Structure of *Escherichia coli* respiratory complex I reconstituted into lipid nanodiscs reveals an uncoupled conformation

Piotr Kolata[*$], and Rouslan G. Efremov[*$1]

[*] Center for Structural Biology, Vlaams Instituut voor Biotechnologie, Brussels, Belgium

[$] Structural Biology Brussels, Department of Bioengineering Sciences, Vrije Universiteit Brussel, Brussels, Belgium

[1] Corresponding author.

E-mail address: rouslan.efremov@vub.vib.be

Abstract

Respiratory complex I is a multi-subunit membrane protein complex that reversibly couples NADH oxidation and ubiquinone reduction with proton translocation against trans-membrane potential. Complex I from *Escherichia coli* is among the best functionally characterized complexes, but its structure remains unknown, hindering further mechanistic studies to understand the enzyme coupling mechanism. Here we describe the single particle cryo-electron microscopy (cryo-EM) structure of the entire catalytically active *E. coli* complex I reconstituted into lipid nanodiscs. The structure of this mesophilic bacterial complex I displays highly dynamic connection between the peripheral and membrane domains. The peripheral domain assembly is stabilized by unique terminal extensions and an insertion loop. The membrane domain structure reveals novel dynamic features. Unusual conformation of the conserved interface between the cytoplasmic and membrane domains suggests an uncoupled conformation of the complex. Based on these structural data we suggest a new simple and testable coupling mechanism for the molecular machine.

Introduction

Complex I, NADH:ubiquinone oxidoreductase, is a multi-subunit enzyme found in many bacteria and most eukaryotes. It facilitates transfer of two electrons from NADH to ubiquinone, or its analogues, coupled reversibly with translocation of four protons across the membrane against trans-membrane potential (Galkin et al., 2006; Sazanov, 2015). Structures of the complete complex I from several eukaryotes (Fiedorczuk et al., 2016; Hunte et al., 2010; Kampjut and Sazanov, 2020; Zhu et al., 2016), one thermophilic bacterium (Baradaran et al., 2013), and the partial structure of the membrane domain of *Escherichia coli* complex I (Efremov and Sazanov, 2011), have been determined.

The composition of complex I differs significantly between species. Mitochondrial complex I has molecular weight 1 MDa and comprises more than 35 subunits (Wirth et al., 2016) whereas bacterial analogues are much smaller with molecular weights approximately 500 kDa. Complex I from all characterized species contains homologues of 14 core subunits; seven subunits each assemble into peripheral and membrane arms, joined at their tips and form the complex with a characteristic L-shape.

The peripheral arm, exposed to the cytoplasm in bacteria or the mitochondrial matrix in eukaryotes, contains binding sites for NADH, ubiquinone, and flavin mononucleotide (FMN) as well as eight or nine iron-sulfur clusters, seven of which connect the NADH and ubiquinone-binding sites (Sazanov, 2015) enabling rapid electron transfer (Verkhovskaya et al., 2008) .

The membrane-embedded arm includes a chain of three antiporter-like subunits, NuoL, NuoM, and NuoN (*E. coli* nomenclature is used for the subunits hereafter) (Efremov and Sazanov, 2011), which are also found in the Mrp family of multisubunit H^+/Na antiporters (Steiner and Sazanov, 2020). Each antiporter-like subunit contains two structural repeats comprising five trans-membrane helices (TMH, TMH4-8, and TMH9-13). TMH7 and

TMH12 are interrupted by an extended loop in the middle of the membrane and the helix TM8 at the interface between symmetric motifs is interrupted by the π-bulge (Baradaran et al., 2013; Efremov and Sazanov, 2011). Membrane-embedded NuoH mediates interaction with the peripheral arm and also contains five-helix structural repeats found in antiporter-like subunits (Baradaran et al., 2013). Together with subunits NuoB and NuoD it forms an extended ubiquinone-binding cavity (Q-cavity) spanning the membrane bilayer hydrophobic region to the ubiquinone-binding site (Q-site) in the proximity of the terminal iron-sulfur cluster N2 (Baradaran et al., 2013).

The membrane arm features a continuous chain of conserved and functionally important ionizable residues positioned in the middle of the membrane. These are suggested to be involved in proton translocation and its coupling to electron transfer (Baradaran et al., 2013; Efremov and Sazanov, 2011). Attempts to visualize conformational changes in the membrane domain (Kampjut and Sazanov, 2020; Parey et al., 2018) have revealed rotation of the cytoplasmic half of TMH3 of NuoJ in mammalian complex I (Agip et al., 2018) and were associated with active-deactive transition. Recently, proton translocation mechanisms without conformational changes in antiporter-like subunits were suggested (Kampjut and Sazanov, 2020; Steiner and Sazanov, 2020). However, all proposed coupling mechanisms remain largely speculative and require further validation by functional, biochemical, and structural methods.

E. coli complex I is among the best functionally characterized complex I. It has been studied using many biophysical and biochemical techniques (Verkhovskaya and Bloch, 2012). Combined with the possibility of fast and extensive mutagenesis (Pohl et al., 2007; Verkhovskaya and Bloch, 2012), it represents a highly attractive system to study the coupling mechanism. However, owing to its fragile and dynamic nature (Verkhovskaya and Bloch,

2012), high-resolution structures of this complex remain limited to a partial structure of the membrane domain (Efremov and Sazanov, 2011).

Here we present a single particle cryo-EM structure of the entire *E. coli* complex I reconstituted into lipid nanodiscs, with the peripheral arm structure solved at 2.1 Å resolution and that of the membrane domain at 3.7 Å.

Results

Overall structure

Twin-strep tag was added to genomically encoded subunit NuoF using a CRISPR-Cas9 based system (Jiang et al., 2015) (Figure 1 - figure supplement 1). This enabled single-step purification of solubilized complex (Figure 1 - figure supplement 2A), which was further reconstituted into lipid nanodiscs comprising *E. coli* polar lipids and membrane scaffold protein MSP2N2 (Grinkova et al., 2010) (Figure 1 - figure supplement 2A,B). Mass photometry indicated that reconstituted complex I was homogeneous and monodispersed (Figure 1 - figure supplement 2C,D).

NADH:potassium ferricyanide (FeCy) and NADH:ubiquinone-1 (Q1) activities of the reconstituted complex I (Figure 1 - figure supplement 2E,F) were similar to those of detergent-purified protein supplemented with native *E. coli* lipids (Sazanov, 2003). Furthermore, NADH:Q1 activity was completely inhibited by piericidin-A (Figure 1 - figure supplement 2F) indicating that complex I reconstituted in lipid nanodiscs was intact and catalytically active in a detergent-free environment.

We determined the single particle cryo-EM structure of the reconstituted complex (Figure 1, Figure 1 - figure supplement 3,4, Table 1, Movie 1). Multiple conformations of the complex that differed by relative positions of the peripheral and membrane arms were revealed by 3D classification (Figure 1 - figure supplement 4,5). Three conformations of the entire complex were reconstructed to average resolutions between 3.3 and 3.7 Å (Figure 1 - figure

supplement 4) resolving the interface between the arms; however, due to high-residual mobility of the arms, the antiporter-like subunits were resolved at below 8 Å (Figure 1 - figure supplement 4).

Focused refinement of each arm separately and subtraction of nanodisc density (Figure 1 - figure supplement 3) improved the resolution of peripheral and membrane arms to 2.7 Å and 3.7 Å, respectively (Figure 1 - figure supplement 3,4, Table 1).

Micrograph analysis, in contrast to mass photometry, revealed that large fraction of the particles corresponds to the peripheral arm only (Figure 1 - figure supplement 3) that may have dissociated during cryo-EM sample preparation. These yielded 3D reconstruction to 2.8 Å resolution (Figure 1 - figure supplement 3), similar to the map of the peripheral arm of intact complex I. Joining two subsets improved resolution of the peripheral arm to 2.1 Å (Table 1, Figure 1 - figure supplement 4). Using the resulting maps, an atomic model of the entire *E. coli* complex I was built, comprising 4618 residues and accounting for 94.7% of the total polypeptide constituting the complex (Table 2).

Arrangement of the arms and individual core subunits in *E. coli* complex I is very similar to that of *Thermus thermophilus* (RMSD 6.3 Å over 2593 Cα atoms, 4.0 Å membrane arm over 1604 atoms, and 2.2 Å for peripheral arm 855 Cα) and mammalian enzyme (8.3 Å over 2420 Cα atoms (3.8 Å MD 1710 atoms, PD 2.0 767 Cα)) (Figure 1B) apart from the relative long-range twisting and bending of arms observed between complex I from different species (Baradaran et al., 2013; Vinothkumar et al., 2014).

Comparison of *E. coli* complex I conformations reconstructed to better than 4 Å resolution revealed two modes of relative arm rotation (Figure 1C): 1) rotation around an axis that passes through the NuoH-NuoB interface and is tilted around 45 degrees out of the plane formed by the arms with an amplitude of at least 13 degrees, and 2) rotation around an axis parallel to the membrane and roughly perpendicular to the long axis of the membrane arm

with an amplitude of approximately 4 degrees. Although relative arm movements were observed in mammalian (Kampjut and Sazanov, 2020; Zhu et al., 2016) and *T. thermophilus* complex I (Gutiérrez-Fernández et al., 2020), their amplitudes were smaller and movement directionality was less diverse. Despite significant relative arm movements, the structure of each arm was rigid and did not reveal different conformations apart from the specific local dynamics discussed below.

Structure of the peripheral arm

Architecture of the peripheral arm reveals a novel evolutionary strategy to stabilize the subcomplex

At an average resolution of 2.1 Å with the local resolution in the core reaching 2.0 Å (Figure 1 - figure supplement 4) conformations of most side chains in the peripheral arm, positions of ions, and multiple water molecules were resolved unambiguously (Figure 2, Figure 1 - figure supplement 6).

The overall architecture of the conserved core of the peripheral arm subunits is very similar to other homologues. Unlike other structurally characterized homologues, *E. coli* subunits NuoC and NuoD are joined in a single polypeptide. The 35 amino acid-long linker includes an α-helix (residues 180-194) that interacts with subunit NuoB (Figure 2A). The relative positions of all redox centers with FMN and nine iron-sulfur clusters, including off path cluster N7 (Sazanov, 2006), are particularly well conserved (Figure 2C).

A distinctive feature of the *E. coli* peripheral arm is the presence of ordered C-terminal extensions in subunits NuoB, NuoI, and NuoF with a length of 22 to 45 residues and a large 94 residue insertion loop in subunit NuoG, referred to as the G-loop (Figure 2A, Table 3). These extensions are unique among structurally characterized complex I homologues and have a well-defined structure. While the G-loop has a compact fold, the conformation of the C-terminal tails is extended. They line the surface of the conserved fold of the peripheral arm

with high shape complementarity (Figure 2A, Figure 2 - figure supplement 1). Apart from a few helical turns, these extensions have no secondary structure elements (Table 3). They create additional inter-subunit contacts with some surface areas exceeding 1,000 Å^2 and involving polar interactions (Table 3). Similarly, the G-loop fills a crevice between NuoCD, NuoI, and NuoG subunits (Figure 2A). Together the extensions and G-loop increase the interaction surface between the electron acceptor module (NuoEFG) and connecting module (NuoICDB) by a factor of three (from 1400 to 4600 Å^2), thus stabilizing the peripheral arm assembly. These structural features are conserved within the Enterobacteriaceae family and are very common in the phylum Gammaproteobacteria. They display high conservation of interfacial residues, particularly for the G-loop (Figure 2 - figure supplement 1) and demonstrate a new evolutionary strategy for complex stabilization that was not observed to date in complex I structures from other species.

A strong density near the NuoG surface coordinated by ^{G}Asp617, ^{G}Gln632, ^{G}Glu647, ^{G}Asp731 and four water molecules (Figure 2B) was assigned to a Ca^{2+} ion. The coordination number, geometry, and ion-ligand distances of ca 2.5 Å (H. Zheng et al., 2008) as well as the 2 mM concentration of Ca^{2+} in the buffer support this assignment. Divalent ions are known to increase both the activity and stability of *E. coli* complex I (Sazanov, 2003). One of the calcium ligands, ^{G}Asp731, is part of the G-loop, suggesting that Ca^{2+} stabilizes the fold of the G-loop and consequently, the peripheral arm.

The extensions spatially overlap with the supernumerary subunits of complex I from *T. thermophilus* (Sazanov, 2006) and the structurally conserved supernumerary subunits of eukaryotic complex I (Zhu et al., 2016) (Table 3), consistent with the suggestion that the primary role of supernumerary subunits is to stabilize the complex (Fiedorczuk et al., 2016).

Bound water molecules

At 2.1 Å resolution, 1165 water molecules associated with the peripheral arm were modelled (Figure 2D). The positions of 180 water molecules are conserved with those identified in the peripheral arm of ovine complex I (Kampjut and Sazanov, 2020) (Figure 2D, red spheres). Most of conserved waters are buried in the interior of the subunits, shielded from the solvent, and most likely play a structural role in maintaining the subunit fold. Only a few water molecules interact closely with iron-sulfur clusters and may influence their potential (Table 4, Figure 1 - figure supplement 5). The water molecules located close to or between iron-sulfur clusters are not more conserved than those in the other parts of the complex, suggesting that they were not evolutionary selected to optimize the rate of electron transfer as was suggested by Schulte et.al. (Schulte et al., 2019a).

At 2.1 Å resolution, several unusual density features were observed next to some surface-exposed histidines and between some cysteine-methionine pairs as listed in Table 6 and depicted in Figure 3 - figure supplement 1.

Electron input and output sites

The FMN conformation and key water molecules in the NADH-binding pocket of *E. coli* complex I are conserved (Kampjut and Sazanov, 2020; Schulte et al., 2019b). This includes the position of W1060 that forms a hydrogen bond with the isoalloxazine ring N5 atom in FMN and with ^{F}Glu92, which likely acts as the activating group during catalysis of hydride transfer from NADH (Fraaije et al., 2000) (Figure 3A). Schulte et.al. (Schulte et al., 2019b) suggested a mechanism for regulation of reactive oxygen species (ROS) generation by *E. coli* complex I that involves flipping the carbonyl oxygen of ^{F}Glu93 upon enzymatic reduction. Our structure unambiguously places the corresponding carbonyl oxygen in a conformation that points away from FMN (Figure 3A) similar to conformations found in the reduced and

oxidized ovine complex I (Kampjut and Sazanov, 2020), which does not support its involvement in ROS regulation.

E. coli-specific features in the FMN-binding pocket include ^{F}His400 that replaces the Leu residues found in other homologues. ^{F}His400 is in Van der Waals contact with the isoalloxazine ring of FMN; its imidazole ring interacts directly with the N3 cluster iron atom and forms a hydrogen bond with Sγ of ^{F}Cys357 coordinating N3 (Figure3A). ^{F}His400 is solvent-accessible even in the presence of NADH, and therefore, may become protonated upon N3 reduction. ^{F}Arg320 is positioned such that it can form hydrogen bonds with the ribose moiety of the NADH nicotinamide group and may stabilize bound dinucleotide (Figure 3A). Both Arg320^{F} and His400^{F} may serve to counter-balance the negative charges of electrons on N1a and N3 clusters and to increase protein stability. The structure does not reveal specific features explaining the decreased affinity for FMN in the reduced enzyme (Holt et al., 2016). This can be attributed to minor conformational changes in the pocket upon enzyme reduction.

The Q-binding site in complex I is formed at the end of a crevice between NuoD and NuoB subunits (Baradaran et al., 2013). In *E. coli,* this wedge is formed by the 58–69 stretch of NuoB and the tip of the 220–225 loop from subunit NuoD. Both ^{D}Tyr273 and ^{D}His224, found in the proximity of bound decylubiquinone (Baradaran et al., 2013) are conserved in *E. coli* and point towards the quinone binding site, whereas the tip of the 218–223 loop is disordered as in most complex I structures.

Environment and potentials of iron-sulfur clusters

At a resolution of 2.1 Å the atoms constituting the iron-sulfur clusters are resolved as independent density blobs. The conformation of side chains as well as the positions of hydrating waters in the primary and secondary interaction spheres are mostly unambiguously resolved (Figure 3). In *E. coli* complex I, cluster N1a can be reduced by NADH due to its

uniquely high potential (~ -0.3 V), differentiating it from other characterized species in which N1a cannot be reduced by NADH (Birrell et al., 2013; Zu et al., 2002). The potential of iron-sulfur clusters in proteins among other factors depends on solvent exposure, proximity of charged residues, and the number of hydrogen bonds formed between the cluster environment and sulfur atoms of clusters or coordinating cysteines (Denke et al., 1998; Fritz et al., 2002). Comparison of the chemical environment of N1a with other high-resolution structures of complex I revealed three specific differences explaining the higher potential of the N1a cluster (Table 5): (1) *E. coli*-specific ^{E}Asn142 forms a hydrogen bond with Sγ of ^{E}Cys97 coordinating the N1a cluster and with N1a S1 (Figure 3B), consistent with its mutation to Met decreasing potential by 53 mV (Birrell et al., 2013). (2) In *E. coli*, water molecule W74 forms a hydrogen bond with Sγ of ^{E}Cys97. This water molecule resides in a hydrophilic cavity created by *E. coli* specific ^{E}Gly140, replacing the alanine residue found in other species. (3) Because of small differences in the backbone conformation of NuoF, the backbone nitrogen of ^{F}Gly97 can form a hydrogen bond with Sγ of ^{E}Cys133 in *E. coli* and *Aquifex aeolicus* but not in *Ovis aries* (Figure 3B, Table 5).

The environment of the other iron-sulfur clusters is mainly conserved. The differences in hydrogen donors to the clusters, cysteine sulfur atoms, and water molecules in the cluster vicinity are listed in Table 4. Clusters N3 and N2 are briefly discussed below as being the most interesting.

Cluster N3 interacts with ^{F}His400, which is absent in other structurally characterized species; however, the potential of N3 is very similar between species (Leif et al., 1995; Yagi and Matsuno-Yagi, 2003). The effect of proximal His residue is likely compensated by ^{F}Trp363 replacing the hydrogen bond donors (Glu or Gln) found in other species (Table 4).

