilted orientation. In our atomistic simulations, lipids in this region showed a clear tendency toward a horizontal orientation, in agreement with the cryo-EM structure (Fig. 3B and fig. S9, A and B). Typically, lipid tails close to the protein surface in this region deviated by 50° to 90° from the membrane normal (fig. S9). The tilt of lipid fatty acid chains was much more pronounced in the matrix leaflet and resulted in a local thinning of the membrane in the interface region (Fig. 3 and fig. S9). Adsorption of an amphipathic helix to a membrane surface can distort the default geometry of a lipid bilayer (*22*). The peptide displaces the lipid head groups and thereby destabilizes the lipid tail region (fig. S9D). The lipid tails in the destabilized region tend to wrap around the hydrophobic face of the amphipathic helix, reducing the thickness of the lipid bilayer locally. The effect on the lipid head groups exerted by the amphipathic helices in the membrane arm/matrix arm junction is enhanced by the tilted TMHs of ND1 and adjacent accessory subunits NDUFA1 and NDUFA13 (fig. S9B). The local membrane thinning by up to 10 Å could help to guide Q to the entry portal of the access pathway in ND1 (Fig. 3B). Considering that complex III oxidizes Q on the distal side of the membrane, we propose that the thinner lipid bilayer reduces the energy barrier for flipping the polar Q head group from one leaflet to the other, which would be required for fast alternate access to the active sites of respiratory chain complexes I and III (*23*).

Active and deactive complex I

Complex I from mammals or the yeast *Y. lipolytica* undergoes a reversible transition between an active A form and a deactive D form (*4*, *24*). In the absence of substrate, the A form spontaneously converts to the D form. The D form is stabilized by divalent cations and is arrested irreversibly upon modification of a cysteine residue in subunit ND3 by the sulfhydryl reagent N-ethylmaleimide (NEM) (*25*). Only the D form is sensitive to NEM. The conversion from D to A is induced by addition of the substrates NADH and oxidized Q. Bovine complex I was isolated as a mixture of A and D forms that gave rise to different three-dimensional (3D) classes in the cryo-EM analysis (*26*). By contrast, the *Y. lipolytica* complex yielded only one major 3D class (fig. S1) that we assign to the D form based on its reactivity toward Mg^{2+} or NEM (fig. S10).

Two alternative models, the "blocked Q site model" and the "unfolded Q site model," were recently developed for structural changes associated with the A/D transition of complex I (*7*, *27*). The blocked Q site model proposed that Q access to its reduction site is restricted in the D form by the loop connecting the first two β strands of the N-terminal β sheet of NDUFS2 (*7*). However, the conformation of the complete β1-β2 loop in our present cryo-EM structure is different and matches that of the corresponding loop in the murine A form (*21*) almost perfectly (fig. S11A). Therefore, it seems unlikely that the difference between A and D is due to conformational changes of the β1-β2 loop in the NDUFS2 subunit, rendering the blocked Q site model obsolete. The different loop conformation observed in *Y. lipolytica* complex I by x-ray crystallography and cryo-EM may be due to the different detergents used for protein purification (fig. S11B). The alternative unfolded Q site model suggested that the D form of mammalian complex I is distinguished by disordered loops in the Q binding pocket and surrounding protein structure (*21*, *27*). The unfolded Q site model does not apply to *Y. lipolytica* complex I because we observe side-chain densities for nearly all the loop regions that were disordered in the D form of murine or bovine complex I (fig. S11, C and D). The resolved regions include the TMH 5/6 loop of ND1, the TMH 1/2 loop of ND3, the β1-β2 loop of the NDUFS2, and the C-terminal domain of NDUFA9. The major structural rearrangement in mammalian complex I reflects a high-energy barrier between the A and D forms (*24*). This barrier is much lower in the yeast complex (*24*), which would require a less extensive conformational change. Restructuring of the Q binding site in mammalian complex I was suggested to be induced by substrate binding (*27*). The observed Q molecule at the entrance of the Q binding pocket might stabilize this region in the D form of the yeast enzyme.

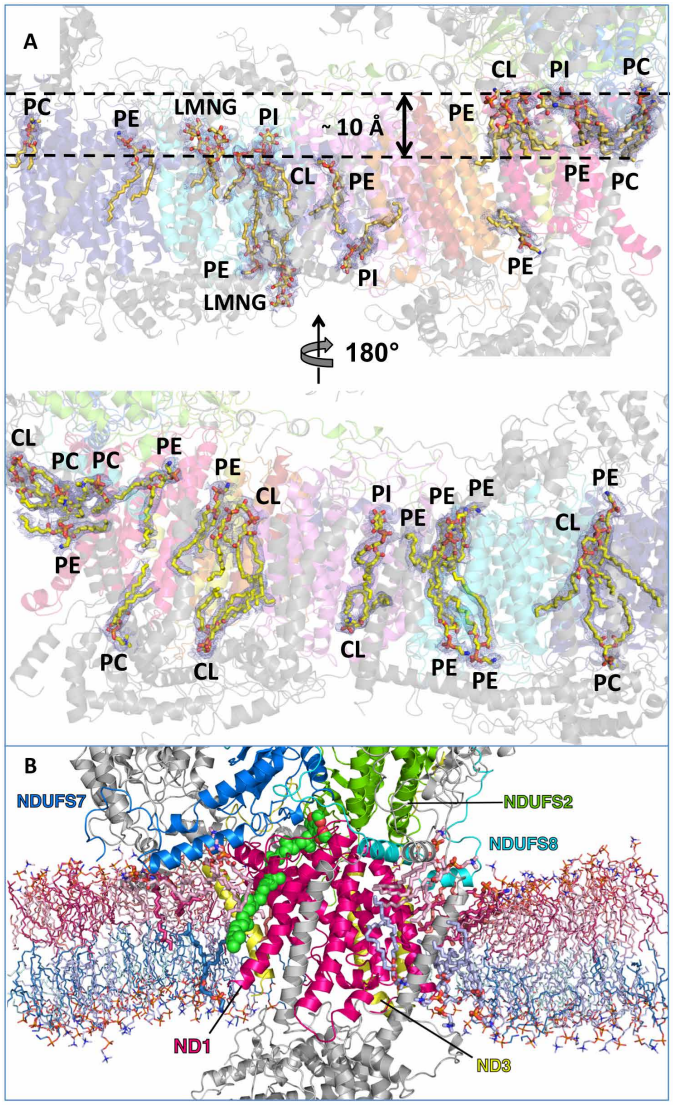

Fig. 3. Lipid binding sites and unique protein lipid arrangement at the junction of the membrane and matrix arms. (**A**) Lipid and detergent binding sites in the membrane arm. CL, cardiolipin; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PI, phosphatidylinositol. (**B**) Cross section of a simulation snapshot reveals that the lipid bilayer thickness is reduced by ~10 Å in the vicinity of protein, apparently due to the unusual local protein architecture. Q is shown as spheres colored by atom (carbon, green; oxygen, red). Lipids in the two leaflets are colored by lipid type. PC/PE/CDL in the matrix leaflet are shown in shades of pink, and those in the other leaflet in shades of blue. Highly bent lipid tails next to protein are drawn as sticks. Subunits are colored as in Fig. 1, and accessory subunits are gray.

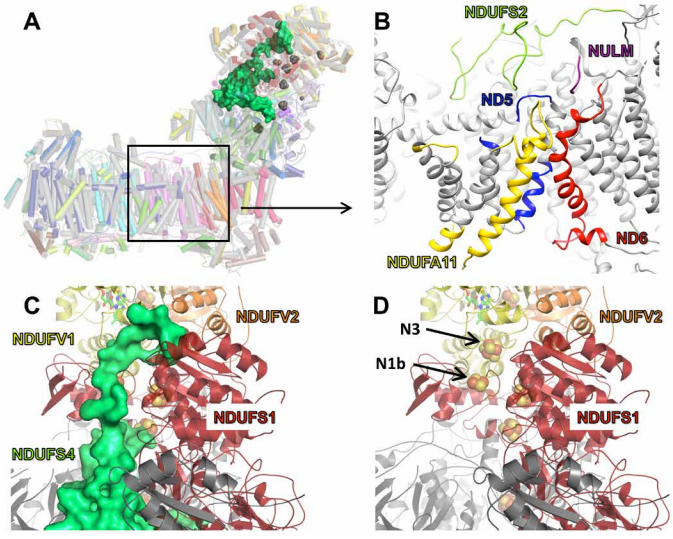

Fig. 4. Mutant lacking NDUFS4. (**A**) Overlay of wild type (gray, NDUFS4 green surface representation) and mutant (cartoon, color as in Fig. 1). (**B**) Zoom into membrane arm. Disordered protein in the mutant is highlighted by color. (**C**) Overlay of NDUFS4 on the matrix arm of the mutant. (**D**) FeS clusters N1b and N3 in the mutant are more exposed to solvent than in the fully assembled complex (see main text).

Structure of a complex I mutant without the disease-related accessory subunit NDUFS4

The accessory subunits (fig. S3, A and B) are generally thought to be dispensable for the core function of complex I because they are not present in the simplest 14-subunit "minimal" prokaryotic complex I. However, orthologs of accessory subunits NDUFS4, NDUFS6, and NDUFA12 were identified as components of complex I in α-proteobacteria, which are thought to be the ancient predecessors of mitochondria (*28*). In eukaryotes, NDUFS4, NDUFS6, and NDUFA12 act together in the final step of complex I assembly (*29*). Different types of mutations in the gene encoding NDUFS4 have been found to cause Leigh syndrome (*10*). Our yeast genetic model system enables us to study the structural basis of NDUFS4-linked Leigh syndrome in detail. In a *Y. lipolytica ndufs4*Δ strain, complex I levels dropped to ~40% of wild type, indicating a less efficient assembly process, and complex I lacking NDUFS4 was associated with the assembly factor NDUFAF2 (*12*). However, NDUFAF2 was lost during protein purification in detergent, suggesting weak interaction. Biochemical and spectroscopic analysis of the mutant complex indicated reduced electron transfer activity, increased ROS production, and prominent changes in the electron paramagnetic resonance (EPR) spectra of FeS clusters N1b and N3 (*12*).

We determined the cryo-EM structure of the mutant complex I at 4.0-Å resolution (Fig. 4, fig. S1, and movie S2) in DDM and compared it to the present wild-type structure presented here and to our recent 4.3-Å cryo-EM map of complex I in the same detergent (*8*). The mutant structure contains all complex I subunits except NDUFS4. Despite changes in the EPR spectrum, there is no indication of a loss or major change in position of any of the eight FeS clusters. The binuclear FeS cluster N1b is located at the bottom of a cleft of NDUFS1 that is traversed by a C-terminal section of NDUFS4 (Fig. 4). The C-terminal end of NDUFS4 runs along the interface of N-module subunits NDUFS1, NDUFV1, and NDUFV2. The C-terminal residue Lys^{161} of NDUFS4 points toward cluster N3 in NDUFV1. In the absence of NDUFS4, both cluster binding sites are exposed and solvent accessible (Fig. 4D). To analyze this in more detail, we performed solvent-accessible surface area (SASA) calculations on the cryo-EM structures and found that the (~7 Å) region around clusters N3 and N1b is more exposed to solvent by ~87 and 228 $Å^2$, respectively, in mutant than in wild type. Similarly, an increase in surface area is also observed for subunits NDUFS1 and NDUFV1 by ~1935 and ~487 $Å^2$, respectively, in the complex I lacking NDUFS4 (see the section on "Computational methods"). The change in chemical environment explains both the different EPR spectra and the increased ROS formation and decreased electron transfer activity of the mutant complex.

Previous studies with patient material had suggested that complete absence of the N module is a typical characteristic of NDUFS4-linked Leigh syndrome and a valuable diagnostic marker in BN-PAGE (blue native polyacrylamide gel electrophoresis) analysis (*30*). Our wild-type structure shows that NDUFS4 is an important connecting element in the matrix arm. In the mutant, domains of NDUFV1, NDUFV2, NDUFS1 and the attached accessory subunit NDUFA2 are tilted backward by 1°. However, the N module was not lost during protein purification. The N module is unlikely to detach in vivo because intact tissue of *Ndufs4* knockout mice exhibited up to 50% of wild-type levels of rotenone-sensitive oxygen consumption (*31*), in agreement with our activity measurements (*12*).

Structural changes in the mutant are not limited to the vicinity of NDUFS4 in the matrix arm. The hydrophobic accessory subunit NDUFA11 in the center of the membrane arm shields the C-terminal membrane anchor of the long lateral helix of ND5 (Fig. 4B). In the mutant, density for the N-terminal half of NDUFA11, for the end of the lateral helix, and for part of the following TMH 16 of ND5 is weak or absent, indicating disorder. Moreover, structural integrity of TMH 4 and adjacent loops of ND6 and the long N-terminal extension of NDUFS2 on the surface of the membrane arm is compromised. The same changes were observed in two independent preparations of the mutant but not in the wild-type control. To address the functionality of the membrane arm, we measured proton translocation of the mutant after reconstitution into lipid vesicles. ACMA (9-amino-6-chloro-2-methoxyacridine) quenching was slower but reached the same level as the wild-type control, indicating that the mutant retained most of its proton pumping capacity, albeit at a slightly reduced rate (fig. S12). Distortion of NDUFA11 and the protein structure around it is in line with the observation that supercomplex formation stabilizes the NDUFS4 mutant in a mouse model of *Ndufs4*-linked complex I dysfunction (*32*). In complex I supercomplexes, NDUFA11 interacts with three subunits of complex III (*33*). We propose that binding to complex III mitigates structural disorder in the membrane arm, which would at least partially rescue the function of the mutant complex.

NDUFS4-linked Leigh syndrome is thus caused by a combination of impaired complex I assembly and changes in the structure of assembled mutant complex I. Specifically, absence of NDUFS4 distorts two of the FeS cluster binding sites in the N module and parts of the membrane arm. Compromised NADH dehydrogenase function interferes with mitochondrial energy metabolism and decreases the NAD^+/NADH ratio, which triggers secondary effects including lactic acidosis. ROS are released at compromised cluster binding sites or because electrons are trapped at the flavin site and might cause further decay of mitochondrial functions in the mutant.

Cryo-EM structure of a complex I assembly intermediate

In contrast to the mutant lacking NDUFS4, deletion of the gene for accessory subunit NDUFS6 in *Y. lipolytica* resulted in the accumulation of a complex I assembly intermediate that tightly bound assembly factor NDUFAF2 but lacked accessory subunits NDUFS6 and NDUFA12 (*34*). Mutations in the NDUFS6 gene cause lethal neonatal deficiency in human complex I (*35*). Reintroduction of NDUFS6 by in vitro import into patient mitochondria lifted a block in the assembly process and facilitated formation of complete complex I from an intermediate harboring assembly factor NDUFAF2 (*36*).

NDUFA12 and NDUFAF2 are paralogs and share a longer N-terminal and a shorter C-terminal region of similar sequence, interrupted by a stretch of sequence in the assembly factor (fig. S13) (*37*). We isolated the NDUFAF2 assembly intermediate in LMNG and determined its structure at 3.3-Å resolution by cryo-EM (Fig. 5, fig. S1, and movie S3). Overall, the structure matched the fully assembled complex I, but a major part of the matrix arm was tilted sideways (Fig. 5B), resulting in a 6-Å displacement at the tip of the N module. At the position occupied by the N-terminal core domain of NDUFA12 in wild-type complex I, we modeled the corresponding domain of NDUFAF2 in the assembly intermediate (Fig. 5 and fig. S13). The resulting structure clearly proves that in the fully assembled complex, NDUFA12 replaces NDUFAF2, as proposed on the basis of evolutionary correlation (*37*). The next part of the assembly factor folds into a long helix pointing toward the distal part of the matrix arm (Fig. 5 and fig. S13). In an overlay with the wild-type structure, this helix clashes with the position of a loop that connects the two major domains of subunit NDUFS6. Therefore, NDUFAF2 prevents binding of NDUFS6 while it occupies the NDUFA12 position. The extensive interaction between NDUFA12 and NDUFS6 implies that these two subunits operate as one unit. Residues 122 to 223 of NDUFAF2 are not resolved, but the C terminus binds to subdomains 1 and 2 of the large C-terminal domain of NDUFS1. In native complex I, the C terminus of NDUFA12 binds in the same position (Fig. 5).