The potential of cluster N2, the electron donor to quinone, varies in different species (Hirst and Roessler, 2016) notably being lower in *E. coli* compared to its mammalian analogues (-

220 mV *vs.* -140 mV, respectively). However, the structure shows that the polar environment of N2 is very conserved (Figure 3C), including two water molecules, W211 and W438. Two arginines found in close proximity to the N2 cluster, Arg270^{D} and Arg250^{D}, have conserved positions despite ^{49kDa}Arg85 in the mammalian homologue (^{D}Arg250) being dimethylated (Carroll et al., 2013). This modification prevents it from forming a hydrogen bond with ^{B}Cys63, which should decrease N2 potential in the mitochondrial enzyme. Therefore, finer structural differences including those in cluster geometry, are likely responsible for differences in potential, which can likely be explained by high-resolution structure-based modeling.

Structure of the membrane arm

The model of complete membrane arm, including the previously missing subunit NuoH (Efremov and Sazanov, 2011), was built into the density map with local resolution better than 3.5 Å at the arm center and approximately 4.0 Å at its periphery (Figure 1A, Figure 1 - figure supplement 4). An additional density belt corresponding to the lipid nanodisc is clearly visible (Figure 1A, Movie 1) around the membrane-embedded region. It is flat in the plane of the membrane with a thickness of approximately 30 Å, and closely matches hydrophobic surface of the membrane arm. The belt locally bends next to the subunit NuoL at the region where it interacts with the long amphipathic helix and is thinned next to the ^{H}TMH1 (Movie 1).

The structure of the membrane arm in the lipid nanodisc is very similar to the crystal structure of the detergent-solubilized membrane arm (Efremov and Sazanov, 2011) (RMSD of 1.1 Å over 12662 atoms) (Figure 4 - figure supplement 1). The curvature of the membrane arm observed previously (Efremov and Sazanov, 2011) was unchanged in the lipid environment, and therefore, is not an artifact of crystallization or solubilization (Verkhovskaya and Bloch, 2012). Local structural differences in crystal structure include expected repositioning of

^{A}TMH1 next to ^{H}TMH2 (Baradaran et al., 2013), and a change in conformation of the ^{M}TMH5-TMH6 loop (Figure 4 - figure supplement 1).

The fold of subunit NuoH is similar to the structures of *T. thermophilus* and of eukaryotic complexes with one important exception. The density for the N-terminus of NuoH (residues 1-52) that includes ^{H}TMH1 and a part of ^{H}TMH1-TMH2 loop, is completely missing in the reconstructions of the membrane fragment and of complete complex I (Figure 1, Movie 1) suggesting that ^{H}TMH1 is very mobile in the lipid nanodisc. This helix is close to the border of the nanodisc, and nanodisc belt is thinned on the cytoplasmic surface of the nanodisc at the position where the density of ^{H}TMH1 disappears (Movie 1). Simultaneously there is sufficient room to accommodate the trans-membrane helix within the nanodisc.

The structures of the membrane fragment and entire complex I visualize a complete chain of charged residues connecting the Q-site with charged residues in antiporter-like subunits. (Figure 4A). We analyzed the environment of ionizable residues found within the 'E-channel' (Baradaran et al., 2013), a region situated between the Q-cavity and antiporter-like subunit NuoN, to evaluate the existence of a continuous proton translocation path linking the Q-cavity with the antiporter-like subunits suggested for ovine complex I (Kampjut and Sazanov, 2020). The trans-membrane region of *E. coli* NuoH contains fewer charged residues than its homologues from other structurally characterized species (Figure 4 - figure supplement 2). Here, only *E. coli*-specific ^{H}His208, separated from ^{H}Glu157 by 12 Å, is found in the center of the membrane-embedded region of NuoH (Figure 4A). However, a large hydrophilic cavity stretches from the Q-site towards the center of subunit NuoH, ending next to the invariant **^{H}Glu157** (hereafter, invariant residues are marked in bold). Although **^{H}Glu157** is not directly linked to the cavity, DOWSER++ (Morozenko and Stuchebrukhov, 2016) placed waters linking it to the cavity, suggesting that this glutamic acid can exchange protons with the Q-cavity.

The region between NuoH and NuoN includes 6 ionizable side chains located in the middle of the membrane bilayer, 4 of which are invariant (Figure 4A,B). The distances between the residues vary from 5 Å to 12 Å which requires either displacement of the side chains or presence of water molecules to enable proton exchange between them. Analysis of cavities and potential hydration sites using DOWSER++ shows that the cluster of **^{H}Glu157** and **^{A}Asp79** along with the carbonyl oxygen of ^{J}Gly61 forming a π-bulge on the ^{J}TM3 (similar to the X-ray structure) (Efremov and Sazanov, 2011), indicate a hydrophilic cavity that can accommodate several water molecules, enabling proton exchange between these two residues. Carboxyl groups of a chain comprising ^{J}Glu55-**^{K}Glu36**-^{K}Glu72-**^{N}Glu133** are separated by cavities that can potentially be hydrated, enabling proton exchange between the residues. In *E. coli* complex I, residues **^{A}Asp79** and ^{J}Glu55/ **^{K}Glu36** are separated by a distance exceeding 12 Å and a region packed with hydrophobic residues, making proton exchange between the Q-site and NuoN unlikely. **^{A}Glu81,** located opposite **^{A}Asp79** on ^{A}TMH2, apparently does not participate in linking **^{A}Asp79** with **^{N}Glu133.** However, it faces hydrophilic environment of ^{J}Ser145, *E. coli*-specific ^{J}Glu142, and ^{A}Glu102, potentially linking it to the periplasmic surface (Figure 4A). Our analysis indicated that in *E. coli,* the E-channel is less pronounced than in *T. thermophilus* and that no continuous proton path exists between the Q-site and NuoN.

Curiously, ^{H}Lys274 almost universally conserved in complex I and related hydrogenases is found in the ^{H}TMH7 off the main pathways proposed for proton translocation. In our structure, its ammonium group is oriented towards the periplasm (Figure 4A); however, the length and flexibility of the side chain would allow it to reach the center of the membrane upon structural rearrangement.

The cytoplasmic half of ^{J}TMH3 was found to assume two alternative conformations in eukaryotic complex I (Agip et al., 2018; Kampjut and Sazanov, 2020). In *E. coli* complex I,

the density in this region is very well-resolved, suggesting the absence of alternative conformations. A peculiar feature is observed in the density of subunit NuoM instead. The density of the cytoplasmic half of ^{M}TM8 is poor and fragmented between residues 255 and 265, indicating the existence of multiple conformations (Figure 4C, Movie 2). This region is buried in the middle of NuoM and the density of surrounding helices is very well-resolved indicating the local character of the disorder. This region spans the invariant **Lys265** including the π-bulge, and found in some bacteria ^{M}Asp258. Interestingly, though the helix structure in mammalian complex I is similar to that in *E. coli* (the equivalent Asp is missing), the corresponding region in *T. thermophilus* differs significantly (Figure 4C). The cytoplasmic region of ^{13}TM8 is rotated by 2 residues. This is achieved by extending TMH8 at *T. thermophilus* **^{13}Lys 235** (equivalent to *E. coli* **^{M}Lys265**). In *E. coli*, **^{M}Lys265** is buried in the center of the second structural repeat whereas in *T. themophilus*, **^{13}Lys 235** is positioned at the interface between the structural repeats facing **^{13}His218** (in *E. coli* ^{M}His248). This rotation places ^{13}Asp228 (equivalent to *E. coli* ^{M}Asp258) pointing towards central axis of the first structural repeat (TM4-8) and exposed to the cytoplasm, whereas in *E. coli*, it is buried on the central axis of the second structural repeat (TM9-TM13). Thus, higher mobility of the helical fragment situated at a critical position at the interface of symmetry-related modules may indicate π-bulge-enabled helical rotation with a possible role in proton translocation.

The peripheral-membrane arm interface

The interface between membrane and peripheral arms presents an important element of the complex that mediates the coupling of ubiquinone reduction above the cytoplasmic surface to proton translocation across the membrane. The interface between peripheral and membrane arms is primarily formed through interaction between subunits NuoB and NuoD of the peripheral arm with the cytoplasmic surface of NuoH and the TMH1-TMH2 loop of subunit

NuoA in the membrane domain. The residues involved in the direct interaction between the arms and the interface structure are highly conserved (Figure 5 - figure supplement 1) between all complex I and related membrane-bound hydrogenases (Baradaran et al., 2013; Grba and Hirst, 2020; Kampjut and Sazanov, 2020; Yu et al., 2020; 2018).

Local resolution (Figure 1 - figure supplement 4) and B-factors (Figure 5 - figure supplement 2) of the *E. coli* peripheral arm show that the membrane-facing surface (subunits NuoD and NuoB) including the residues lining the Q-cavity are significantly less ordered than the remaining subcomplex. The mobility in the interfacial region of the arm is very similar in the reconstructions of the peripheral arm dissociated from and complexed with the membrane arm. However, several interfacial regions of subunits NuoB, NuoD, and NuoI become more ordered upon complex formation and their density can be observed in the reconstruction of isolated conformations of the entire complex I (Figure 5 - figure supplement 2, Table 2). Similar to complex I from *T. thermophilus* (Baradaran et al., 2013), there are no specific conformational changes at the interface upon association of the arms. These results suggest that the interfacial region of the peripheral arm is inherently flexible and likely responsible, at least in part, for the high relative mobility of the arms.

Structure of the surface of subunit NuoH and relative arrangement of subunits NuoB and NuoD in *E. coli* complex I is similar to that of complex I from other species (RMSD 1.2 Å over 324 Cα atoms with *T. thermophilus* complex I). However, their relative positions differ. Thus, in *E. coli* complex I, NuoB and NuoD are rotated around an axis passing through the center of NuoH and the interface between NuoF and NuoG anticlockwise when observed from the top of peripheral arm at approximately 15 degrees (Figure 5A). This results in the shift of NuoD interfacial regions with an amplitude exceeding 10 Å and the separation of NuoD from NuoH, which reduces the interaction between the four-helical bundle domain of NuoD with NuoH (Figure 5B). The highly conserved fragment of the ^{A}TMH1-TMH2 loop

(residues 46-53), that forms a plug between subunits NuoD and NuoB (Figure 5 - figure supplement 1B) and interacts with the ^{D}221-228 loop containing ubiquinone-coordinating His224, is also disordered in our structure (Figure 5B).

On the opposite side of the interface, structural rearrangements include a 7-degree tilt of ^{A}TMH1 that becomes more perpendicular to the membrane plane and approximately 15-degree rotation of the amphipathic helix in the loop connecting ^{H}TMH1-TMH2, residues 57–68, in the direction of ^{H}TMH1 and towards the membrane center (Figure 5C). Analysis of the three full conformations identifies this helix as the main membrane arm element that performs rearrangement together with the cytoplasmic domain. Its rotation reduces the opening to the Q-cavity (Figure 5D). Homology modeling indicates that the observed rearrangements are still compatible with ^{H}TMH1 occupying its expected position without any steric clashes (Figure 5D) suggesting that ^{H}TMH1 is highly mobile in the lipid environment rather than being absent from its expected position.

Rotation of the NuoB/NuoD subunits module creates multiple openings on the interface between the arms (Figure 5B). The size of the openings is compatible with the diffusion of water molecules and likely, of protons from the outer space towards the Q-cavity. Such openings suggest that ubiquinone bound to the Q-site can receive protons directly from the solvent.

Discussion

Structural features of *E. coli* complex I

E. coli complex I is composed of the smallest number of subunits among all complex I structures characterized so far. Yet it still evolved a strategy to stabilize peripheral arm assembly without involving additional subunits. The interactions between subunits are stabilized by extended C-termini and a large G-loop (Figure 2), in turn stabilized by the Ca^{2+}

ion, which is known to modulate the complex stability (Sazanov, 2003). This indicates an evolutionary pressure on maintaining the peripheral arm integrity, which was 'solved' in a species-dependent manner.

There is no apparent continuous proton translocation path between the Q-cavity and subunit NuoN in *E. coli* complex I. Further, there are no indications for the existence of different conformations in the cytoplasmic half of ^{J}TMH3 observed in mammalian complex I (Agip et al., 2018; Kampjut and Sazanov, 2020) attributed to deactive-active transition (Agip et al., 2018) or more recently, to different catalytic intermediates (Kampjut and Sazanov, 2020). This suggests that these states are either suppressed in the resting state of the bacterial complex or do not occur at all. Conversely, ^{M}TMH8 displays localized disorder next to the π-bulge, indicating involvement of this helix in the structural rearrangements associated with proton translocation, and to our knowledge, represents the first indication of specific conformational changes in antiporter-like subunits.

E. coli complex I is known to be a dynamic complex (Morgan and Sazanov, 2008; Sazanov, 2003). Our cryo-EM reconstructions reveal the reasons for its high flexibility. The peripheral and membrane arms are mainly rigid, whereas the connection between arms is flexible (Figure 1C, Figure 1 – figure supplement 5). Two reasons can be identified for this: 1) high mobility of the interfacial regions of subunits NuoB and NuoD (Figure 1 - figure supplement 4, Figure 5 - figure supplement 2) and 2) the 15-degree rotation of the interfacial subunits NuoB and NuoD relative to NuoH, observed uniquely in *E. coli* (Figure 5A). The rotation disrupts many conserved complimentary interactions between the arms and renders the interface porous such that the Q-cavity is exposed to the solvent. This is different from all the other known structures of complex I and evolutionarily related complexes in which the interface is solvent-inaccessible. Therefore, we interpreted the observed conformation of *E. coli* complex I as an uncoupled state. Our preparation of the complex is competent in

ubiquinone reduction (Figure 1 - figure supplement 2); however, unless conformational changes sealing the Q-cavity occur during the catalytic cycle, ubiquinone reduction by NADH in the present conformation is expected to occur without proton translocation.

The reasons for this difference in interface conformation with other structurally characterized complexes are not clear. It may represent a resting state described in *E. coli* complex I (Belevich et al., 2017), which like in eukaryotes (Babot et al., 2014), is characterized by lower catalytic activity and is activated by NADH:ubiquinone oxidoreduction cycles. However, reactive states in eukaryotic complex I are associated with local conformational changes involving loop rearrangement (Agip et al., 2018; Parey et al., 2018), rather than displacement of the complete domains observed in *E. coli* complex I. Other reasons for the observed rotation of arms may include the higher concentration of divalent ions used in our sample that weakens multiple salt bridges linking the arms, or displacement of the amphipathic termini of subunits NuoB and NuoI due to the limited size of the nanodisc, resulting in weakened interaction between the arms.

The absence of density for ^{H}TMH1 is another unique feature of *E. coli* complex I. It likely reflects the higher dynamics of this helix in lipid environment. Possibly, tilted ^{A}TMH1 (Figure 5C) displaces the periplasmic end of ^{H}TMH1 disrupting its interaction with the TMH2-3 loop and TMH6 of NuoH, rendering this helix dynamic. It may also have a functional role because helix displacement facilitates the otherwise too narrow access to the Q-cavity for ubiquinone (Baradaran et al., 2013). Further insights into the role of ^{H}TMH1 might be obtained once conditions stabilizing the coupled complex conformation are identified.

Revised coupling mechanism

Based on structural features of *E. coli* complex I and the wealth of available structural information, we would like to propose a coupling mechanism that differs from those

suggested previously (Figure 6), which is simple, compatible with the microscopic reversibility principle (Onsager et al., 1996), has evolutionary meaning, and is applicable to the entire class of evolutionarily related complexes.

We propose that the key to the coupling is the formation of a cavity isolated from external protons and accessible to ubiquinone such that ubiquinone can exchange electrons with the N2 cluster. The necessity of having a tightly coupled cavity explains the high conservation of the subunit interface. Notably, as the ubiquinone entrance is situated in the hydrophobic region of the bilayer, these two requirements do not contradict each other.

The potential of benzoquinone-hydroquinone couple depends on the pH (Chambers, 1988), similar to that of any redox reaction involving protons, and has been shown experimentally to decrease by over 400 mV to below -300 mV upon pH change from 2 to 10 (Lemmer et al., 2011). Rough estimations indicate that addition or extraction of a single proton from a cavity with the characteristic dimensions of the Q-cavity, alters the activity of protons within the cavity by hundreds of millivolts. Thus, the redox potential of ubiquinone bound within the cavity enclosed from the environment will be strongly modulated by the extraction/addition of single protons from/to the cavity. *Vice versa*, reduction or oxidation of ubiquinone/ubiquinol is equivalent to adding/removing proton binding groups to/from the Q-cavity. Thus, ubiquinone serves as a transformer that converts the energy of electrons to the chemical potential of protons in a fully reversible manner. In the coupled complex during the forward cycle, ubiquinone reduction decreases proton activity in the cavity, which is rectified by protons entering the cavity and performing work. Questions of how the protons perform the work and where they come from are thus critical to formulate the coupling mechanism.

Multiple proton pathways have been suggested (Baradaran et al., 2013; Efremov and Sazanov, 2012; Kampjut and Sazanov, 2020; Verkhovskaya and Bloch, 2012; Yu et al., 2020; 2018). However, they all end up on intracellular/matrix side of the membrane, which makes it

difficult to explain the energy conversion mechanism. Instead, we propose that the protons re-protonate ubiquinone through NuoH and/or adjacent trans-membrane subunits from the periplasmic/inter membrane side of the membrane as shown in Figure 6. The fold of subunit NuoH contains a set of 5 TMH with a helical arrangement similar to the symmetric module in antiporter-like subunits including broken TMH (Baradaran et al., 2013) and invariant ^{H}Glu157 in a position similar to the invariant ^{M}Glu144, suggesting that it can translocate protons. Proton transport through NuoH is coupled to the transport of three protons by three antiporter-like subunits in the opposite direction. This coupling must involve both the interaction of ionizable residues in the middle of the membrane (Baradaran et al., 2013; Efremov and Sazanov, 2011) as well as conformational changes, and likely proceeds through a classical alternating access mechanism (Jardetzky, 1966). Thus, the entire membrane module functions as a reversible proton antiporter with the stoichiometry of $1H^{+}_{in}/3H^{+}_{out}$. Four protons are translocated outside in two pumping cycles per one reduced ubiquinone molecule (Figure 6). In this mechanism, ubiquinone reduction creates a local enhanced membrane potential on the NuoH subunit between the Q-cavity and periplasmic space. Moreover, only the equilibrium potential of NADH and ubiquinone as well as the transmembrane potential are important for the directionality of the reaction and energy balance as expected for a molecular machine (Astumian et al., 2016).