In the assembly intermediate, density extending along the interface region including the C-terminal part of subunit NDUFA9, the long TMH 1/2 loop of ND3, and a large part of NDUFA7 is poorly ordered or absent (fig. S8C). The N- and C-terminal helices of NDUFS7 are shifted, and lipids in the interface region are unresolved, in contrast to lipids in the major part of the membrane arm. The Q content of the preparation is only 0.5 nmol per nmol complex I as determined by quantitative mass spectrometry. There is no density

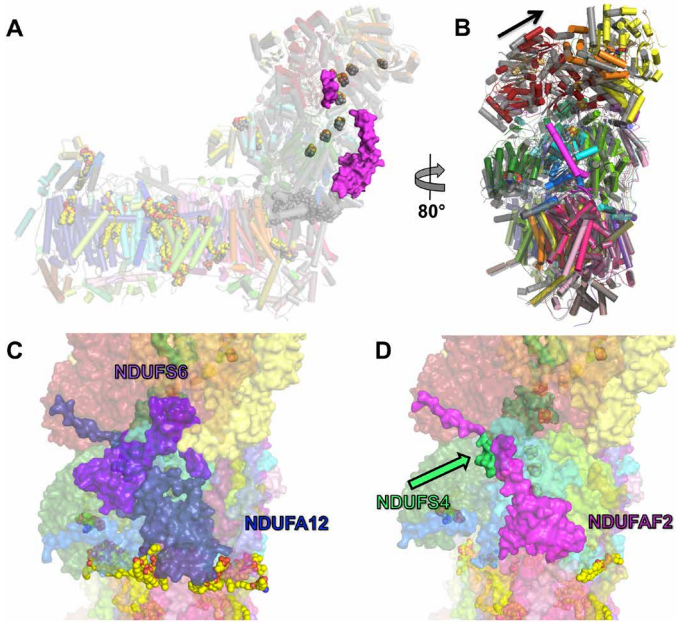

Fig. 5. Cryo-EM structure of a complex I assembly intermediate. (**A**) Overlay of wild type (gray) and assembly intermediate (color) with assembly factor NDUFAF2 (purple surface). (**B**) Overlay as in (A) viewed from the back. The arrow indicates the tilted matrix arm. (**C**) Membrane arm/matrix arm interface of wild-type complex I viewed as in (B) with NDUFA12 and NDUFS6 highlighted. (**D**) Assembly intermediate with NDUFAF2 highlighted, viewed as in (B); compare fig. S13.

for a Q substrate in the access pathway. NDUFA12 is the cornerstone of the membrane arm/matrix arm interface, and its two surface helices are critical for the unusual lipid arrangement in complex I (figs. S8 and S9). We propose that binding of NDUFAF2 instead of NDUFA12 prevents formation of the native interface structure and, thus, restricts Q access before the assembly process is completed. This would minimize the risk of ROS formation by reverse electron flow from reduced Q to an incomplete electron transfer chain.

We previously found that the EPR signal of FeS cluster N4 (Fig. 1) was absent in the assembly intermediate (*34*). Our maps show strong density for the N4 cluster, suggesting that loss of the N4 signal in the earlier study was due to a change of chemical environment rather than absence of the cluster. In the wild type, the C-terminal Zn-binding domain of the deleted subunit NDUFS6 is situated at the interface of the N and Q modules, close to the binding site of N4. Therefore, we propose that the loss of the EPR signal is due to the N4 site being solvent accessible and that the assembly factor has no critical role in the insertion of this FeS cluster.

Immunological experiments had suggested that in *Neurospora crassa* NDUFAF2 associates with a complex I subassembly containing the central subunits of the Q module (NDUFS2, 3, 7, and 8) but lacking the complete N module (*29*). In agreement with this finding, a recent complexome profiling study on the assembly pathway of human complex I demonstrated that NDUFAF2 attaches to a subassembly that includes part of the membrane arm, ND1, and the central subunits of the Q module (*38*). The N module (including NDUFS1, FV1, and FV2) is added later as a separate building block. It has been shown that a concerted interplay of NDUFS4, NDUFS6, and NDUFA12 is required to release the assembly factor (*29*, *36*). Our cryo-EM structure (Fig. 5) shows that the C-terminal end of the assembly factor anchors the N module to the Q module. Our data therefore concur with the concerted action of NDUFA12 and NDUFS6 in complex I assembly. The role of NDUFS4 in this process seems to be more intricate. NDUFAF2 is associated with complex I in patients carrying mutations in NDUFS4 (*39*). In agreement with an earlier study (*29*), we observed that in mitochondrial membranes, NDUFAF2 remained associated with complex I in our *ndufs4*Δ strain (*12*). The assembly factor was lost during purification of this mutant protein complex but is an integral part of the NDUFS6 mutant structure described here. Since NDUFA12 and NDUFS6 assemble in the NDUFS4 mutant (see above), there must be an additional but weaker attachment site in the absence of NDUFS4, probably at the unoccupied position of this subunit. In our structure of the assembly intermediate, NDUFS4 is present, and a 102-residue stretch of the assembly factor is unresolved. We propose that at least a part of this disordered segment is displaced by NDUFS4 when the complex assembles. Note that an extension of NDUFS4 reaches through a narrow opening between NDUFS1 and NDUFS3 to the side of the peripheral arm decorated by NDUFAF2 (Fig. 5D). NDUFAF2 binding in this region might provide a platform for the attachment of NDUFS1, which is then firmly anchored by the C-terminal end of the assembly factor (Fig. 6). Once NDUFS1 is in the correct position, NDUFAF2 detaches by joint action of NDUFA12, NDUFS4, and NDUFS6. In NDUFS4- and NDUFS6-linked diseases, complex I assembly is compromised by disturbing the interplay of the three subunits and the assembly factor. Moreover, binding of NDUFAF2 to mutant complex I might reduce the amount of free assembly factor that is available to support a sufficient level of complex I assembly. In our NDUFS6 mutant, NDUFS4 binds and probably sets the stage for release of NDUFAF2, but the assembly factor remains firmly associated with complex I because NDUFS6, and therefore the NDUFS6/NDUFA12 tandem, is absent.

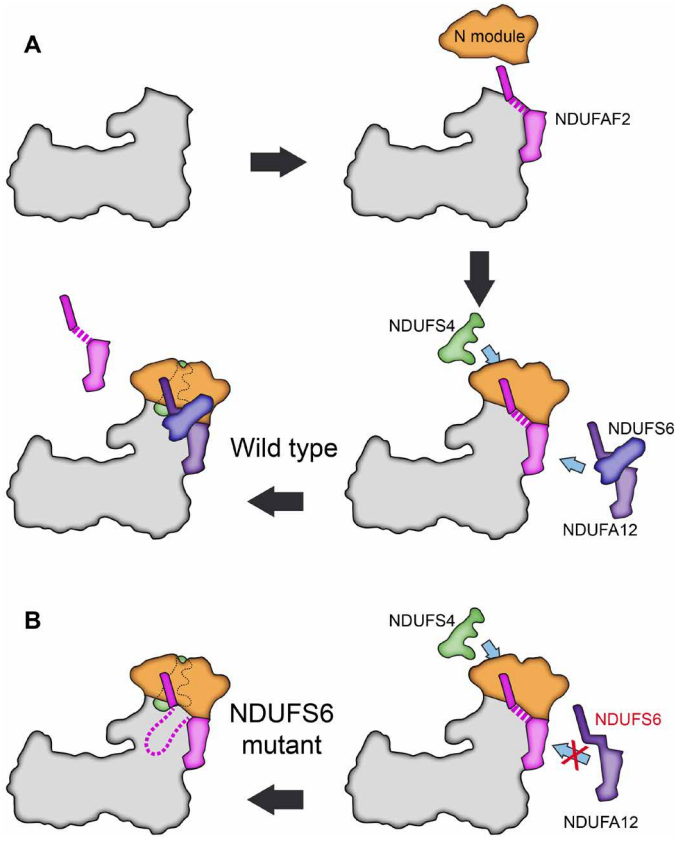

Fig. 6. Function of the assembly factor NDUFAF2. (**A**) Assembly factor NDUFAF2 associates with a complex I assembly intermediate that lacks the N module (NDUFS1, NDUFV1, and NDUFV2). NDUFAF2 provides a platform for N module attachment (dashed line) and anchors NDUFS1 with its C-terminal tail. In the wild type, the concerted interplay of NDUFS4, NDUFS6, and NDUFA12 releases NDUFAF2. NDUFS6 and NDUFA12 form a structural unit that is obligatory for function. (**B**) In the NDUFS6Δ mutant, NDUFS4 can bind and then detaches a part of NDUFAF2 (dotted line), but the assembly factor remains firmly bound because NDUFS6 is absent.

In summary, our new high-resolution cryo-EM structure of respiratory complex I, together with our large-scale molecular simulations, sheds new light on the unique interactions of the 1-MDa membrane protein complex with Q and the lipid bilayer of the membrane. Cryo-EM structures of complex I variants generated with our yeast genetic model system offer clues on the structural basis of NDUFS4-linked Leigh syndrome and the function of assembly factor NDUFAF2. NDUFAF2 anchors the N module to nascent complex I and is released by joint action of three accessory subunits. Combining cryo-EM, molecular simulations, and yeast genetics opens a new avenue for investigating complex I, the most intricate molecular machine of the respiratory chain.

MATERIALS AND METHODS

Purification of complex I variants from *Y. lipolytica*

Complex I was purified from *Y. lipolytica* strain GB30 carrying a deletion for the sulfur transferase (ST1) gene to enhance sample homogeneity (*40*). ST1 is a substoichiometric component of *Y. lipolytica* complex I, and the enzyme complex is fully functional in

the absence of this polypeptide. Protein purification was carried out by His-tag affinity and size exclusion chromatography as described in (*41*), but on the Ni-NTA Sepharose column, the detergent was exchanged from 0.025% DDM to 0.025% LMNG. The complex I assembly intermediate accumulating in the Δ*numm* strain (*34*) (NUMM corresponds to NDUFS6; see table S2) was purified in the same way and polished by ion exchange chromatography on a MonoQ column. The complex I mutant from strain Δ*nuym* (*12*) (NUYM corresponds to NDUFS4; see table S2) was purified in the same way as the GB30 complex but without the detergent exchange.

Activity measurements

NADH:decylubiquinone (DBQ) oxidoreductase activity was measured in 20 mM Na^+-Mops (pH 7.2), 50 mM NaCl, 2 mM KCN, and 0.025% LMNG using 150 μM DBQ and 100 μM NADH as substrates. The residual activity in the presence of 2 μM 2-*n*-decyl-quinazolin-4-yl-amine (DQA) was subtracted. For assignment of our preparation to the A or D form, we tested the sensitivity of complex I toward 5 mM $MgCl_2$ or 2 mM NEM. Activity measurements were carried out in 50 mM tris-HCl (pH 8.5), 80 mM KCl, 0.2 mM EDTA, 1.1 mM NaCN, and 0.025% LMNG with 150 μM DBQ and 100 μM NADH (*9*).

Proton pumping assay with reconstituted complex I

Asolectin (10 mg ml^{-1}) from soybean was solubilized in 1.6% *n*-octyl glucopyranoside, 20 mM K^+-Mops (pH 7.2), and 80 mM KCl. Purified complex I (200 μg) was reconstituted into proteoliposomes at a protein-to-lipid ratio of 1:50 (w/w) according to (*42*), but with 30-min intervals for the addition of BioBeads (Bio-Rad). Proteoliposomes were centrifuged at 100,000*g* (1 hour, 4°C) and resuspended in 200 μl of 20 mM K^+-Mops (pH 7.2) and 80 mM KCl. Proteoliposomes containing 5 μg of complex I were added to 2 ml of buffer [20 mM K^+-Mops (pH 7.2), 50 mM KCl, and 0.5 μM valinomycin] in a stirred cuvette. Proton pumping was monitored by ACMA fluorescence quenching. The fluorophore was added to a final concentration of 0.5 μM. Measurements were performed in a Shimadzu RF-5001 fluorimeter at an excitation wavelength of 430 nm and an emission wavelength of 475 nm (bandpass, 5 nm each; integration time, 1 s) at 30°C. The reaction was started by the successive addition of 60 μM DBQ and 100 μM NADH. Last, 1 μM FCCP was added to collapse the pH gradient.

Mass spectrometry

Purified complex I from *Y. lipolytica* was denatured in 2 M guanidine hydrochloride. Proteins were digested with trypsin (sequencing grade, Promega) and analyzed by mass spectrometry using Thermo Scientific Q Exactive Plus equipped with an ultra-high performance liquid chromatography (UHPLC) unit (Thermo Scientific Dionex UltiMate 3000). Peptides and proteins (false discovery ratio, <0.01) were identified by PEAKS Studio 7.0 (Bioinformatics Solutions Inc.) using all 6456 entries of *Y. lipolytica* from UniProtKB (released in July 2016) supplemented with the information of YALI1_D30003p.

Sample extraction for lipid mass spectrometry

Lipid and Q9 were extracted with minor modification as described (*43*). Briefly, 1.5 ml of methanol containing 5 nmol phosphatidylethanolamine (PE) 24:0 as extraction control standard and 5 ml of methyl tert-butyl ether (MTBE) were added to the *Y. lipolytica* protein pellets. After shaking, addition of 1.25 ml of deionized water, and centrifugation, the upper organic phase was collected and the lower aqueous phase was extracted a second time with MTBE/methanol/water (10/3/2.5, v/v/v). The combined organic extracts were dried in a stream of nitrogen and redissolved in 500 μl of chloroform/methanol (1/1, v/v). Twenty-five microliters of this extract was supplemented with 0.2 nmol phosphatidylcholine (PC) 25:0, 17:4, 43:6, and 31:1; 0.3 nmol PE 25:0, 37:4, 43:6 and 31:1; 0.3 nmol phosphatidylserine (PS) 25:0, 37:4, 43:6, and 31:1; 0.135 nmol lysophosphatidylcholine 17:1; 0.2 nmol phosphatidylinositol 25:0, 37:4, 43:6, and 31:1; 0.89 nmol CL 57:4; 0.81 nmol CL 61:1; 0.71 nmol CL 80:4; 0.68 nmol CL 86:4 (all lipid species are quantitative LM standards from Avanti Polar Lipids, USA); and 5 nmol Coenzyme Q10 (Merck, Germany), evaporated and redissolved in 100 μl of isopropanol/chloroform/methanol (95/5/5, v/v/v).

Reverse-phase liquid chromatography mass spectrometry analysis

Phospholipid and ubiquinone were quantified as described (*44*). For chromatographic separation, a BEH C8 reversed-phase column (100 mm by 1 mm, 1.7 μm) from Waters (Waters, USA) was used at a column oven temperature of 50°C in a Dionex UltiMate 3000 RS UHPLC system (Thermo Fisher Scientific Inc., USA). Gradients were eluted with 50% eluent A, containing 1% of 1 M aqueous ammonium formate and 0.1% formic acid, raising to 100% eluent B, containing acetonitrile/isopropanol (5/2 v/v) with the same additives, over 40 min, followed by 10 min at 100% eluent B. The flow rate was set to 150 μl min^{-1}, and the autosampler temperature was 8°C.