Under equilibrium, the potential of ubiquinone in the Q-cavity is equilibrated with the potential of NADH, and of protons in the Q-cavity, which results in trans-membrane potential-dependent semiquinone species as observed by EPR spectroscopy in tightly coupled submitochondrial particles (Yano et al., 2005) .

The proposed mechanism is applicable to all the complexes that are evolutionary related to complex I (Efremov and Sazanov, 2012; Yu et al., 2020; 2018). In all of these, the cavity formed between the peripheral arm and NuoH subunit is sealed. The peripheral arm-NuoH

complex is undoubtably one of the stand-alone evolutionary modules. This is supported by the differences in its position between complex I and membrane-bound hydrogenases (Yu et al., 2020; 2018), and its susceptibility to dissociation from the membrane arm in *E. coli* (Baranova et al., 2007; Efremov and Sazanov, 2011) as expected for a late evolutionary addition (Levy et al., 2008). The initial association of the hydrogen-evolving module with an antiporter may have had an evolutionary advantage with the proton-translocating module serving as a source of protons (Yu et al., 2018), biasing H_2 evolution towards the reaction product or to enhance Na^+ extraction from the cells (Boyd et al., 2014).

Complex uncoupling is achieved by opening the Q-cavity to the solvent, consistent with elegant experiments by Cabrera-Orefice et.al (Cabrera-Orefice et al., 2018) in which the locking conserved the ^{A}TMH1-2 plug with the cysteine bridge reversibly uncoupled the enzyme. Close examination of the crosslinked structure indicates that crosslinking fixes the plug conformation in a way that the Q-cavity is accessible to the solvent.

The exact proton translocation mechanism within the antiporter-like module is unknown and requires further experimental and computational investigation. Here, we can only speculate that given the high conservation of ^{H}Glu157, it plays an important role in the coupling and likely changes its protonation state during the pumping cycle. Thus, it can influence the pKa of neighboring ionizable residues. In membrane-bound hydrogenase (MBH), an equivalent ^{M}Glu141 is separated from the closest ionizable ^{H}Lys409 by distance of 20 Å, which in a hydrophobic environment with a dielectric constant of 10, allow them to mutually modulate the pKa of each other by approximately 1 pH unit, similar to the free energy conserved upon ferredoxin oxidation by the protein complex. This distance is reduced to around 13 Å in membrane-bound sulfane sulfur reductase (MBS) and to around 6 Å in complex I, consistent with the proportionally higher free energy of catalyzed reactions (Yu et al., 2020; 2018).

The proposed coupling mechanism also suggests how the different conformational states associated with proton translocation might be trapped. A pH jump applied to purified coupled complex I will create a difference in potential between the Q-cavity and periplasmic surface, which depending on the direction of the jump, may trap different equilibrium conformations of this molecular machine.

Materials and Methods

Generation of an *E. coli* strain expressing Twin-Strep-tagged respiratory complex I

The native *nuo* operon encoding the 13 subunits of respiratory complex I (NuoA-N) was engineered with a Twin-Strep-tag (WSHPQFEKGGGSGGGSGGSAWSHPQFEK, IBA GmbH) at the N-terminus of NuoF using CRISPR-Cas9-enabled recombineering(Jiang et al., 2015). The DNA sequence encoding the C-terminal region of NuoE and N-terminus of NuoF was retrieved from GenBank (Acc. No. NC_012971.2 region 2288438 – 2289174). The tag-coding sequence followed by a TEV protease recognition site (Tropea et al., 2009) was appended upstream of the NuoF N-terminus and was codon-optimized, together with the 2288766–2288807 region of the genomic fragment. Such designed, linear DNA knock-in cassette was synthesized (GenScript). The vectors pCas and pTargetF were gifts from Sheng Yang (Addgene plasmids #62225 and #62226). The N20 sequence (GGTCAGCGGATGCGTTTCGG) was introduced into pTargetF by inverse PCR. Genomic engineering was performed according as described by Jiang *et al* (Jiang et al., 2015). Briefly, pCas vector was transformed into the chemically competent *E. coli* BL21AI strain (Thermo Fisher Scientific Inc.). The transformants were grown in shake-flask culture at 30°C in Lysogeny Broth (LB) medium containing 25 μg mL^{-1} (w/v) kanamycin monosulfate and 10 mM L-arabinose. Upon reaching OD_{600} 0.5, the bacteria were rendered electrocompetent and

were co-electroporated with the linear DNA cassette and the mutated pTargetF vector. The transformants were selected on LB-agar plates supplemented with 25 μg mL^{-1} (w/v) kanamycin and 50 μg mL^{-1} (w/v) streptomycin, or 50 μg mL^{-1} (w/v) spectinomycin. The positives, identified by colony PCR and DNA sequencing, were cured of the plasmids as described previously (Jiang et al., 2015). We further refer to the modified strain as *E. coli* BL21FS (NuoF-Strep).

Expression and purification of respiratory complex I

E. coli BL21FS was cultivated in LB medium for 48 hours at 37°C in a microaerobic environment. The cells were harvested by centrifugation and the membrane fraction was isolated as described by Sazanov *et al* (Sazanov, 2003). All subsequent steps were performed at 4°C. The homogenate was solubilized in 2% (w/v) n-Dodecyl β-D-maltoside (DDM, Anatrace) for 2 hours while stirring, after which the non-solubilized fraction was removed by ultracentrifugation at 225 000 × *g* for 1 hour. The supernatant was adjusted to 200 mM NaCl and loaded on a 5 mL Strep-Tactin® Superflow® high capacity column (IBA GmbH). After washing with 25 column volumes (CV) of buffer A (50 mM Bis-tris pH 6, 2 mM $CaCl_2$, 200 mM NaCl, 0.04% (w/v) DDM, 10% (v/v) sucrose, 0.003% (w/v) *E. coli* polar lipid extract (Avanti Polar Lipids, EPL), 0.2 mM PMSF), complex I was eluted with 2 CV of buffer B (buffer A containing 5 mM D-desthiobiotin (IBA GmbH). The purity of the eluted protein was assessed by SDS-PAGE and activity assays (Figure 1 - figure supplement 1). The purified complex I was concentrated using an Amicon Ultra-4 100K centrifugal filter (Merck) to 0.5 mg mL^{-1} (w/v), fast-frozen in liquid nitrogen and stored at -80°C.

Reconstitution of respiratory complex I into lipid nanodiscs

The membrane scaffold protein MSP2N2 was expressed and purified following a published protocol (Grinkova et al., 2010). Purified complex I at concentration 520 nM was mixed with 10.4 μM MSP2N2 (1:20 protein:MSP molar ratio) and incubated for 1 hour at 4°C.

Subsequently, the detergent was removed by adding 0.5 g mL^{-1} (w/v) Bio-Beads (Bio-Rad) overnight at 4°C. The reconstituted protein was further purified on the Superose 6 Increase 10/300 GL column (GE Healthcare) equilibrated in a buffer comprising 20 mM Bis-Tris pH 6.8, 200 mM NaCl and 2 mM $CaCl_2$. The protein-containing fractions were pooled and concentrated to 0.1–0.2 mg mL^{-1} (w/v) using Amicon Ultra-0.5 100K centrifugal concentrators.

Activity assays

NADH:ferricyanide (FeCy) and NADH:ubiquinone-1 (Q1) activities were measured as described previously (Sazanov, 2003). For the assays, 3 nM of detergent-solubilized or nanodisc-reconstituted complex I and either 1 mM FeCy (Sigma Aldrich BVBA) or 100 µM Q1 (Sigma Aldrich BVBA) were added to the assay buffer (10 mM Bis-Tris pH 6.8, 200 mM NaCl, 10 mM $CaCl_2$) in a stirred quartz cuvette at 30°C. The reaction was initiated by adding 100 µM NADH (Carl-Roth GmbH) and followed as reduction in absorbance at 340 nm using a Varian Cary 300 UV-Vis spectrophotometer (Agilent Technologies, Inc). During the inhibition assay, complex I was incubated with 20 µM Piericidin A (Cayman Chemical) for 5 min in the assay buffer at 30°C prior to Q1 addition.

Mass photometry

The composition of the protein preparation was assessed using mass photometry on a Refeyn OneMP instrument (Refeyn Ltd.), which was calibrated using an unstained native protein ladder (NativeMark™ Unstained Protein Standard A, Thermo Fisher Scientific Inc.). Measurements were performed on the reconstituted complex I at a concentration of 0.015 mg ml^{-1} using AcquireMP 2.2.0 software and were analyzed using the DiscoverMP 2.2.0 package.

Preparation of cryo-EM samples

The cryo-EM samples were prepared using a CP3 cryoplunge (Gatan). Quantifoil R0.6/1 Cu300 holey carbon grids were cleaned with chloroform, acetone, and isopropanol as described by Passmore *et al* (Passmore and Russo, 2016) . The grids were glow discharged in the ELMO glow discharge system (Corduan Technologies) from both sides for 2 min at 11 mA and 0.28 mbar. Four microliters of the reconstituted protein solution at 0.15 mg ml^{-1} concentration were applied on a grid and blotted from both sides for 2.2 s with Whatman No. 3 filter paper at 97 % relative humidity. The grid was then plunge-frozen in liquid ethane at -176°C and stored in liquid nitrogen.

Cryo-EM data collection

Cryo-EM images were collected on a JEOL CryoARM 300 microscope equipped with an in-column Ω energy filter (Fislage et al., 2020) at 300 kV, automatically using SerialEM 3.0.8 (Mastronarde, 2005) at a nominal magnification of 60,000 and the corresponding calibrated pixel size of 0.771 Å. Five images per single stage position were collected using a cross pattern with 3 holes along each axis (Efremov & Stroobants, 2021 in press). The 3 s exposures were dose-fractionated into 61 frames with an electron dose of 1.06 e- $Å^{-2}$ per frame. The energy filter slit was set to 20 eV width. In total, 9122 zero-loss micrographs were recorded with the defocus varying between -0.9 and -2.2 µm (Table 1).

EM image processing

The dose-fractionated movies were motion-corrected using MotionCor2 (Zheng et al., 2017) in the patch mode. The Contrast Transfer Function (CTF) parameters were estimated using CTFFIND-4.1 (Rohou and Grigorieff, 2015). 40 micrographs of various defoci were selected, manually picked, and used to train the neural network of crYOLO 1.4 (Wagner et al., 2019). After training, 1,256,734 particles were picked automatically from the complete dataset,

615 extracted in RELION 3.0 (Zivanov et al., 2018), and imported into cryoSPARC 2.11[11].
616 Following 2D classification, six initial models were generated, among which one
617 corresponded to the cytoplasmic arm-only and another corresponded to the complete complex
618 I. Using hetero-refinement, 441,265 and 525,680 particles were assigned to the cytoplasmic
619 arm and complete complex, respectively. Further processing was performed in RELION
620 3.1(Zivanov et al., 2020). After per-particle CTF estimation and Bayesian polishing, 3D auto-
621 refinement of the complete complex produced a map at an average resolution of 3.4 Å (Figure
622 1 - figure supplement 3). However, the map was very heterogeneous with the peripheral arm
623 resolved at 3.0–3.6 Å whereas the membrane arm was resolved at over 10 Å.
624 To address this heterogeneity, both arms were refined independently using multi-body
625 refinement (Nakane et al., 2018) (Figure 1 - figure supplement 3) and the peripheral domain
626 signal was subtracted. After two rounds of 3D classification applied to the membrane domain
627 and nanodisc signal subtraction, a subset of 48,745 particles was 3D refined to an average
628 resolution of 3.6 Å. However, the density map was anisotropic. To improve the
629 reconstruction, the original stack of 525,680 particles was refined against the masked
630 cytoplasmic arm, followed by subtraction of the signal from the peripheral arm. Next,
631 membrane arm map obtained above was filtered to 9 Å and used as an initial model for the 3D
632 refinement of all resulting membrane arm particles. Next, to prevent model bias, the refined
633 map was low-pass filtered to 20 Å and used in the subsequent 3D classification with 10
634 classes, τ of 12 and 24° local angular search range and 1.8° angular step. The best class
635 (110,258 particles and 8 Å resolution) was auto-refined using the starting model low-pass
636 filtered to 15 Å, which produced the reconstruction to a resolution of 4.4 Å. Next, the
637 nanodisc density was subtracted, which further improved the resolution to 3.9 Å. Following
638 3D classification without alignment with τ of 40, 8 classes, and resolution in the E-step
639 limited to 4 Å, a subset of 37,441 particles was identified, which after auto refinement,

produced a density map at an average resolution of 3.9 Å with better resolved peripheral regions. Finally, density modification with the resolve_cryo_em tool available in PHENIX 1.18.2 (Terwilliger et al., 2020) improved the resolution to 3.7 Å (Figure 1 - figure supplement 3,4).

After multibody refinement of the arms described above, peripheral arm particles with the subtracted membrane arm were 3D classified into 12 classes without alignment using τ of 40 and resolution of the expectation step limited to 4 Å. The best class contained 134,976 particles and was further refined to 2.9 Å resolution.

A subset of 166,580 particles was selected after a similar 3D classification procedure that was applied to the 441,265 particles of dissociated peripheral arm particles. It was further cleaned from the remaining particles of the complete complex I by 2D classification, resulting in a subset of 151,357 particles that produced a density map to a resolution of 3.0 Å. As the reconstructions of the dissociated and membrane arm-subtracted cytoplasmic arms were virtually identical, both stacks were combined. After two cycles of per-particle CTF refinement, aberration corrections, and Bayesian particle polishing in RELION 3.1, the resolution improved to 2.4 Å. Consecutive density modification in PHENIX further improved the resolution to 2.1 Å (Figure 1 - figure supplement 3,4, Table 1).

To resolve the conformation of entire complex I, a stack of 525,680 particles was aligned to the peripheral arm using auto-refinement with a mask around the peripheral arm in RELION 3.1. Next, 3D classification without alignment into 30 classes with resolution of the expectation step limited to 20 Å and τ of 4 was performed, followed by auto-refinement of each resulting class, which produced maps to a resolution in the range of 9-20 Å (some of the classes are shown in Figure 1 - figure supplement 5).

Three high-resolution conformations of complete complex I were obtained as follows. Conformation 1 was resolved by applying the 3D classification into 15 classes, τ of 6, a 24°

665 local angular search range, and 1.8° sampling interval to the subset of 110,258 particles that
666 produced the 3.9 Å reconstruction of the membrane arm (see above). The best class consisted
667 of 23,445 particles that were refined to a resolution of 3.9 Å.
668 Conformations 2 and 3, were identified by applying 3D classification without image
669 alignment into 12 classes with τ of 40 and resolution of the expectation step limited to 4 Å, to
670 the stack of 525,680 intact complex I particles. Two of the best classes, consisting of 21,620
671 and 21,234 particles were refined to 4.6 Å and 4.5 Å, respectively. Following density-
672 modification in PHENIX, the resolution of the maps was improved to 3.3, 3.8, and 3.7 Å, for
673 conformations 1, 2, and 3, respectively (Figure 1 - figure supplement 4C, Table 1).

674 ## Model building

675 Peripheral arm subunits constituting NuoB, CD, E, F, G, and I were first homology modelled
676 in the SWISS-MODEL server (Waterhouse et al., 2018) based on the structure of *T.*
677 *thermophilus* (PDB ID:4HEA (Baradaran et al., 2013)) and were rigid-body fitted into the
678 density map in UCSF Chimera 1.13.1. Following manual rebuilding in Coot 0.9 (Casañal et
679 al., 2020), the model was subjected to real-space refinement against the final 2.1 Å map of
680 the cytoplasmic arm in PHENIX 1.18.2 using the default parameters. Secondary structure
681 restrains were applied only to the interfacial region resolved at a lower resolution. The value
682 of the nonbonded_weight parameter was optimized. Water molecules were added to the map
683 and validated using the "Check/delete waters" tool in Coot 0.9 . Molecular dynamics-based
684 model idealization was conducted in ISOLDE 1.0b5 (Croll, 2018), followed by several
685 iterations of real-space refinement without atomic displacement parameter (ADP) restraints
686 and manual rebuilding in Coot 0.9.
687 For the membrane domain, the previously obtained *E. coli* model (PDB ID: 3RKO) was real-
688 space-refined in PHENIX. The missing NuoH subunit was homology-modelled using the *T.*
689 *thermophilus* structure (PDB ID: 4HEA) in Coot 0.9. The final model was obtained after

several rounds of manual rebuilding and real-space refinement using standard parameters with Ramachandran restrains, secondary-structure restrains applied to the NuoL TMH9-13, without ADP restrains, and with the optimized nonbonded_weight parameter. To generate the model of the complete complex I, the separate peripheral and membrane arm structures were combined and the missing parts at the interface (Table 2) were built manually. As the density of NuoL and NuoM was very poor in all the resolved full conformations, these subunits were subjected to rigid-body refinement in PHENIX, whereas the others were subjected to real-space refinement with minimization_global, local_grid_search, morphing, and ADP refinement. Ramachandran, ADP, and secondary-structure restrains were used. After manual rebuilding in Coot, real-space refinement of the full complex was performed with standard parameters and restrains. The models were validated in MolProbity (Williams et al., 2018) . Structural conservation was evaluated using the ConSurf server (Ashkenazy et al., 2016). The figures and movies were generated in UCSF ChimeraX version 1.1. (Goddard et al., 2018) and PyMOL (The PyMOL Molecular Graphics System, Version 2.4.1 Schrödinger, LLC).

Acknowledgements

We are indebted to Henri De Greve for help with establishing CRISPR-Cas9 for *E. coli* complex I. We are thankful to Dr. Adam Schröfel and Dr. Marcus Fislage for providing support during cryo-EM data collection, to Annelore Stroobants for technical assistance and to Lukasz Milewski for assistance in data processing. We kindly thank VIB Tech Watch fund for facilitating access to the Refeyn instrument. We would like to acknowledge the funding provided by Vlaams Instituut voor Biotechnologie, Fonds Wetenschappelijk Onderzoek (grant Nos. G0H5916N, G.0266.15N) and by the European Research Council.

Data availability

Cryo-EM density maps and atomic models are deposited into the PDB and EMDB databases with the following accession codes: cytoplasmic domain (PDB ID: 7NZ1, EMD-12661), membrane domain (PDB ID: 7NYH, EMD-12652), entire complex conformation 1 (PDB ID: 7NYR, EMD-12653), conformation 2 (PDB ID: 7NYU, EMD-12654), conformation 3 (PDB ID: 7NYV, EMD-12655).

The following data sets were generated:

Kolata P, Efremov RG (2021). Electron Microscopy Data Bank ID EMD-12661. Respiratory complex I from Escherichia coli - focused refinement of cytoplasmic arm. https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-12661
Kolata P, Efremov RG (2021). Electron Microscopy Data Bank ID EMD-12652. Respiratory complex I from Escherichia coli - focused refinement of membrane arm. https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-12652
Kolata P, Efremov RG (2021). Electron Microscopy Data Bank ID EMD-12653. Respiratory complex I from Escherichia coli - conformation 1. https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-12653

Kolata P, Efremov RG (2021). Electron Microscopy Data Bank ID EMD-12654. Respiratory complex I from Escherichia coli - conformation 2. https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-12654
Kolata P, Efremov RG (2021). Electron Microscopy Data Bank ID EMD-12655. Respiratory complex I from Escherichia coli - conformation 3. https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-12655

Kolata P, Efremov RG (2021). RCSB Protein Data Bank ID 7NZ1. Respiratory complex I from Escherichia coli - focused refinement of cytoplasmic arm. https://www.rcsb.org/structure/7NZ1
Kolata P, Efremov RG (2021). RCSB Protein Data Bank ID 7NYH. Respiratory complex I from Escherichia coli - focused refinement of membrane arm. https://www.rcsb.org/structure/7NYH
Kolata P, Efremov RG (2021). RCSB Protein Data Bank ID 7NYR. Respiratory complex I from Escherichia coli - conformation 1. https://www.rcsb.org/structure/7NYR
Kolata P, Efremov RG (2021). RCSB Protein Data Bank ID 7NYU. Respiratory complex I from Escherichia coli - conformation 2. https://www.rcsb.org/structure/7NYU
Kolata P, Efremov RG (2021). RCSB Protein Data Bank ID 7NYV. Respiratory complex I from Escherichia coli - conformation 3. https://www.rcsb.org/structure/7NYV

Competing interests

Authors declare no competing interests.

References

Agip A-NA, Blaza JN, Bridges HR, Viscomi C, Rawson S, Muench SP, Hirst J. 2018. Cryo-EM structures of complex I from mouse heart mitochondria in two biochemically defined states. *Nat Struct Mol Biol* **25**:1–13. doi:10.1038/s41594-018-0073-1

Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, Ben-Tal N. 2016. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Research* **44**:W344–50. doi:10.1093/nar/gkw408

Astumian RD, Mukherjee S, Warshel A. 2016. The Physics and Physical Chemistry of Molecular Machines. *ChemPhysChem* **17**:1719–1741. doi:10.1002/cphc.201600184

Babot M, Birch A, Labarbuta P, Galkin A. 2014. Characterisation of the active/de-active transition of mitochondrial complex I. *Biochim Biophys Acta* **1837**:1083–1092. doi:10.1016/j.bbabio.2014.02.018

Baradaran R, Berrisford JM, Minhas GS, Sazanov LA. 2013. Crystal structure of the entire respiratory complex I. *Nature* **494**:443–448. doi:10.1038/nature11871

Baranova EA, Holt PJ, Sazanov LA. 2007. Projection structure of the membrane domain of Escherichia coli respiratory complex I at 8 A resolution. *Journal of Molecular Biology* **366**:140–154. doi:10.1016/j.jmb.2006.11.026

Belevich N, Ballmoos von C, Verkhovskaya M. 2017. Activation of Proton Translocation by Respiratory Complex I. *Biochemistry* **56**:5691–5697. doi:10.1021/acs.biochem.7b00727

Birrell JA, Morina K, Bridges HR, Friedrich T, Hirst J. 2013. Investigating the function of [2Fe–2S] cluster N1a, the off-pathway cluster in complex I, by manipulating its reduction potential. *Biochemical Journal* **456**:139–146. doi:10.1042/BJ20130606

Boyd ES, Schut GJ, Adams M, Microbe JP. 2014. Hydrogen metabolism and the evolution of biological respiration. *Microbe* **9**:361–367.

Cabrera-Orefice A, Yoga EG, Wirth C, Siegmund K, Zwicker K, Guerrero-Castillo S, Zickermann V, Hunte C, Brandt U. 2018. Locking loop movement in the ubiquinone pocket of complex I disengages the proton pumps. *Nature Communications* **9**:1–10. doi:10.1038/s41467-018-06955-y

775 Carroll J, Ding S, Fearnley IM, Walker JE. 2013. Post-translational Modifications near the
776 Quinone Binding Site of Mammalian Complex I. *Journal of Biological Chemistry*
777 **288**:24799–24808. doi:10.1074/jbc.M113.488106
778 Casañal A, Lohkamp B, Emsley P. 2020. Current developments in Coot for macromolecular
779 model building of Electron Cryo-microscopy and Crystallographic Data. *Protein Sci*
780 **29**:1069–1078. doi:10.1002/pro.3791
781 Chambers JQ. 1988. Electrochemistry of quinones, The Quinonoid Compounds: Vol. 1
782 (1988). Chichester, UK: John Wiley & Sons, Ltd. doi:10.1002/9780470772119.ch12
783 Croll TI. 2018. ISOLDE: a physically realistic environment for model building into low-
784 resolution electron-density maps. *Acta Crystallogr D Struct Biol* **74**:519–530.
785 doi:10.1107/S2059798318002425
786 Denke E, Merbitz-Zahradnik T, Hatzfeld OM, Snyder CH, Link TA, Trumpower BL. 1998.
787 Alteration of the Midpoint Potential and Catalytic Activity of the Rieske Iron-Sulfur
788 Protein by Changes of Amino Acids Forming Hydrogen Bonds to the Iron-Sulfur Cluster.
789 *Journal of Biological Chemistry* **273**:9085–9093. doi:10.1074/jbc.273.15.9085
790 Efremov RG, Sazanov LA. 2012. The coupling mechanism of respiratory complex I - a
791 structural and evolutionary perspective. *Biochim Biophys Acta* **1817**:1785–1795.
792 doi:10.1016/j.bbabio.2012.02.015
793 Efremov RG, Sazanov LA. 2011. Structure of the membrane domain of respiratory complex I.
794 *Nature* **476**:414–420. doi:10.1038/nature10330
795 Fiedorczuk K, Letts JA, Degliesposti G, Kaszuba K, Skehel M, Sazanov LA. 2016. Atomic
796 structure of the entire mammalian mitochondrial complex I. *Nature* **538**:406–410.
797 doi:10.1038/nature19794
798 Fislage M, Shkumatov AV, Stroobants A, Efremov RG. 2020. Assessing the JEOL CRYO
799 ARM 300 for high-throughput automated single-particle cryo-EM in a multiuser
800 environment. *IUCrJ* **7**:707–718. doi:10.1107/S2052252520006065
801 Fraaije MW, sciences AMTIB, 2000. 2000. Flavoenzymes: diverse catalysts with recurrent
802 features. *Elsevier*
803 **25**:126–132. doi:10.1016/S0968-0004(99)01533-9
804 Fritz G, Roth A, Schiffer A, Büchert T, Bourenkov G, Bartunik HD, Huber H, Stetter KO,
805 Kroneck PMH, Ermler U. 2002. Structure of adenylylsulfate reductase from the
806 hyperthermophilic Archaeoglobus fulgidus at 1.6-Å resolution. *Proceedings of the*
807 *National Academy of Sciences* **99**:1836–1841. doi:10.1073/pnas.042664399
808 Galkin A, Drose S, Brandt U. 2006. The proton pumping stoichiometry of purified
809 mitochondrial complex I reconstituted into proteoliposomes. *Biochimica et Biophysica*
810 *Acta (BBA) - Bioenergetics* **1757**:1575–1581. doi:10.1016/j.bbabio.2006.10.001
811 Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, Ferrin TE. 2018.
812 UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci*
813 **27**:14–25. doi:10.1002/pro.3235
814 Grba DN, Hirst J. 2020. Mitochondrial complex I structure reveals ordered water molecules
815 for catalysis and proton translocation. *Nat Struct Mol Biol* 1–20. doi:10.1038/s41594-020-
816 0473-x
817 Grinkova YV, Denisov IG, Sligar SG. 2010. Engineering extended membrane scaffold
818 proteins for self-assembly of soluble nanoscale lipid bilayers. *Protein Engineering Design*
819 *and Selection* **23**:843–848. doi:10.1093/protein/gzq060
820 Gutiérrez-Fernández J, Kaszuba K, Minhas GS, Baradaran R, Tambalo M, Gallagher DT,
821 Sazanov LA. 2020. Key role of quinone in the mechanism of respiratory complex I.
822 *Nature Communications* **11**:1–17. doi:10.1038/s41467-020-17957-0

Hirst J, Roessler MM. 2016. Energy conversion, redox catalysis and generation of reactive oxygen species by respiratory complex I. *BBA - Bioenergetics* **1857**:872–883. doi:10.1016/j.bbabio.2015.12.009

Holt PJ, Efremov RG, Nakamaru-Ogiso E, Sazanov LA. 2016. Reversible FMN dissociation from Escherichia coli respiratory complex I. *BBA - Bioenergetics* **1857**:1777–1785. doi:10.1016/j.bbabio.2016.08.008

Hunte C, Zickermann V, Brandt U. 2010. Functional Modules and Structural Basis of Conformational Coupling in Mitochondrial Complex I. *Science* **329**:448–451. doi:10.1126/science.1191046

Jardetzky O. 1966. Simple allosteric model for membrane pumps [27]. *Nature* **211**:969–970. doi:10.1038/211969a0

Jiang Y, Chen B, Duan C, Sun B, Yang J, Yang S. 2015. Multigene editing in the Escherichia coli genome via the CRISPR-Cas9 system. *Appl Environ Microbiol* **81**:2506–2514. doi:10.1128/AEM.04023-14

Kampjut D, Sazanov LA. 2020. The coupling mechanism of mammalian respiratory complex I. *Science* **370**:eabc4209–18. doi:10.1126/science.abc4209

Leif H, Sled VD, Ohnishi T, Weiss H, Friedrich T. 1995. Isolation and characterization of the proton-translocating NADH: ubiquinone oxidoreductase from Escherichia coli. *Eur J Biochem* **230**:538–548.

Lemmer C, Bouvet M, Meunier-Prest R. 2011. Proton coupled electron transfer of ubiquinone Q2 incorporated in a self-assembled monolayer. *Phys Chem Chem Phys* **13**:13327–6. doi:10.1039/c0cp02700f

Levy ED, Erba EB, Robinson CV, Teichmann SA. 2008. Assembly reflects evolution of protein complexes. *Nature* **453**:1262–1265. doi:10.1038/nature06942

Mastronarde DN. 2005. Automated electron microscope tomography using robust prediction of specimen movements. *Journal of Structural Biology* **152**:36–51. doi:10.1016/j.jsb.2005.07.007

Morgan DJ, Sazanov LA. 2008. Three-dimensional structure of respiratory complex I from Escherichia coli in ice in the presence of nucleotides. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1777**:711–718. doi:10.1016/j.bbabio.2008.03.023

Morozenko A, Stuchebrukhov AA. 2016. Dowser++, a new method of hydrating protein structures. *Proteins* **84**:1347–1357. doi:10.1002/prot.25081

Nakane T, Kimanius D, Lindahl E, Scheres SH. 2018. Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. *eLife* **7**:1485. doi:10.7554/eLife.36861

Onsager L, Hemmer PC, Holden H. 1996. Lars Onsager 1075.

Parey K, Brandt U, Xie H, Mills DJ, Siegmund K, Vonck J, Kühlbrandt W, Zickermann V. 2018. Cryo-EM structure of respiratory complex I at work. *eLife* **7**. doi:10.7554/eLife.39213

Passmore LA, Russo CJ. 2016. Specimen Preparation for High-Resolution Cryo-EM. *Meth Enzymol* **579**:51–86. doi:10.1016/bs.mie.2016.04.011

Rohou A, Grigorieff N. 2015. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of Structural Biology* **192**:216–221. doi:10.1016/j.jsb.2015.08.008

Sazanov LA. 2015. A giant molecular proton pump: structure and mechanism of respiratory complex I. *Nat Rev Mol Cell Biol* **16**:375–388. doi:10.1038/nrm3997

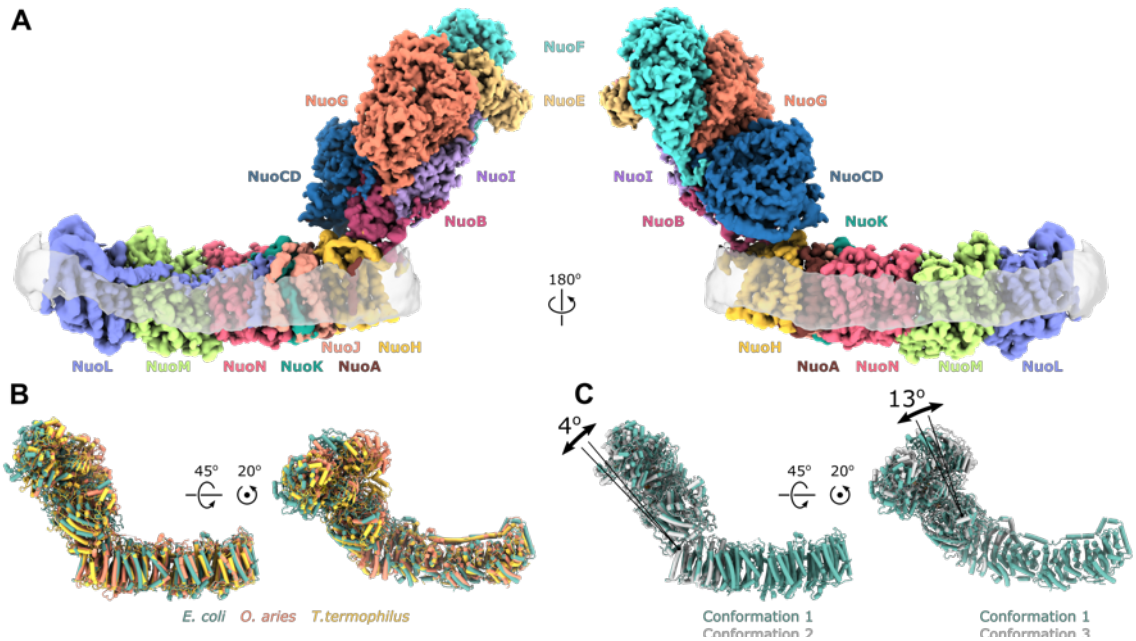
Sazanov LA. 2006. Structure of the Hydrophilic Domain of Respiratory Complex I From Thermus Thermophilus. *Science* **311**:1430–1436. doi:10.1126/science.1123809

Sazanov LA. 2003. A Role for Native Lipids in the Stabilization and Two-dimensional Crystallization of the Escherichia coli NADH-Ubiquinone Oxidoreductase (Complex I). *Journal of Biological Chemistry* **278**:19483–19491. doi:10.1074/jbc.M208959200