For ionization and analysis, an Orbitrap Velos Pro Mass Spectrometer (Thermo Fisher Scientific Inc., USA) in data-dependent acquisition mode with HESI-II ion source was used. Samples were measured in positive electrospray mode at 4.5-kV source voltage, 275°C source temperature, and 300°C capillary temperature in full scan [mass/charge ratio (*m*/*z*), 320 to 1.000]. In negative electrospray mode, the source voltage was at 3.8 kV, the source temperature was at 325°C, and the capillary temperature was at 300°C in full scan (*m*/*z* 320 to 1600). Resolution was set at 100,000 (*m*/*z*, 400). The 10 most abundant ions of the spectrum were sequentially fragmented with collision-induced dissociation mode in the ion trap using He as collision gas (normalized collision energy, 50; isolation width, 1.5; activation Q, 0.2; and activation time, 10). Exclusion time was set at 11 s. Data were processed using a Lipid Data Analyzer described in (*45*) with a cutoff filter at 5% quantified with the added internal standards as a one-point calibration and normalized to the protein weight. Annotation was according to the official LIPID MAPS shorthand nomenclature (*46*).

Cryo-EM sample preparation and imaging

Complex I or assembly intermediate purified in LMNG was applied at a concentration of 1.5 to 2 mg/ml to freshly glow-discharged C-flat 1/1 or 1.2/1.3 holey carbon grids (Protochips, USA). Samples of the ΔNDUFS4 mutant, containing DDM as detergent, were diluted to a final concentration of 2.7 mg/ml. Grids were automatically blotted for 12 to 14 s in a Vitrobot Mark IV (Thermo Fisher Scientific Inc., USA) at 10°C and 80% humidity [drain and wait time, 0 s; blot force, −2 a.u. (arbitrary units)], plunge frozen in liquid ethane, and stored in liquid nitrogen until further use.

Cryo-EM data of wild-type complex I and the assembly intermediate were collected automatically on a 300-kV FEI Titan Krios microscope (Thermo Fisher Scientific Inc., USA) using EPU software (Thermo Fisher Scientific Inc., USA) on a K2 Summit detector (Gatan) in counting mode. Cryo-EM images were acquired at a nominal

magnification of ×125,000 with a calibrated pixel size of 1.077 Å and a total exposure of 30 to 40 $e^-/Å^2$, at defocus ranges from −1.5 to −2.5 μm. Cryo-EM data of the ΔNDUFS4 mutant were collected automatically with Leginon (*47*) on an FEI Tecnai Polara microscope (Thermo Fisher Scientific Inc., USA) at 300 kV equipped with a Gatan K2 direct electron detector operating in counting mode. Cryo-EM images were acquired at a calibrated pixel size of 1.09 Å (nominal magnification of ×200,000), a defocus range from −2.0 to −3.0 μm, an exposure time of 8 s, and a total electron exposure on the specimen of ~60 $e^-/Å^2$.

Image processing

A set of 4043 dose-fractionated micrographs from wild-type complex I were subjected to motion correction and dose-weighting of frames by MotionCor2 (*48*). The micrograph-based contrast transfer function (CTF) was determined by CTFFIND4.1, and the resulting images were used for further analysis with the software package RELION-3.0 (*49*). Particles were picked using Autopick within RELION-3.0, yielding 666,666 particles, which were extracted with a box size of 456 by 456 pixels. Initial reference-free 2D classification was performed with fourfold binned data to remove false positives and imperfect particles. Visual selection of class averages with interpretable features resulted in a dataset of 664,841 particles, which were further sorted by 3D classification with a previous cryo-EM map of complex I from *Y. lipolytica* (*8*) low-pass filtered to 30 Å as an initial reference for the first round of 3D classification. The resulting best 3D class of 297,066 particles was autorefined, yielding a map with a 3.5-Å resolution. In a final step, the resolution was improved using algorithms for CTF refinement and Bayesian polishing implemented in RELION-3.0 (*49*). The final map at 3.2-Å resolution was sharpened using an isotropic B-factor of −66 $Å^2$ (fig. S1 and table S1). Local map resolution was estimated with ResMap (http://resmap.sourceforge.net) (fig. S2) (*50*). All resolutions were estimated using the 0.143 cutoff criterion (*51*) with gold standard Fourier shell correlation (FSC) between two independently refined half maps (*52*).

A similar image processing workflow was applied to the complex I assembly intermediate and ΔNDUFS4 mutant. For the assembly intermediate, a total of 250,406 particles were extracted from 1386 motion-corrected micrographs. 3D classification resulted in a final dataset of 112,418 particles, which yielded a 3.3-Å map after CTF refinement and Bayesian polishing. An isotropic B-factor of −50 $Å^2$ was applied (fig. S2). For the ΔNDUFS4 mutant, CTF parameters were estimated by Gctf within the RELION-3.0 workflow (*49*). A total of 418,740 particles were extracted from 5964 micrographs after discarding images that showed contamination or poor CTF estimation. The extracted particles were subjected to reference-free 2D classification and sorted by 3D classification. The best class, comprising 145,767 particles, was autorefined and CTF refined, and Bayesian polishing was applied in RELION-3.0. The final map of 4.0-Å resolution was sharpened with an isotropic B-factor of −30, and the local resolution was estimated with ResMap (fig. S2).

Model building

The cryo-EM structure of deactive *Y. lipolytica* complex I [Protein Data Bank (PDB) ID 6GCS] was used as template for building the wild-type complex I in COOT (*53*). Full-atom models of all accessory subunits were built. A previously unidentified density with one TM helix fitted the sequence of NUUM, a subunit without homolog in mammalian complex I. The density assigned as NUUM in the previous map (*8*) occupies the position of the mammalian subunit NDUFB2. A BLAST search in fungi yielded several homologs, including one in the yeast *Pichia pastoris* (*Komatagaella phaffii*). A second BLAST search with this yeast homolog against the *Y. lipolytica* proteome (https://blast.ncbi.nlm.nih.gov) yielded the sequence of YALI1_D3003g, which fitted the density. We refer to this previously unidentified *Y. lipolytica* complex I subunit as NIGM. Lipids were built wherever possible, and the fatty acid tails truncated to the visible density. Two LMNG molecules were identified and built. The structure was refined using *phenix.real_space_refine* (*54*), followed by several rounds of rebuilding in COOT. The ΔNDUFS4 and ΔNDUFS6 structures were built with the wild-type model as template. The model was cross-validated against overfitting by the method as described earlier (*55*): The FSC between the refined model and the final map was determined (FSC_{SUM}; fig. S2B). Random shifts (up to 0.3 Å) were introduced into the coordinates of the final model, followed by refinement in PHENIX against the first unfiltered half-map. Then, the FSC was determined between this refined model and the half-map used for refinement (FSC_{work}), and the FSC against the second half-map, which was not used at any point during refinement (FSC_{free}). The marginal gap between the curves describing FSC_{work} and FSC_{free} indicates that the model was not overfitted (see fig. S2B). A quality check indicated excellent stereochemistry with 91.51 to 93.89% of the nonglycine and nonproline residues found in the most-favored regions and with 0.15 to 0.17% outliers (all-atom clashscore, 8.36 to 9.29). Refinement and validation statistics were summarized in Table 1. Figures were drawn with Chimera (*56*) and PyMOL (The PyMOL Molecular Graphics System, Version 2.0, Schrödinger, LLC).

Computational methods

Classical atomistic MD simulations were performed on the entire structure of complex I from *Y. lipolytica*. Three short stretches that were missing from the structure (residues 355 to 359 in the NDUFA9 subunit and residues 77 to 83 and 138 to 151 in the NDUFB11 subunit, respectively) were modeled using the Modeller software (*57*). All structurally resolved lipids, the ligands, and the metal centers were included in the model [Q9, flavin mononucleotide (FMN), FeS clusters, NADPH, and phosphopantetheine (ZMP)]. The entire protein system was embedded into a hybrid lipid bilayer, which was generated using CHARMM-GUI (*58*). The membrane composition was tailored to mimic the inner mitochondrial membrane with 50% 3-palmitoyl-2-oleoyl-D-glycero-1-phosphatidylcholine (POPC), 35% 3-palmitoyl-2-oleoyl-D-glycero-1-phosphatidylethanolamine (POPE), and 15% CL. The protein-membrane complex was then solvated with TIP3 water molecules and Na^+ and Cl^- ions (100 mM). The CHARMM36 force field was used for protein, lipids, water, and ions (*59*, *60*), and the available force field parameters for oxidized Q9 (*61*), FMN and FeS clusters [see (*62*)], NADPH (*63*), and ZMP (*58*). The total system size was about 1.3 million atoms.