873 Schulte M, Frick K, Gnandt E, Jurkovic S, Burschel S, Labatzke R, Aierstock K, Fiegen D,
874 Wohlwend D, Gerhardt S, Einsle O, Friedrich T. 2019a. A mechanism to prevent
875 production of reactive oxygen species by Escherichia coli respiratory complex I. *Nature*
876 *Communications* **10**:2551–9. doi:10.1038/s41467-019-10429-0
877 Schulte M, Frick K, Gnandt E, Jurkovic S, Burschel S, Labatzke R, Aierstock K, Fiegen D,
878 Wohlwend D, Gerhardt S, Einsle O, Friedrich T. 2019b. A mechanism to prevent
879 production of reactive oxygen species by Escherichia coli respiratory complex I. *Nature*
880 *Communications* **10**:1–9. doi:10.1038/s41467-019-10429-0
881 Steiner J, Sazanov L. 2020. Structure and mechanism of the Mrp complex, an ancient
882 cation/proton antiporter. *eLife* **9**:1–47. doi:10.7554/eLife.59407
883 Terwilliger TC, Ludtke SJ, Read RJ, Adams PD, Afonine PV. 2020. Improvement of cryo-
884 EM maps by density modification. *Nat Meth* **17**:923–927. doi:10.1038/s41592-020-0914-
885 9
886 Tropea JE, Cherry S, Waugh DS. 2009. Expression and purification of soluble His(6)-tagged
887 TEV protease. *Methods Mol Biol* **498**:297–307. doi:10.1007/978-1-59745-196-3_19
888 Verkhovskaya M, Bloch DA. 2012. Energy-converting respiratory Complex I: On the way to
889 the molecular mechanism of the proton pump. *International Journal of Biochemistry and*
890 *Cell Biology* **45**:1–21. doi:10.1016/j.biocel.2012.08.024
891 Verkhovskaya ML, Belevich N, Euro L, Wikström M, Verkhovsky MI. 2008. Real-time
892 electron transfer in respiratory complex I. *Proceedings of the National Academy of*
893 *Sciences* **105**:3763–3767. doi:10.1073/pnas.0711249105
894 Vinothkumar KR, Zhu J, Hirst J. 2014. Architecture of mammalian respiratory complex I.
895 *Nature* **515**:80–84. doi:10.1038/nature13686
896 Wagner T, Merino F, Stabrin M, Moriya T, Antoni C, Apelbaum A, Hagel P, Sitsel O, Raisch
897 T, Prumbaum D, Quentin D, Roderer D, Tacke S, Siebolds B, Schubert E, Shaikh TR,
898 Lill P, Gatsogiannis C, Raunser S. 2019. SPHIRE-crYOLO is a fast and accurate fully
899 automated particle picker for cryo-EM. *Communications Biology* **2**:1–13.
900 doi:10.1038/s42003-019-0437-z
901 Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer
902 TAP, Rempfer C, Bordoli L, Lepore R, Schwede T. 2018. SWISS-MODEL: homology
903 modelling of protein structures and complexes. *Nucleic Acids Research* **46**:W296–W303.
904 doi:10.1093/nar/gky427
905 Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, Verma V, Keedy
906 DA, Hintze BJ, Chen VB, Jain S, Lewis SM, Arendall WB, Snoeyink J, Adams PD,
907 Lovell SC, Richardson JS, Richardson DC. 2018. MolProbity: More and better reference
908 data for improved all-atom structure validation. *Protein Sci* **27**:293–315.
909 doi:10.1002/pro.3330
910 Wirth C, Brandt U, Hunte C, Zickermann V. 2016. Structure and function of mitochondrial
911 complex I. *BBA - Bioenergetics* **1857**:902–914. doi:10.1016/j.bbabio.2016.02.013
912 Yagi T, Matsuno-Yagi A. 2003. The Proton-Translocating NADH−Quinone Oxidoreductase
913 in the Respiratory Chain: The Secret Unlocked †. *Biochemistry* **42**:2266–2274.
914 doi:10.1021/bi027158b
915 Yano T, Dunham WR, Ohnishi T. 2005. Characterization of the delta muH+-sensitive
916 ubisemiquinone species (SQ(Nf)) and the interaction with cluster N2: new insight into the
917 energy-coupled electron transfer in complex I. *Biochemistry* **44**:1744–1754.
918 doi:10.1021/bi048132i
919 Yu H, Haja DK, Schut GJ, Wu C-H, Meng X, Zhao G, Li H, Adams MWW. 2020. Structure
920 of the respiratory MBS complex reveals iron-sulfur cluster catalyzed sulfane sulfur
921 reduction in ancient life. *Nature Communications* **11**:1–13. doi:10.1038/s41467-020-
922 19697-7

923 Yu H, Wu C-H, Schut GJ, Haja DK, Zhao G, Peters JW, Adams MWW, Li H. 2018. Structure
924 of an Ancient Respiratory System. *Cell* **173**:1636–1649.e16.
925 doi:10.1016/j.cell.2018.03.071
926 Zheng H, Chruszcz M, Lasota P, Lebioda L, Minor W. 2008. Data mining of metal ion
927 environments present in protein structures. *Journal of Inorganic Biochemistry* **102**:1765–
928 1776. doi:10.1016/j.jinorgbio.2008.05.006
929 Zheng SQ, Palovcak E, Armache J-P, Verba KA, Cheng Y, Agard DA. 2017. MotionCor2:
930 anisotropic correction of beam-induced motion for improved cryo-electron microscopy.
931 *Nature Publishing Group* **14**:331–332. doi:10.1038/nmeth.4193
932 Zhu J, Vinothkumar KR, Hirst J. 2016. Structure of mammalian respiratory complex I. *Nature*
933 **536**:1–17. doi:10.1038/nature19095
934 Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJH, Lindahl E, Scheres SHW.
935 2018. New tools for automated high-resolution cryo-EM structure determination in
936 RELION-3. *eLife* **7**:163. doi:10.7554/eLife.42166
937 Zivanov J, Nakane T, Scheres SHW. 2020. Estimation of high-order aberrations and
938 anisotropic magnification from cryo-EM data sets in RELION-3.1. *IUCrJ* **7**:253–267.
939 doi:10.1107/S2052252520000081
940 Zu Y, Di Bernardo S, Yagi T, Biochemistry JH, 2002. 2002. Redox properties of the [2Fe-2S]
941 center in the 24 kDa (NQO2) subunit of NADH: ubiquinone oxidoreductase (complex I).
942 *ACS Publications*
943 **41**:10056–10069. doi:10.1021/bi026026f
944

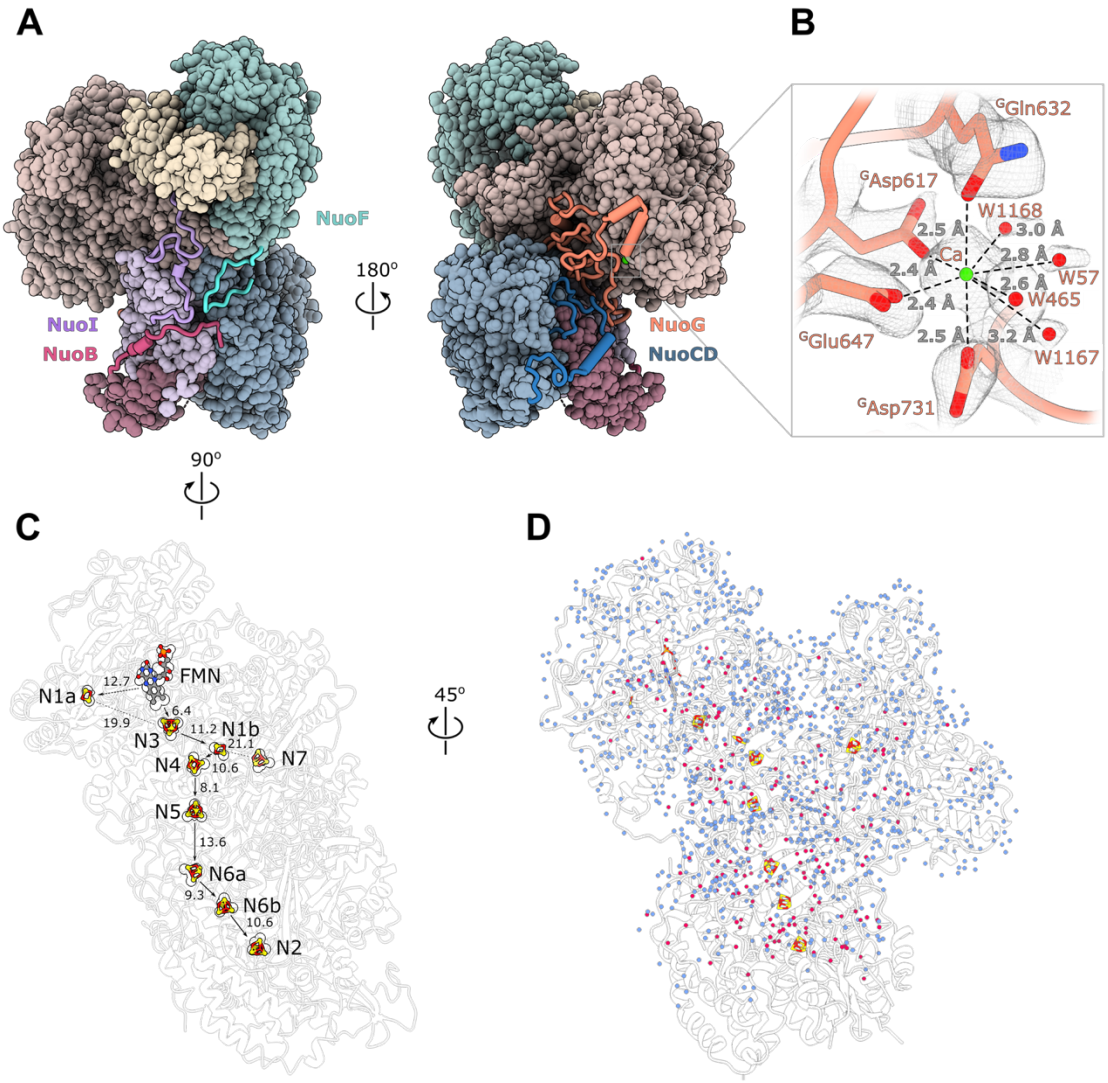
945 # Figures



946
947

948 **Figure 1 Architecture of *Escherichia coli* respiratory complex I**. (A) Segmented density
949 map of the complete complex I shown together with the nanodisc density. (B) Comparison of
950 the structures of the *E. coli* (green), *Thermus thermophilus* (PDB ID: 4HEA, yellow), and the
951 core subunits of ovine (PDB ID: 6ZKD, orange) complex I. (C) Conformational differences
952 between three conformations resolved at high resolution. The structures are aligned on the
953 membrane arm. The rotation axes and angles are indicated.

954



955

956 **Figure 2 Structure of the peripheral arm.** (A, B) The *Escherichia coli*-specific extensions
957 in the peripheral arm subunits: (A) C-termini of NuoI (violet), NuoB (pink), NuoF
958 (turquoise), (B) NuoG insertion (orange), and NuoCD linker (blue). (C) Structural details of
959 the calcium-binding site. (D) Comparison of the FMN and Fe-S clusters positions in *E. coli*
960 (shown as atoms) and *Thermus thermophilus* (shown as outline around *E. coli* atoms). Edge-
961 to-edge distances and the electron pathway are indicated. (E) Water molecules modelled into
962 the 2.1 Å resolution density of the peripheral arm are show in blue. Water molecules
963 conserved with the peripheral arm of ovine complex I (PDB ID: 6ZK9, red) are shown as red
964 spheres. FMN and iron-sulfur (Fe-S) clusters are shown as spheres.

965

966

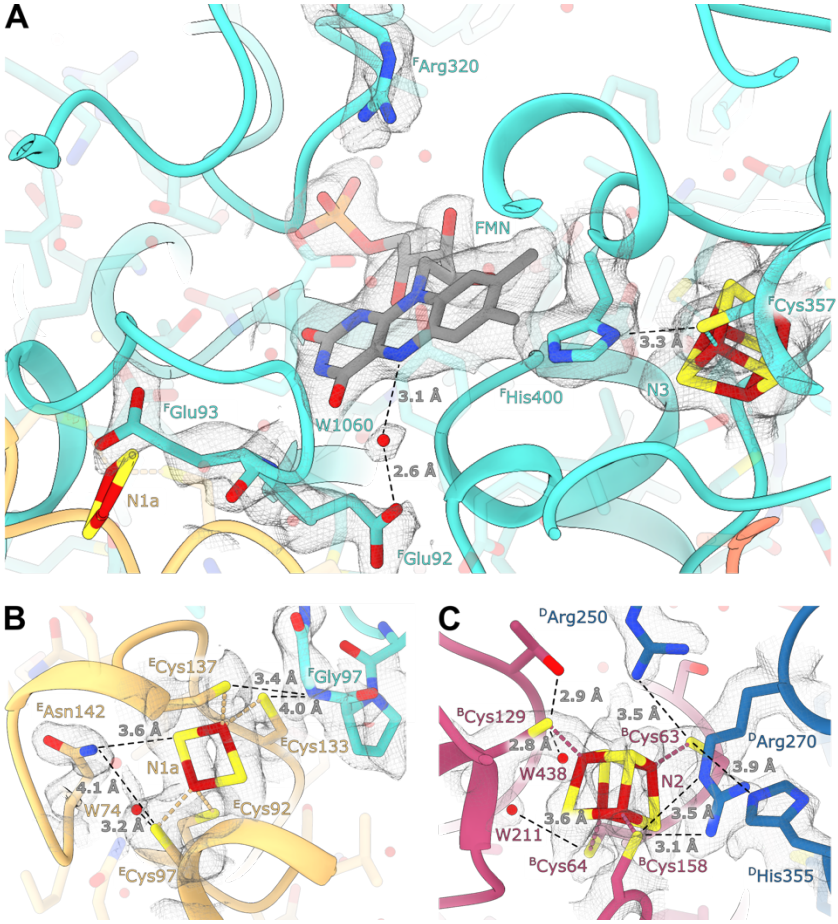


Figure 3 Details of the electron transport chain. (A) The NADH-binding pocket and environment of the Fe-S cluster N3. (B, C) Environment of the Fe-S clusters N1a and N2.

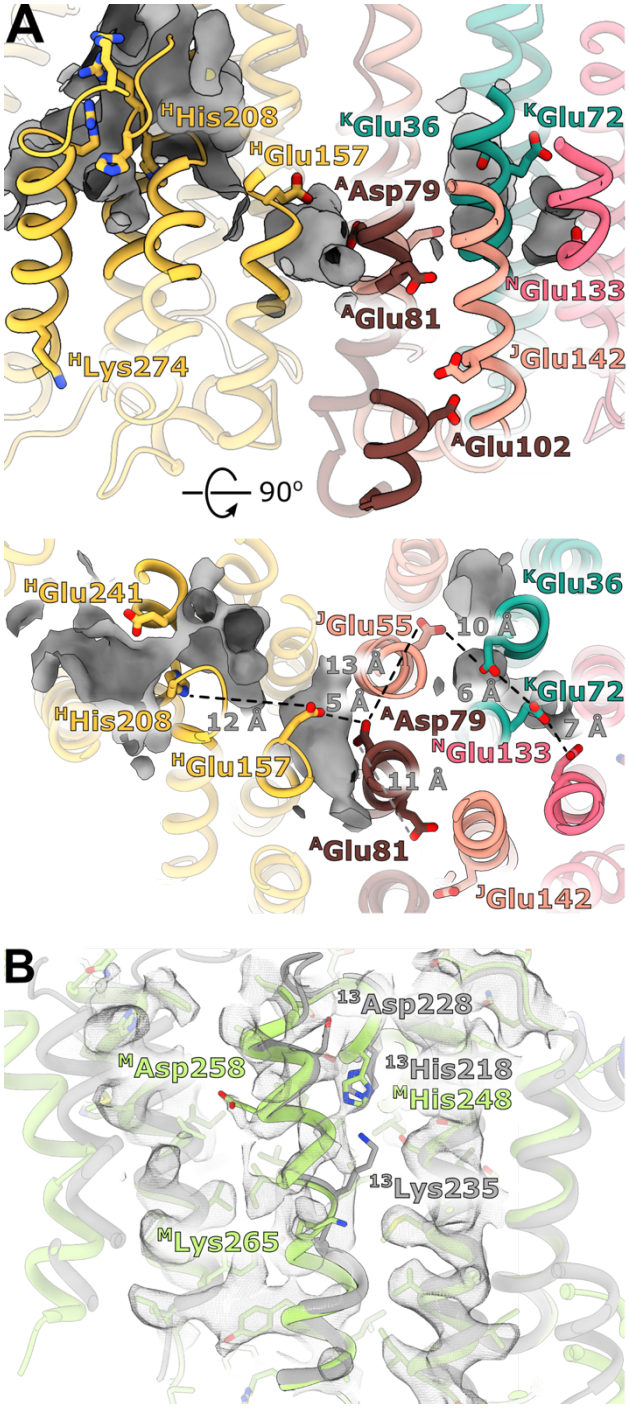


Figure 4 Structural details of the membrane arm. (A) The E-channel. Top: Side view, bottom: view from the cytoplasm. Charged residues between NuoH and NuoN subunits are indicated as along with the distances between them. The cavities allowing entrance of ions and water molecules are shown as grey surfaces. (B) Conformational heterogeneity within the NuoM subunit. The *E. coli* structure and amino acids are green-colored, whereas the aligned structure of *T. thermophilus* is grey. The *E. coli* membrane arm density is depicted.

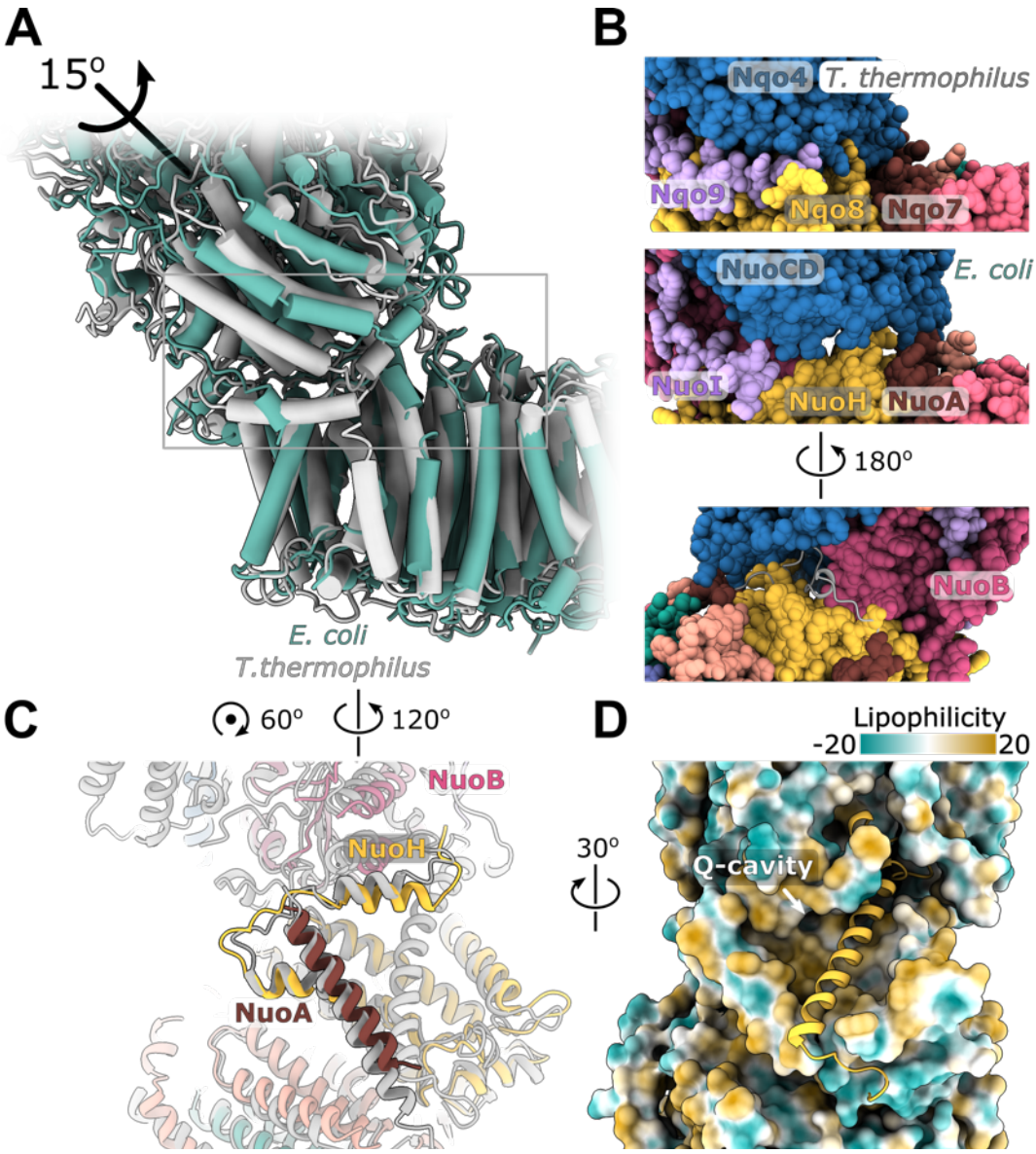
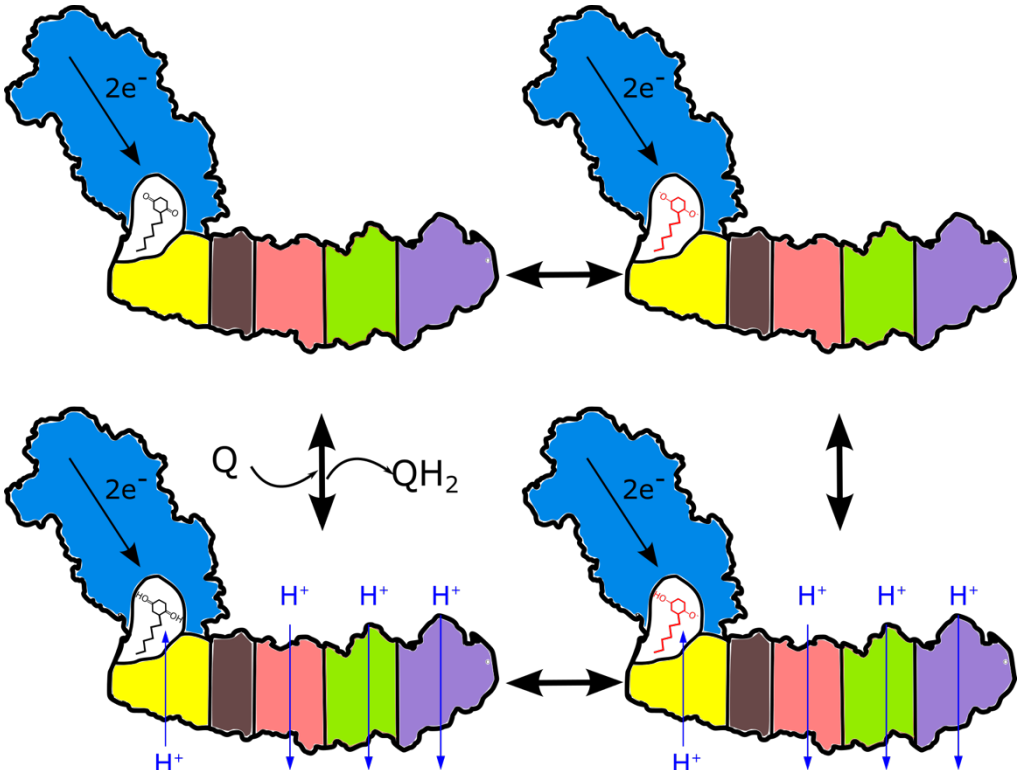


Figure 5 Interface between the peripheral and membrane arms. (A) Comparison of the interface between *E. coli* (green) and *T. thermophilus* (PDB ID:4HEA, grey) complex I. Structures were aligned on the subunit NuoH/Nqo8. The rotation axis of the subunits NuoB/D module relative to Nqo6/4 is indicated. (B) Interfacial contacts between the peripheral and membrane arms in *T. thermophilus* (upper panel) and *E. coli* (middle and bottom panel). A gap in the subunit interface is apparent in the absence of the conserved ^{A}TMH1-2 loop fragment. The corresponding loop from *T. thermophilus* is shown in grey in the cartoon representation (bottom panel). (C) Differences in the structures of NuoH and NuoA subunits between *E. coli* (colour coded as in Figure 1) and *T. thermophilus* (grey). (D) View from the membrane on the entrance to the Q-cavity. Homology model of ^{H}TM1, absent in the *E. coli* structure, is shown in the cartoon representation. The protein surface is coloured by lipophilicity.

997



998

999 **Figure 6** Proposed mechanism of coupling in respiratory complex I. Ubiquinone reduction
1000 decreases the proton potential in the Q-cavity, generating/enhancing electrochemical potential
1001 between the Q-cavity and periplasmic space. It is subsequently neutralized by protons
1002 translocated through NuoH from the periplasm into the Q-cavity and this translocation is
1003 coupled with the reversible translocation of three protons into the periplasm. Colour coding of
1004 schematic subunits in the membrane arm is similar to that described in Figure 1, negatively
1005 charged states of ubiquinone are shown in red.

1006

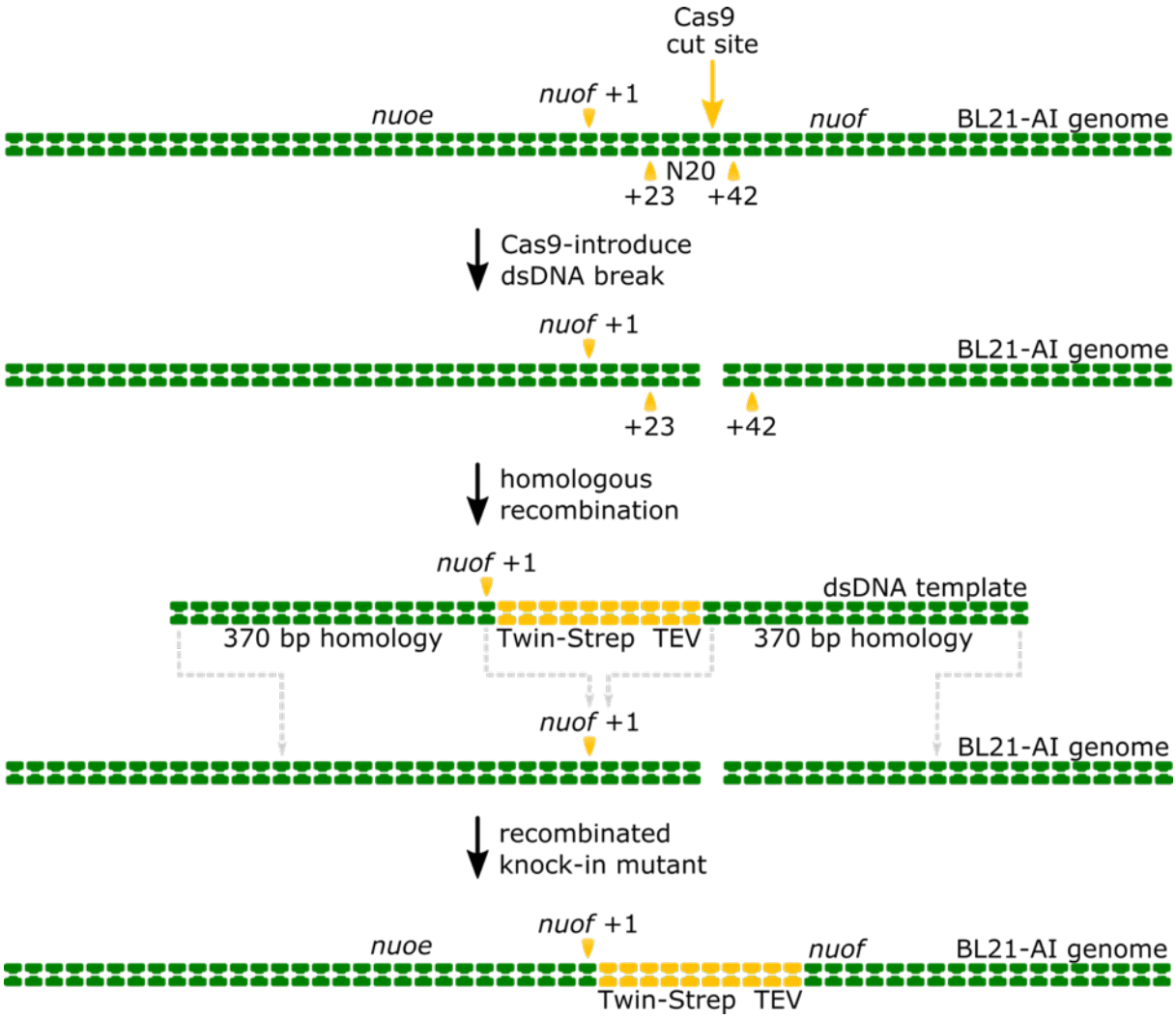
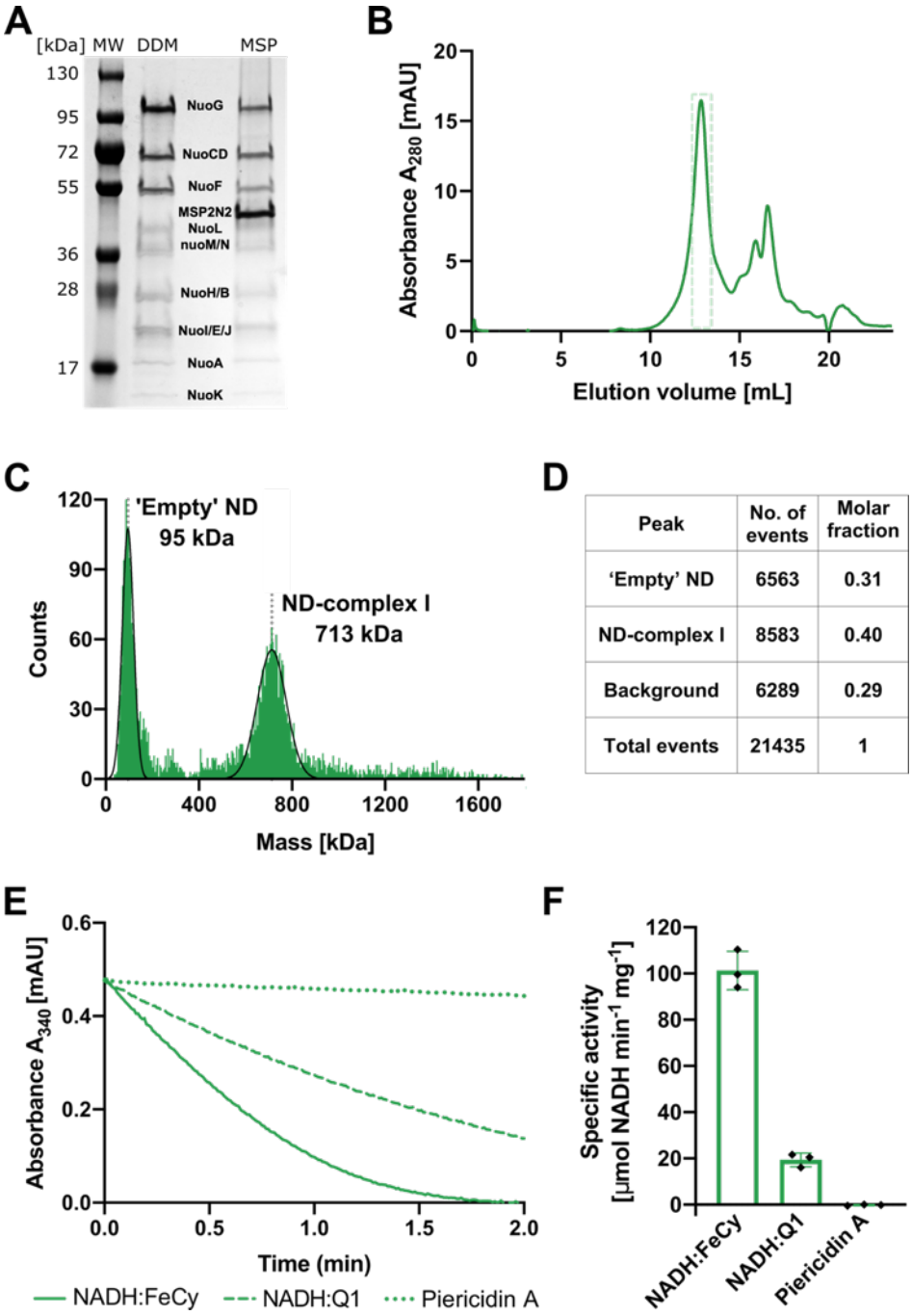


Figure 1 - figure supplement 1 Schematic representation of Crispr-Cas9-enabled incorporation of the twin-strep tag into the N-terminus of the genomically-encoded NuoF subunit. The Cas9 enzyme introduces a double-stranded DNA (dsDNA) break into the nuof locus within the *E. coli* BL21-AI genome at the 20-nucleotide target sequence (N20), located 23 base-pairs (bp) downstream of the nuof +1 site. Subsequently, the λ-Red mediated homologous DNA recombination incorporates the supplied dsDNA template comprising the knock-in cassette flanked by 370 bp long homologous regions into the genome. The recombination results in insertion of sequences coding for the Twin-Strep tag and TEV protease recognition site right after the nuof transcription start site.



Peak	No. of events	Molar fraction
'Empty' ND	6563	0.31
ND-complex I	8583	0.40
Background	6289	0.29
Total events	21435	1

Figure 1 - figure supplement 2 Purification and biochemical characterization of *E. coli* complex I reconstituted into lipid nanodiscs. (A) SDS-PAGE depicting the sample solubilized in detergent after affinity chromatography (DDM) and the sample reconstituted into lipid nanodiscs after size-exclusion chromatography (ND). (B) Size-exclusion chromatography profile after complex I reconstitution into lipid nanodiscs. The main peak at 13 mL consists mostly of the intact reconstituted complex I. Fractions marked by the green, dashed-rectangle were pooled, concentrated, and used for cryo-EM and activity assays. The second peak (retention volume 15 ml and higher) contains empty nanodiscs and the dissociated cytoplasmic arm. (C,D) Mass photometry of the reconstituted complex I pooled form the main gel filtration peak. (C) The representative mass histogram, showing two main peaks: 'Empty' nanodiscs at 95 kDa and the nanodisc-reconstituted complex I at 713 kDa. (D) Molar fractions of components identified in the histogram C. (E) Representative traces of the spectrophotometric activity assays: NADH:FeCy (solid line), NADH:Q1 (dashed line), and NADH:Q1 in the presence of 20 μM Piericidin A (dotted line). The concentration of complex I was 2.5 times lower for NADH:FeCy compared to that for the NADH:Q1 assay. (F) Values of V_{max} for the three assay conditions described in panel (E). The graph shows mean ± SD, n=3. Individual measurement results are indicated as diamonds.

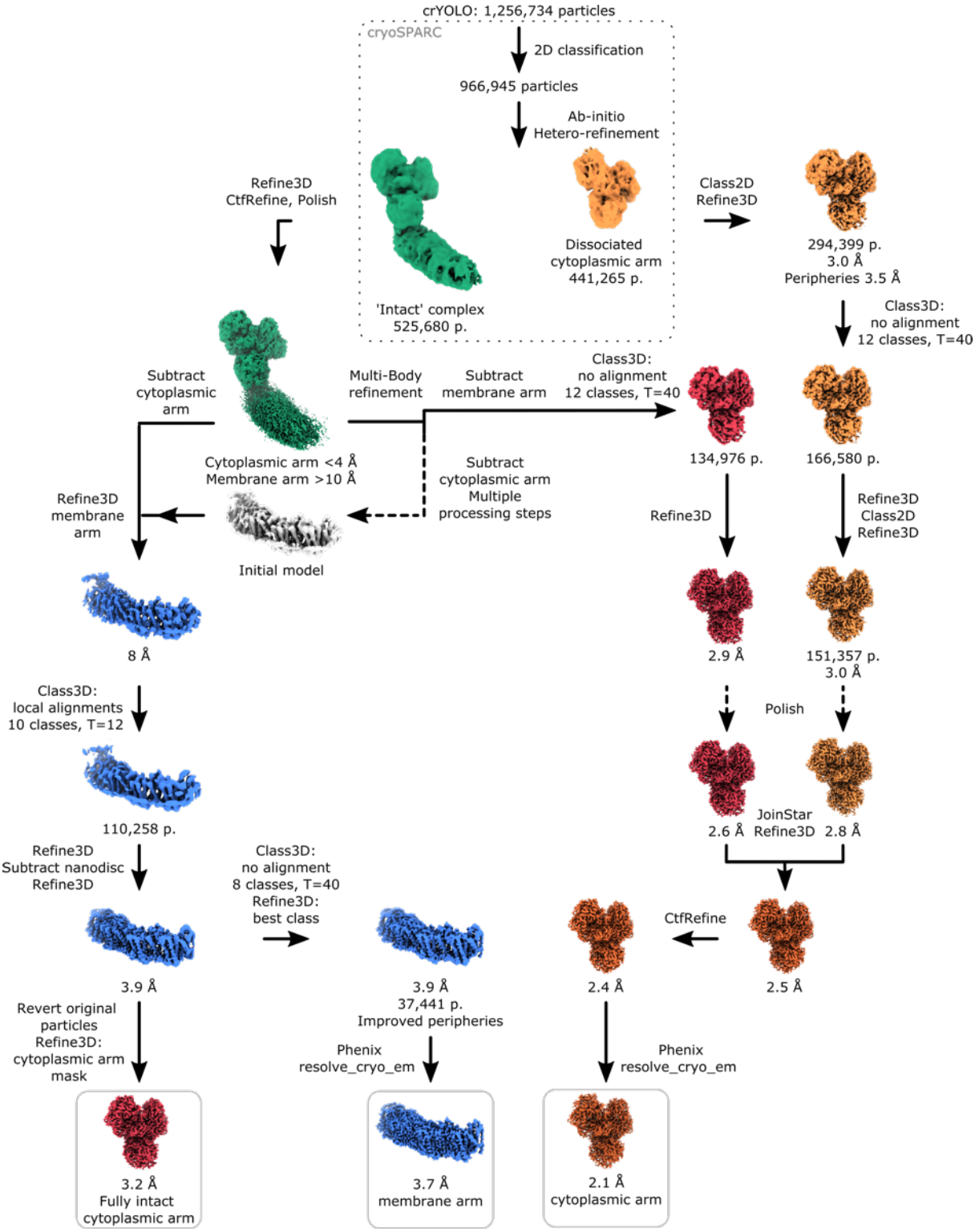


Figure 1 - figure supplement 3 Image processing diagram. The scheme indicates the principal steps during image processing that resulted in reconstructions of the peripheral and membrane arms.