GROMACS (2016.4) (*64*) was used to perform the MD simulations. First, the energy of the entire system was minimized to fix the geometry of all the components and release strain from the system. Furthermore, keeping harmonic restraints on the heavy protein atoms (force constant 20,000 kJ mol^{-1} nm^{-2}) and on the membrane lipid phosphates (2000 kJ mol^{-1} nm^{-2} force constant in the *z* direction), we ran a short (100 ps) NVT equilibration using velocity-rescaling thermostat (*65*). Subsequently, we performed a 10-ns NPT equilibration run with harmonic restraints on the backbone of the protein and on heavy atoms of Q9 with the Berendsen barostat (*66*). In the production

MD, all atoms were kept free. The Nose-Hoover thermostat (*67*, *68*) and Parrinello-Rahman barostat (*69*) were applied, keeping the temperature at 310 K and a pressure of 1 atmosphere. The time step for integration during MD was 2 fs using LINCS (*70*). Cryo-EM structures and simulation trajectories were analyzed with Visual Molecular Dynamics (VMD) (*71*). The total simulation time was ~1 μs.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/5/12/eaax9484/DC1

Supplementary Materials and Methods

Fig. S1. Structure determination of wild-type complex I, assembly intermediate, and mutant without NDUFS4.

Fig. S2. Representative cryo-EM densities of wild-type complex I with fitted models.

Fig. S3. Assignment of accessory subunits.

Fig. S4. A second Q binding site in complex I and quantitative mass spectrometry.

Fig. S5. Atomic model of *Y. lipolytica* complex I in membrane-solvent environment and hydrogen bond analysis from MD simulation trajectory.

Fig. S6. Interactions between protein and the Q head group and tail.

Fig. S7. Lipid binding sites in yeast and mammalian complex I.

Fig. S8. Surface helices and lipids modeled at the interface of the membrane arm and the peripheral arm.

Fig. S9. Bending of the lipid tails seen in MD simulation.

Fig. S10. Spectrophotometric assay of NADH:DBQ activity.

Fig. S11. Conformation of the β1-β2 loop of NDUFS2 and loops and regions critical for the active/deactive transition.

Fig. S12. Proton pumping activity of the NDUFS4 mutant.

Fig. S13. Cartoon representation and sequence alignment of NDUFAF2 and NDUFA12 from *Y. lipolytica* and *Homo sapiens*.

Table S1. Data collection, refinement, and model statistics.

Table S2. Subunit names and chain identifiers for *Y. lipolytica* and human and bovine complex I.

Table S3. Lipid species detected in *Y. lipolytica* complex I.

Movie S1. Cryo-EM structure of respiratory complex I from *Y. lipolytica* in LMNG at 3.2-Å resolution.

Movie S2. Cryo-EM structure of ΔNDUFS4 at 4.0-Å resolution.

Movie S3. Cryo-EM structure of the ΔNDUFS6 assembly intermediate at 3.3-Å resolution.

References (*73*–*78*)