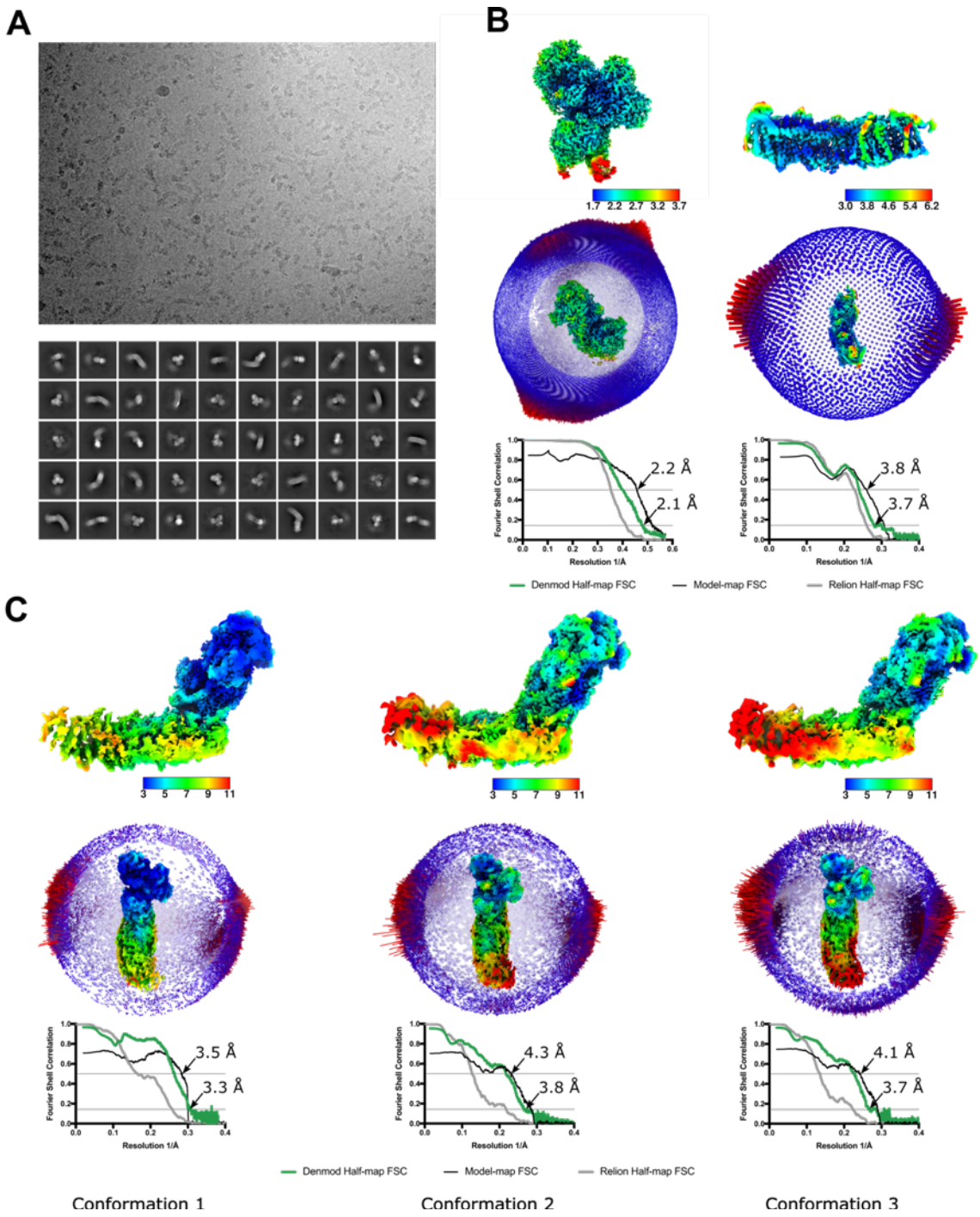
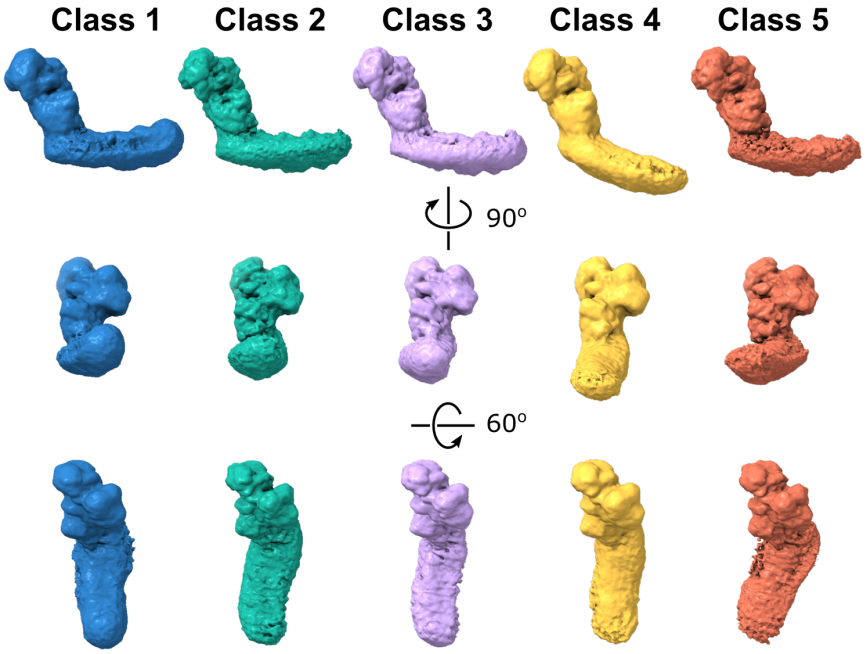


Figure 1 - figure supplement 4 Properties of the cryo-EM sample and the final reconstructions. (A) A representative micrograph (top) and 2D classes of the entire complex I (bottom). (B,C) Local resolution maps, angular distribution, and FSC plots for reconstructions of the cytoplasmic (B, left) and membrane (B, right) arms as well as for the three conformations of the intact complex I (C).

1045



1046
1047 **Figure 1 - figure supplement 5 Dynamic connection between peripheral and membrane arms.** Representative
1048 3D classes show flexibility between the arms in the *E. coli* complex. The conformations were obtained by 3D
1049 classification of all intact complex I particles aligned to the peripheral arm.
1050
1051
1052

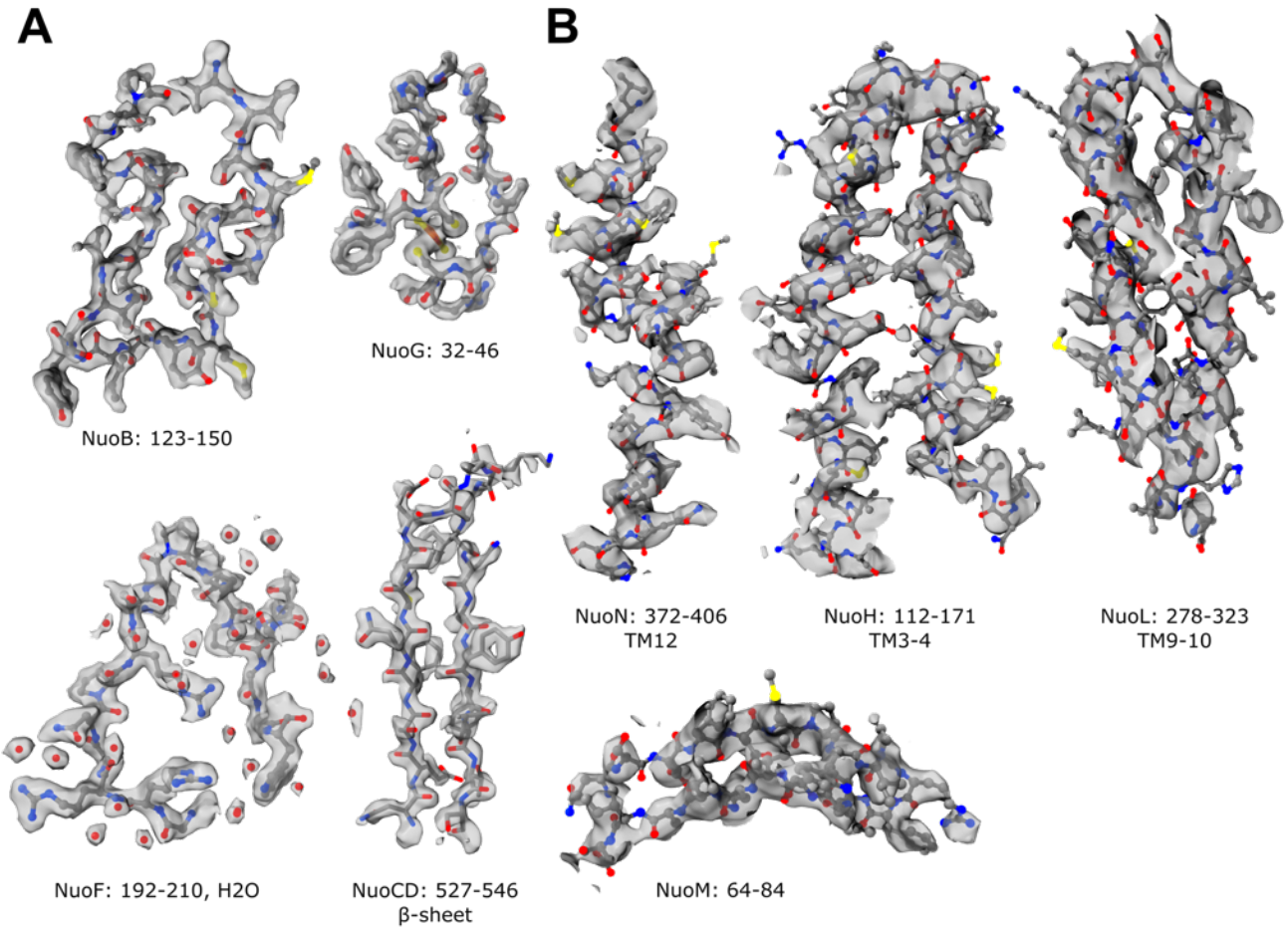


Figure 1 - figure supplement 6 Representative cryo-EM map densities. Examples of density maps for the reconstructions of the (A) cytoplasmic and (B) membrane arm.

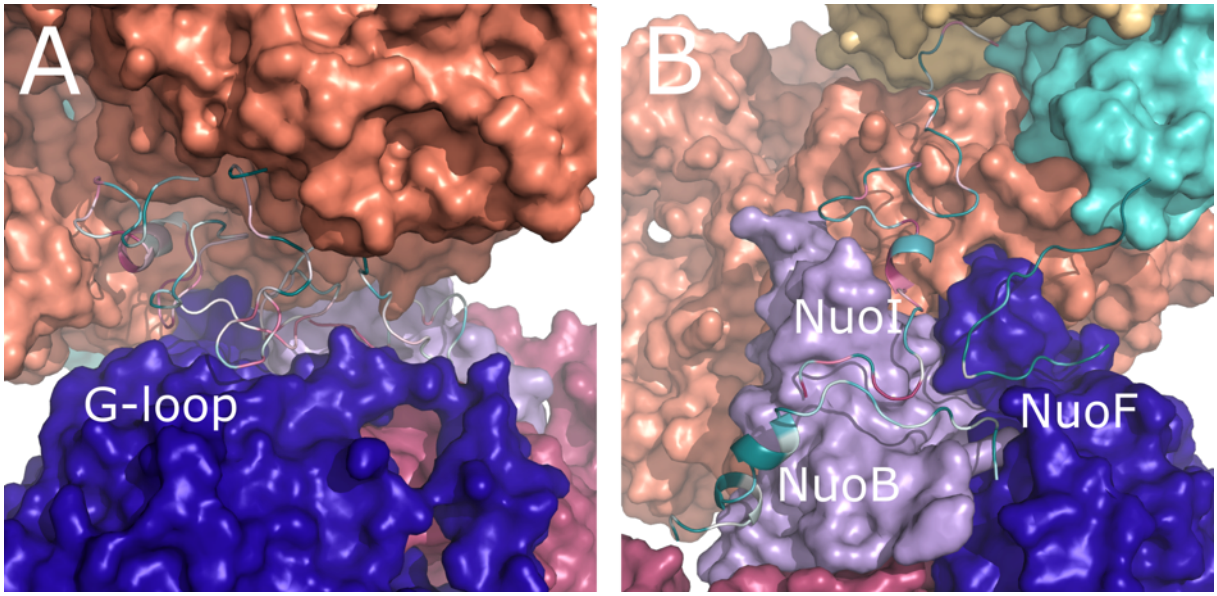


Figure 2 - figure supplement 1 Conservation of *E. coli*-specific tails in the peripheral arm subunits. (A) Insertion into subunit NuoG, the G-loop. (B) C-terminal tails in subunits NuoF, NuoI and NuoB. Color coding of the subunits is the same as in Figure 1. Conservation was calculated using ConSurf server and is color coded from green to white to purple as the degree of conservation increases.

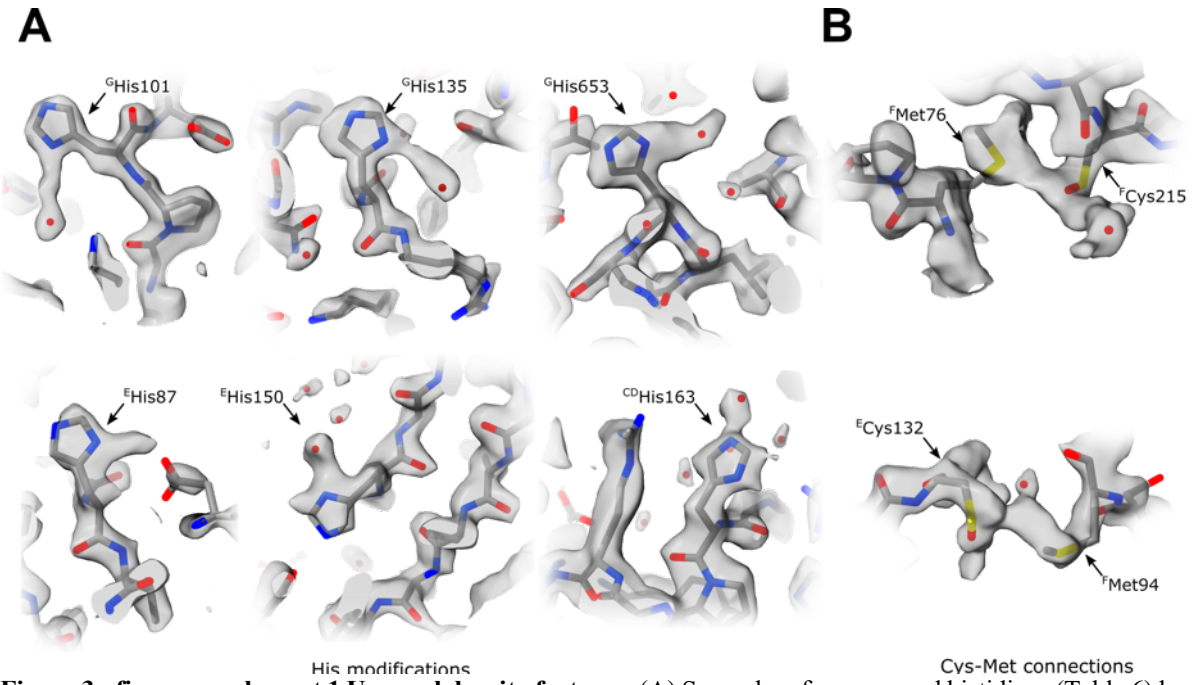


Figure 3 - figure supplement 1 Unusual density features. (A) Several surface-exposed histidines (Table 6) have extended density protruding from the imidazole ring in the plane of the ring. These extensions may represent a tightly bound heavy atom, but these features are heterogeneous (length between 2 and 4 Å) and have a very heterogeneous chemical environment; therefore, they cannot be attributed to a single type of bound atom or modification. We tentatively modelled them as water molecules. (B) In three locations, density bridging sulfur atoms of surface-exposed and closely positioned cysteine-methionine couples are bridged by density with reproducible elongated features (Cys-Met: ^{F}Met76-^{F}Cys215, ^{F}Met94-^{E}Cys132, ^{B}Met106-^{B}Cys102). We could not assign the density to any known modification or tightly bound chemical present in the protein purification buffer. Because complex I was purified without reducing agents, the corresponding Cys residues were tentatively modelled as cysteines oxidized to sulfenic acid.