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. J. Hirst, Mitochondrial complex I. *Annu. Rev. Biochem.* **82**, 551–575 (2013).
2. R. J. Rodenburg, Mitochondrial complex I-linked disease. *Biochim. Biophys. Acta* **1857**, 938–945 (2016).
3. E. T. Chouchani, C. Methner, S. M. Nadtochiy, A. Logan, V. R. Pell, S. Ding, A. M. James, H. M. Cochemé, J. Reinhold, K. S. Lilley, L. Partridge, I. M. Fearnley, A. J. Robinson, R. C. Hartley, R. A. J. Smith, T. Krieg, P. S. Brookes, M. P. Murphy, Cardioprotection by S-nitrosation of a cysteine switch on mitochondrial complex I. *Nat. Med.* **19**, 753–759 (2013).
4. A. B. Kotlyar, A. D. Vinogradov, Slow active/inactive transition of the mitochondrial NADH-ubiquinone reductase. *Biochim. Biophys. Acta* **1019**, 151–158 (1990).
5. A.-N. A. Agip, J. N. Blaza, J. G. Fedor, J. Hirst, Mammalian respiratory complex I through the lens of Cryo-EM. *Annu. Rev. Biophys.* **48**, 165–184 (2019).
6. R. Baradaran, J. M. Berrisford, G. S. Minhas, L. A. Sazanov, Crystal structure of the entire respiratory complex I. *Nature* **494**, 443–448 (2013).
7. V. Zickermann, C. Wirth, H. Nasiri, K. Siegmund, H. Schwalbe, C. Hunte, U. Brandt, Mechanistic insight from the crystal structure of mitochondrial complex I. *Science* **347**, 44–49 (2015).
8. K. Parey, U. Brandt, H. Xie, D. J. Mills, K. Siegmund, J. Vonck, W. Kühlbrandt, V. Zickermann, Cryo-EM structure of respiratory complex I at work. *eLife* **7**, e39213 (2018).
9. A. Cabrera-Orefice, E. G. Yoga, C. Wirth, K. Siegmund, K. Zwicker, S. Guerrero-Castillo, V. Zickermann, C. Hunte, U. Brandt, Locking loop movement in the ubiquinone pocket of complex I disengages the proton pumps. *Nat. Commun.* **9**, 4500 (2018).
10. J. D. Ortigoza-Escobar, A. Oyarzabal, R. Montero, R. Artuch, C. Jou, C. Jiménez, L. Gort, P. Briones, J. Muchart, E. López-Gallardo, S. Emperador, E. R. Pesini, J. Montoya, B. Pérez, P. Rodríguez-Pombo, B. Pérez-Dueñas, Ndufs4 related Leigh syndrome: A case report and review of the literature. *Mitochondrion* **28**, 73–78 (2016).
11. M. E. Breuer, P. H. G. M. Willems, J. A. M. Smeitink, W. J. H. Koopman, M. Nooteboom, Cellular and animal models for mitochondrial complex I deficiency: A focus on the NDUFS4 subunit. *IUBMB Life* **65**, 202–208 (2013).
12. F. Kahlhöfer, K. Kmita, I. Wittig, K. Zwicker, V. Zickermann, Accessory subunit NUYM (NDUFS4) is required for stability of the electron input module and activity of mitochondrial complex I. *Biochim. Biophys. Acta* **1858**, 175–181 (2017).
13. I. Ogilvie, N. G. Kennaway, E. A. Shoubridge, A molecular chaperone for mitochondrial complex I assembly is mutated in a progressive encephalopathy. *J. Clin. Invest.* **115**, 2784–2792 (2005).
14. K. Fiedorczuk, J. A. Letts, G. Degliesposti, K. Kaszuba, M. Skehel, L. A. Sazanov, Atomic structure of the entire mammalian mitochondrial complex I. *Nature* **538**, 406–410 (2016).
15. E. G. Yoga, O. Haapanen, I. Wittig, K. Siegmund, V. Sharma, V. Zickermann, Mutations in a conserved loop in the PSST subunit of respiratory complex I affect ubiquinone binding and dynamics. *Biochim. Biophys.* **7**, 573–581 (2019).
16. M. Verkhovsky, D. A. Bloch, M. Verkhovskaya, Tightly-bound ubiquinone in the *Escherichia coli* respiratory complex I. *Biochim. Biophys. Acta* **1817**, 1550–1556 (2012).
17. J. Warnau, V. Sharma, A. P. Gamiz-Hernandez, A. Di Luca, O. Haapanen, I. Vattulainen, M. Wikström, G. Hummer, V. R. I. Kaila, Redox-coupled quinone dynamics in the respiratory complex I. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E8413–E8420 (2018).
18. O. Haapanen, A. Djurabekova, V. Sharma, Role of Second Quinone Binding Site in Proton Pumping by Respiratory Complex I. *Front. Chem.* **7**, 221 (2019).
19. M. Wikström, V. Sharma, V. R. Kaila, J. P. Hosler, G. Hummer, New perspectives on proton pumping in cellular respiration. *Chem. Rev.* **115**, 2196–2221 (2015).
20. M. A. Tocilescu, V. Zickermann, K. Zwicker, U. Brandt, Quinone binding and reduction by respiratory complex I. *Biochim. Biophys. Acta* **1797**, 1883–1890 (2010).
21. A.-N. A. Agip, J. N. Blaza, J. G. Fedor, J. Hirst, Cryo-EM structures of complex I from mouse heart mitochondria in two biochemically defined states. *Nat. Struct. Mol. Biol.* **48**, 165–184 (2018).
22. S. Ludtke, K. He, H. Huang, Membrane thinning caused by magainin 2. *Biochemistry* **34**, 16764–16769 (1995).
23. P. Kaurola, V. Sharma, A. Vonk, I. Vattulainen, T. Róg, Distribution and dynamics of quinones in the lipid bilayer mimicking the inner membrane of mitochondria. *Biochim. Biophys. Acta* **1858**, 2116–2122 (2016).
24. E. Maklashina, A. B. Kotlyar, G. Cecchini, Active/de-active transition of respiratory complex I in bacteria, fungi, and animals. *Biochim. Biophys. Acta* **1606**, 95–103 (2003).
25. A. Galkin, B. Meyer, I. Wittig, M. Karas, H. Schägger, A. Vinogradov, U. Brandt, Identification of the mitochondrial ND3 subunit as a structural component involved in the active/deactive enzyme transition of respiratory complex I. *J. Biol. Chem.* **283**, 20907–20913 (2008).
26. J. Zhu, K. R. Vinothkumar, J. Hirst, Structure of mammalian respiratory complex I. *Nature* **536**, 354–358 (2016).
27. J. N. Blaza, K. R. Vinothkumar, J. Hirst, Structure of the Deactive State of Mammalian Respiratory Complex I. *Structure* **26**, 312–319.e3 (2018).
28. C.-y. Y. Yip, M. E. Harbour, K. Jayawardena, I. M. Fearnley, L. A. Sazanov, Evolution of respiratory complex I: "supernumerary" subunits are present in the alpha-proteobacterial enzyme. *J. Biol. Chem.* **286**, 5023–5033 (2011).
29. B. Pereira, A. Videira, M. Duarte, Novel insights into the role of Neurospora crassa NDUFAF2, an evolutionarily conserved mitochondrial complex I assembly factor. *Mol. Cell. Biol.* **33**, 2623–2634 (2013).
30. Z. Assouline, M. Jambou, M. Rio, C. Bole-Feysot, P. de Lonlay, C. Barnerias, I. Desguerre, C. Bonnemains, C. Guillermet, J. Steffann, A. Munnich, J. P. Bonnefont, A. Rötig, A. S. Lebre, A constant and similar assembly defect of mitochondrial respiratory chain complex I allows rapid identification of NDUFS4 mutations in patients with Leigh syndrome. *Biochim. Biophys. Acta* **1822**, 1062–1069 (2012).
31. S. E. Kruse, W. C. Watt, D. J. Marcinek, R. P. Kapur, K. A. Schenkman, R. D. Palmiter, Mice with mitochondrial complex I deficiency develop a fatal encephalomyopathy. *Cell Metab.* **7**, 312–320 (2008).
32. M. A. Calvaruso, P. Willems, M. van den Brand, F. Valsecchi, S. Kruse, R. Palmiter, J. Smeitink, L. Nijtmans, Mitochondrial complex III stabilizes complex I in the absence of NDUFS4 to provide partial activity. *Hum. Mol. Genet.* **21**, 115–120 (2012).
33. J. S. Sousa, D. J. Mills, J. Vonck, W. Kühlbrandt, Functional asymmetry and electron flow in the bovine respirasome. *eLife* **5**, e21290 (2016).
34. K. Kmita, C. Wirth, J. Warnau, S. Guerrero-Castillo, C. Hunte, G. Hummer, V. R. I. Kaila, K. Zwicker, U. Brandt, V. Zickermann, Accessory NUMM (NDUFS6) subunit harbors a Zn-binding site and is essential for biogenesis of mitochondrial complex I. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 5685–5690 (2015).
35. D. M. Kirby, R. Salemi, C. Sugiana, A. Ohtake, L. Parry, K. M. Bell, E. P. Kirk, A. Boneh, R. W. Taylor, H.-H. M. Dahl, M. T. Ryan, D. R. Thorburn, NDUFS6 mutations are a novel cause of lethal neonatal mitochondrial complex I deficiency. *J. Clin. Invest.* **114**, 837–845 (2004).
36. M. Lazarou, M. McKenzie, A. Ohtake, D. R. Thorburn, M. T. Ryan, Analysis of the assembly profiles for mitochondrial- and nuclear-DNA-encoded subunits into complex I. *Mol. Cell. Biol.* **27**, 4228–4237 (2007).
37. T. Gabaldon, D. Rainey, M. A. Huynen, Tracing the evolution of a large protein complex in the eukaryotes, NADH:Ubiquinone oxidoreductase (Complex I). *J. Mol. Biol.* **348**, 857–870 (2005).
38. S. Guerrero-Castillo, F. Baertling, D. Kownatzki, H. J. Wessels, S. Arnold, U. Brandt, L. Nijtmans, The Assembly Pathway of Mitochondrial Respiratory Chain Complex I. *Cell Metab.* **25**, 128–139 (2017).
39. R. O. Vogel, M. A. van den Brand, R. J. Rodenburg, L. P. van den Heuvel, M. Tsuneoka, J. A. Smeitink, L. G. Nijtmans, Investigation of the complex I assembly chaperones B17.2L and NDUFAF1 in a cohort of CI deficient patients. *Mol. Genet. Metab.* **91**, 176–182 (2007).

40. C. Hunte, V. Zickermann, U. Brandt, Functional modules and structural basis of conformational coupling in mitochondrial complex I. *Science* **329**, 448–451 (2010).
41. N. Kashani-Poor, S. Kerscher, V. Zickermann, U. Brandt, Efficient large scale purification of his-tagged proton translocating NADH:ubiquinone oxidoreductase (complex I) from the strictly aerobic yeast *Yarrowia lipolytica*. *Biochim. Biophys. Acta* **1504**, 363–370 (2001).
42. S. Dröse, A. Galkin, U. Brandt, Proton pumping by complex I (NADH:ubiquinone oxidoreductase) from *Yarrowia lipolytica* reconstituted into proteoliposomes. *Biochim. Biophys. Acta* **1710**, 87–95 (2005).
43. V. Matyash, G. Liebisch, T. V. Kurzchalia, A. Shevchenko, D. Schwudke, Lipid extraction by methyl-tert-butyl ether for high-throughput lipidomics. *J. Lipid Res.* **49**, 1137–1146 (2008).
44. A. Triebl, M. Trötzmüller, J. Hartler, T. Stojakovic, H. C. Köfeler, Lipidomics by ultrahigh performance liquid chromatography-high resolution mass spectrometry and its application to complex biological samples. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **1053**, 72–80 (2017).
45. J. Hartler, M. Trötzmüller, C. Chitraju, F. Spener, H. C. Köfeler, G. G. Thallinger, Lipid Data Analyzer: unattended identification and quantitation of lipids in LC-MS data. *Bioinformatics* **27**, 572–577 (2011).
46. G. Liebisch, J. A. Vizcaíno, H. Köfeler, M. Trötzmüller, W. J. Griffiths, G. Schmitz, F. Spener, M. J. O. Wakelam, Shorthand notation for lipid structures derived from mass spectrometry. *J. Lipid Res.* **54**, 1523–1530 (2013).
47. C. Suloway, J. Pulokas, D. Fellmann, A. Cheng, F. Guerra, J. Quispe, S. Stagg, C. S. Potter, B. Carragher, Automated molecular microscopy: The new Leginon system. *J. Struct. Biol.* **151**, 41–60 (2005).
48. S. Q. Zheng, E. Palovcak, J.-P. Armache, K. A. Verba, Y. Cheng, D. A. Agard, MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
49. J. Zivanov, T. Nakane, B. O. Forsberg, D. Kimanius, W. J. H. Hagen, E. Lindahl, S. H. W. Scheres, New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
50. A. Kucukelbir, F. J. Sigworth, H. D. Tagare, Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–65 (2014).
51. P. B. Rosenthal, R. Henderson, Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
52. S. H. W. Scheres, S. Chen, Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* **9**, 853–854 (2012).
53. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D* **60**, 2126–2132 (2004).
54. P. V. Afonine, B. K. Poon, R. J. Read, O. V. Sobolev, T. C. Terwilliger, A. Urzhumtsev, P. D. Adams, Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta. Crystallogr. Sect. D* **74**, 531–544 (2018).
55. A. Brown, F. Long, R. A. Nicholls, J. Toots, P. Emsley, G. Murshudov, Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta Crystallogr. Sect. D* **71**, 136–153 (2015).
56. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
57. A. Sali, T. L. Blundell, Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 (1993).
58. S. Jo, T. Kim, V. G. Iyer, W. Im, CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865 (2008).
59. J. B. Klauda, R. M. Venable, J. A. Freites, J. W. O'Connor, D. J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A. D. MacKerell Jr., R. W. Pastor, Update of the CHARMM all-atom additive force field for lipids: Validation on six lipid types. *J. Phys. Chem. B* **114**, 7830–7843 (2010).
60. A. D. MacKerell Jr., D. Bashford, M. Bellott, R. L. Dunbrack Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).
61. V. V. Galassi, G. M. Arantes, Partition, orientation and mobility of ubiquinones in a lipid bilayer. *Biochim. Biophys. Acta* **1847**, 1560–1573 (2015).
62. V. Sharma, G. Belevich, A. P. Gamiz-Hernandez, T. Róg, I. Vattulainen, M. L. Verkhovskaya, M. Wikström, G. Hummer, V. R. I. Kaila, Redox-induced activation of the proton pump in the respiratory complex I. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 11571–11576 (2015).
63. J. J. Pavelites, J. L. Gao, P. A. Bash, A. D. Mackerell Jr., A molecular mechanics force field for NAD^+, NADH, and the pyrophosphate groups of nucleotides. *J. Comput. Chem.* **18**, 221–239 (1997).
64. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19–25 (2015).
65. G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
66. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, Molecular-Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).
67. W. G. Hoover, Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A. Gen. Phys.* **31**, 1695–1697 (1985).
68. S. Nosé, A unified formulation of the constant temperature molecular-dynamics methods. *J. Chem. Phys.* **81**, 511–519 (1984).
69. M. Parrinello, A. Rahman, Polymorphic transitions in single-crystals - a new molecular-dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
70. B. Hess, P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *J. Chem. Theory Comput.* **4**, 116–122 (2008).
71. W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics. *J. Mol. Graph. Model.* **14**, 33–38 (1996).
72. M. M. Roessler, M. S. King, A. J. Robinson, F. A. Armstrong, J. Harmer, J. Hirst, Direct assignment of EPR spectra to structurally defined iron-sulfur clusters in complex I by double electron-electron resonance. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 1930–1935 (2010).
73. M. Babot, A. Birch, P. Labarbuta, A. Galkin, Characterisation of the active/de-active transition of mitochondrial complex I. *Biochim. Biophys. Acta* **1837**, 1083–1092 (2014).
74. N. Howell, L. A. Bindoff, D. A. McCullough, I. Kubacka, J. Poulton, D. Mackey, L. Taylor, D. M. Turnbull, Leber hereditary optic neuropathy: Identification of the same mitochondrial ND1 mutation in six pedigrees. *Am. J. Hum. Genet.* **49**, 939–950 (1991).
75. A. Jussupow, A. Di Luca, V. R. I. Kaila, How cardiolipin modulates the dynamics of respiratory complex I. *Sci. Adv.* **5**, eaav1850 (2019).
76. M. Fry, D. E. Green, Cardiolipin requirement for electron transfer in complex I and III of the mitochondrial respiratory chain. *J. Biol. Chem.* **256**, 1874–1880 (1981).
77. R. Guo, S. Zong, M. Wu, J. Gu, M. Yang, Architecture of Human Mitochondrial Respiratory Megacomplex $I_2III_2IV_2$. *Cell* **170**, 1247–1257.e12 (2017).
78. A. Drozdetskiy, C. Cole, J. Procter, G. J. Barton, JPred4: A protein secondary structure prediction server. *Nucleic Acids Res.* **43**, W389–W394 (2015).

Acknowledgments: We thank D. Januliene for help and advice during cryo-EM data collection using Leginon, and Ö. Yildiz and J. F. Castillo-Hernández for computer support. We thank P. Lastrico for the artwork and B. Murphy for the critical reading of the manuscript. We thank J. Meisterknecht for the excellent technical assistance. We acknowledge PRACE for awarding us access to MareNostrum at Barcelona Supercomputing Center (BSC), Spain. We also thank the Center for Scientific Computing, Finland, for the generous computational resources, including the Grand Challenge resources. **Funding:** This work was supported by the Deutsche Forschungsgemeinschaft (grant Zl 552/4-2 to V.Z. and SFB 815/Z1 to I.W.). V.S. acknowledges research funding from the Academy of Finland (294652), the University of Helsinki, and the Sigrid Jusélius Foundation. O.H. acknowledges research funds from the CHEMS doctoral school of the University of Helsinki. H.K. acknowledges funding from the Austrian Federal Ministry of Education, Science and Research grant number BMWFW-10.420/0005-WF/V/3c/2017. I.W. acknowledges the BMBF mitoNET-German Network for Mitochondrial Disorders 01GM1906D. EM data were collected and processed through the cryo-EM suite of the MPI of Biophysics, funded by the Max Planck Society. **Author contributions:** K.P. purified and characterized complex I, prepared cryo-EM grids, acquired and processed cryo-EM data, built the model, analyzed the data, and drew the figures. O.H. performed the modeling and simulation of complex I, analyzed the simulation data, and drew the figures. V.S. analyzed the simulation data and interpreted its mechanistic implications and supervised the modeling and simulation work. H.K. and T.Z. performed and analyzed quantitative mass spectrometry. S.P. prepared cryo-EM grids and collected data. K.S. purified and characterized complex I. I.W. performed mass spectrometry and analyzed and interpreted the MS data. D.J.M. optimized the high-resolution EM alignment and the data collection procedure and collected cryo-EM data. J.V. built the model and interpreted the maps. W.K. provided cryo-EM infrastructure, supervised the cryo-EM work, and cowrote the paper. V.Z. designed the study, interpreted the mechanistic implications of the structures, and wrote the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The cryo-EM structures of parental complex I, ΔNUDFS6, and ΔNDUFS4 have been deposited in the PDB with PDB IDs 6RFR, 6RFQ, and 6RFS, and the respective cryo-EM maps in the EMDB under accession numbers EMD-4873, EMD-4872, and EMD-4874. Additional data related to this paper may be requested from the authors.

Submitted 7 May 2019
Accepted 22 October 2019
Published 11 December 2019
10.1126/sciadv.aax9484

Citation: K. Parey, O. Haapanen, V. Sharma, H. Köfeler, T. Züllig, S. Prinz, K. Siegmund, I. Wittig, D. J. Mills, J. Vonck, W. Kühlbrandt, V. Zickermann, High-resolution cryo-EM structures of respiratory complex I: Mechanism, assembly, and disease. *Sci. Adv.* **5**, eaax9484 (2019).

High-resolution cryo-EM structures of respiratory complex I: Mechanism, assembly, and disease

Kristian Parey, Outi Haapanen, Vivek Sharma, Harald Köfeler, Thomas Züllig, Simone Prinz, Karin Siegmund, Ilka Wittig, Deryck J. Mills, Janet Vonck, Werner Kühlbrandt and Volker Zickermann

Sci Adv **5** (12), eaax9484.
DOI: 10.1126/sciadv.aax9484