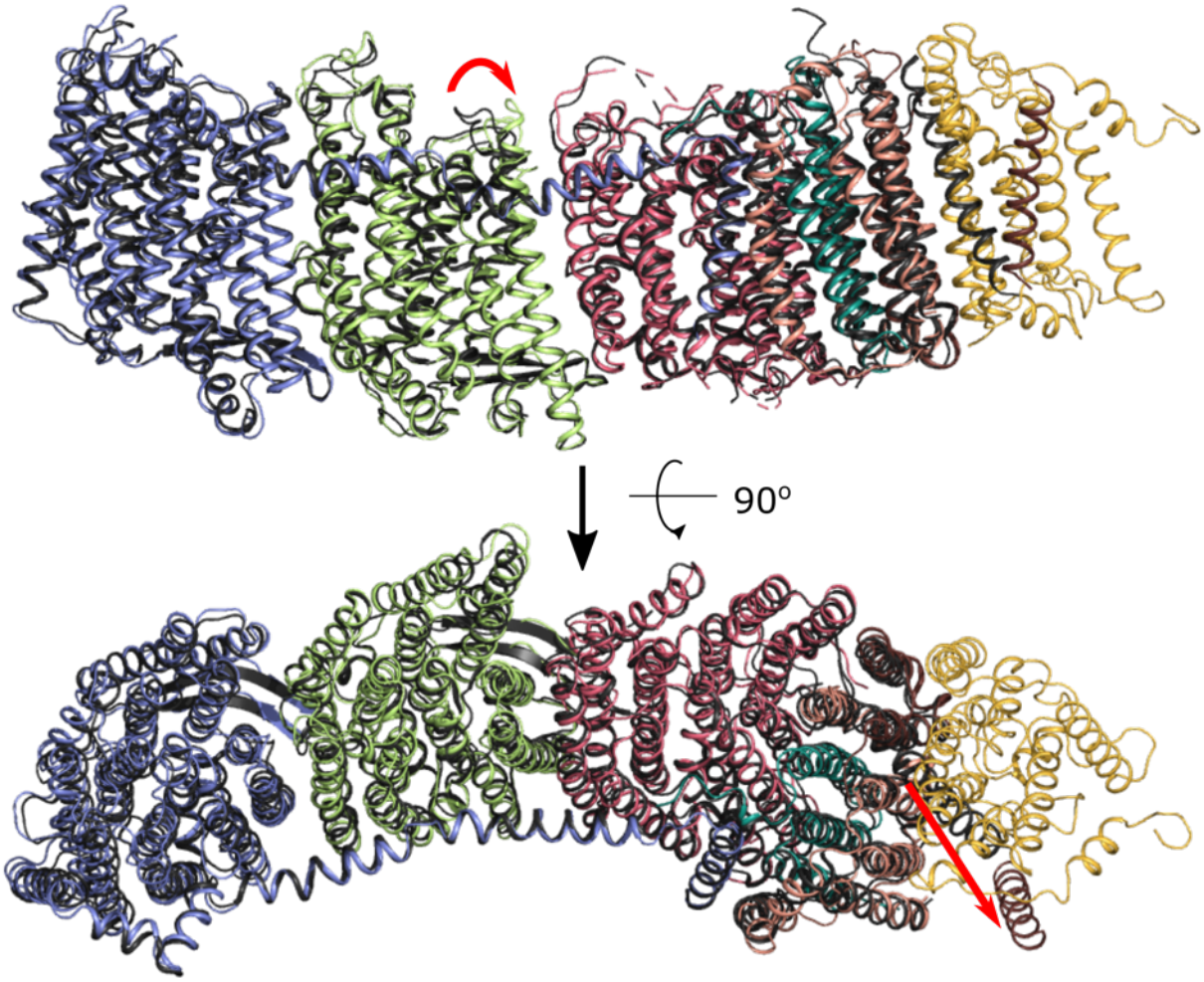


Figure 4 - figure supplement 1 Comparison of X-ray and cryo-EM structures of the membrane domain. The cryo-EM structure is colored as in Figure 1, the X-ray model is shown in black. Significant shifts in the cytoplasmic loop of NuoM and the shift of ^{A}TM1 are indicated with red arrows. Part of the β-hairpin observed in the crystal structure forms an extension of ^{M}TM6a in the cryo-EM structure. This happens despite this loop not being involved in direct crystal contact.

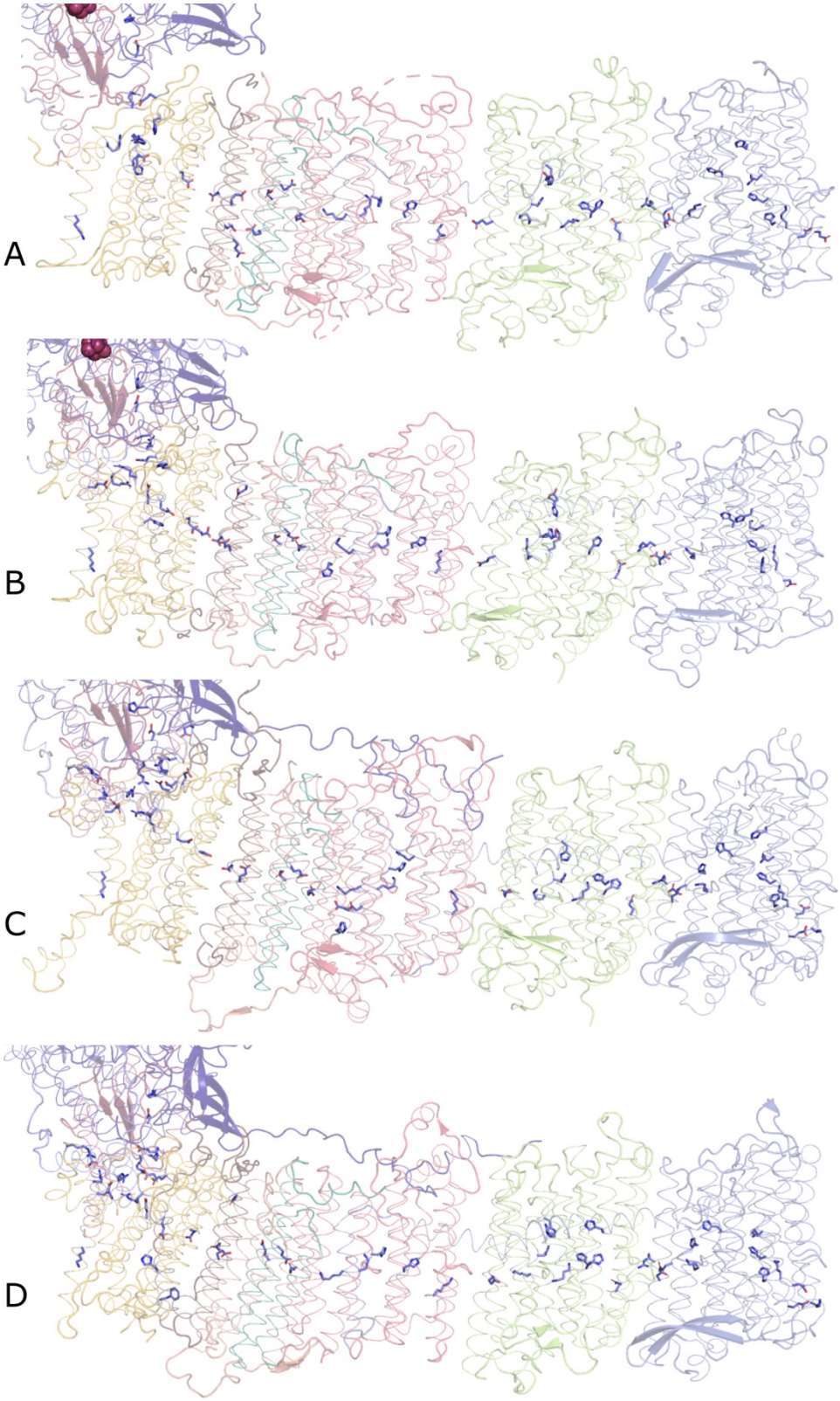


Figure 4 - figure supplement 2 Conserved chain of ionizable residues. Comparison of the chains of ionizable residues lining the Q-cavity and positioned in the hydrophobic region of the lipid membrane in (A) *E. coli*, (B) *T. thermophilus*, (C) *Yarrowia lipolytica,* and (D) *Ovis aries* complex I. In C and D, only the core subunits are shown.

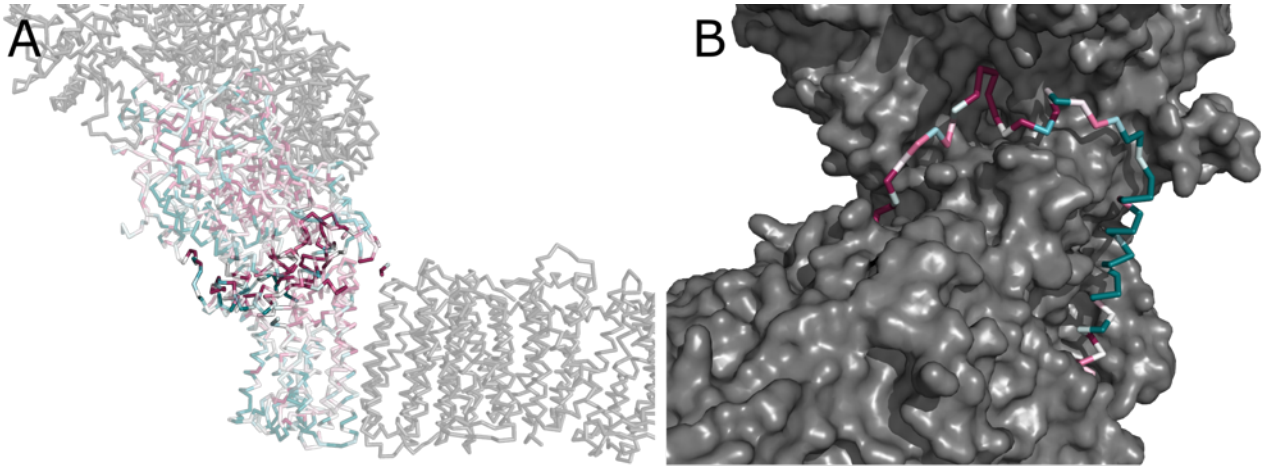


Figure 5 - figure supplement 1 Conserved interface between the arms. (A) Conservation of subunits contributing to the interface between the peripheral and membrane arms, calculated in ConSurf and color-coded from green to white to purple as the degree of conservation increases. Residues directly contributing to the interaction between arms are highlighted. (B) The conserved region of the NuoA TMH1-TMH2 loop forms a plug, which fills a crevice between the subunits NuoD, NuoB, and NuoH.

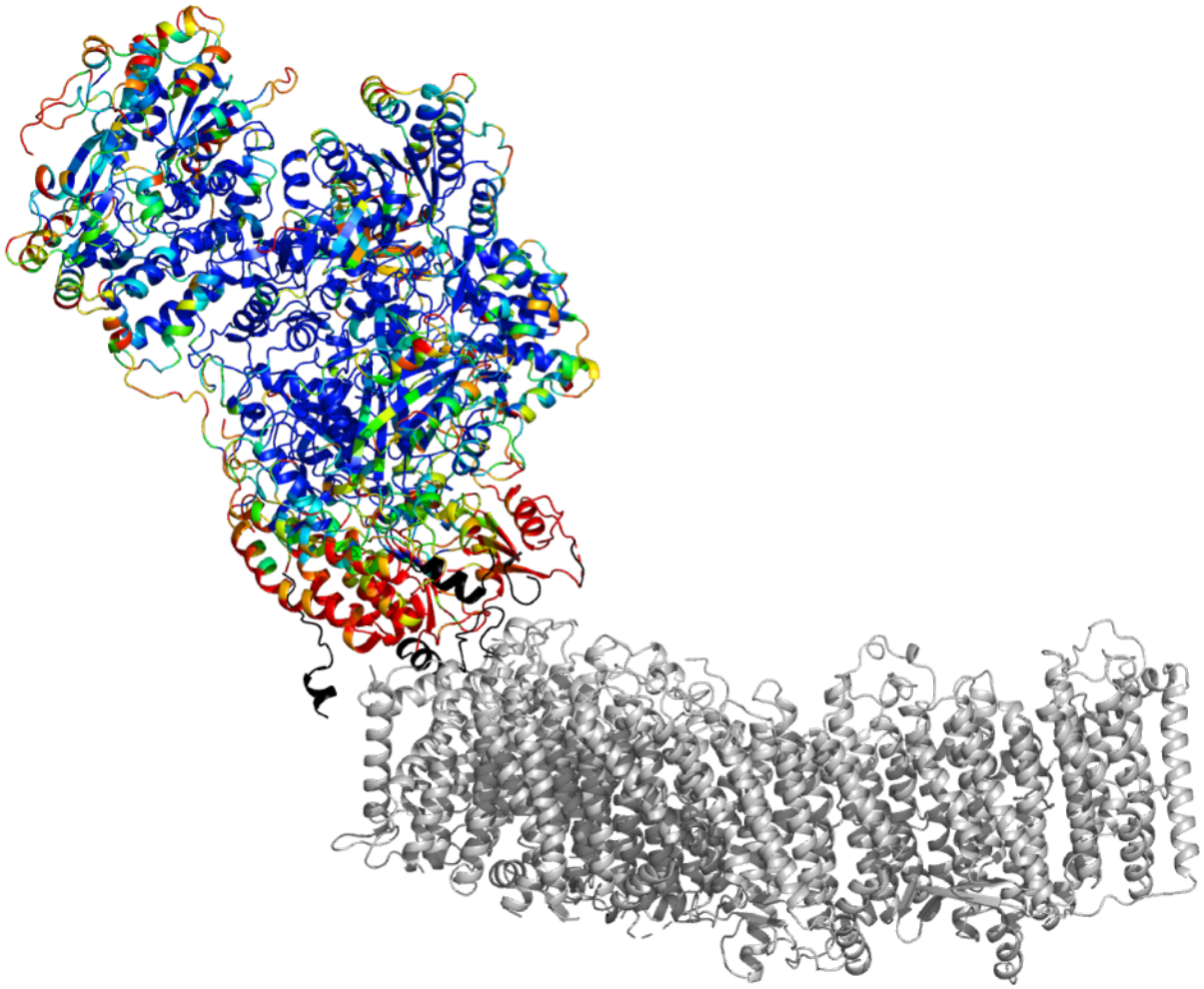


Figure 5 - figure supplement 2 B-factors of the peripheral arm show higher mobility at the arms interface. The cartoon representation is shown. The peripheral arm is colored by B-factors. The rainbow color palette scales from 5 Å^2 - blue to 100 Å^2- red. Regions completely disordered in the map of focused reconstruction of the peripheral arm but showing density in in the reconstructions of the intact complex are colored in black. The membrane arm is shown in grey for reference.

Movie 1 Composite density map of *E. coli* complex I is shown along with density of the lipid nanodisc. The homology model of TMH1is shown in ribbon representation.

Movie 2 The fragmented density of the NuoM TM8 is surrounded by well-resolved TMHs.

Table 1 Statistics of cryo-EM data collection, data processing, and model refinement

Data collection					
Microscope	JEOL CRYOARM300				
Acceleration voltage [kV]	300				
Energy filter	In-column Omega energy filter				
Energy filter slit width [eV]	20				
Magnification	60 000 x				
Detector	Gatan K3				
Physical pixel size [Å]	0.771				
Exposure time [s]	3				
Number of frames	61				
Total electron dose [$e^-/Å^2$]	65				
Defocus range [μm]	0.9 - 2.2				
Number of micrographs collected	9,122				
Total number of particles extracted	1,256,734				
Data processing					
	Entire complex				
	Conformation 1	Conformation 2	Conformation 3	Membrane domain	Cytoplasmic domain
PDB ID:	7NYR	7NYU	7NYV	7NYH	7NZ1
EMDB ID:	EMD-12653	EMD-12654	EMD-12655	EMD-12652	EMD-12661
Imposed symmetry	C1	C1	C1	C1	C1
Final number of particles	23,445	21,620	21,234	37,441	286,384
Final resolution, RELION, FSC=0.143	3.9	4.6	4.5	3.9	2.4
Final resolution, RELION, FSC=0.5	6.0	7.9	7.4	4.4	2.8
Sharpening B-factor, RELION [$Å^2$]	-67	-126	-116	-70	-52
Final resolution, PHENIX resolve_cryo_em, FSC=0.143	3.3	3.8	3.7	3.7	2.1
Final resolution, PHENIX resolve_cryo_em, FSC=0.5	3.8	4.6	4.4	4.2	2.5
Local resolution range (Å)	2.9-11	3.2-15	3.1-13	3.0-6.2	1.8-3.8
Model refinement					
Initial model	4HEA, 3RKO	4HEA, 3RKO	4HEA, 3RKO	4HEA, 3RKO	4HEA
Refinement package	PHENIX 1.18.2, Real-space refinement				
Model resolution at FSC=0.5 (Å)	3.5	4.3	4.1	3.8	2.2
Cross-correlation					
Mask	0.68	0.58	0.59	0.71	0.80
Volume	0.66	0.57	0.58	0.68	0.75
Model composition					
Non-hydrogen atoms	36075	36075	36075	16908	19830
Protein residues	4618	4618	4618	2195	2361
Waters	0	0	0	0	1170
Ligands	11	11	11	0	11
B-factors mean (A^2)					
Protein	48	72	50	84	31
Ligand	32	50	37	-	29
Waters	-	-	-	-	21
R.M.S. deviations					
Bond lengths (Å)	0.005	0.004	0.004	0.006	0.006
Bond angles (°)	0.840	0.814	0.805	0.950	0.938
Validation					
MolProbity score	1.00	0.98	0.98	1.33	0.92
Clashscore	1.20	1.19	1.19	2.41	1.36
Poor rotamers (%)	0.96	0.93	0.93	0.97	1.02
C-beta outliers (%)	0	0	0	0	0
CaBLAM outliers (%)	2.05	1.96	2.05	2.05	1.60
Ramachandran plot (%)					
Favored	97.11	97.18	97.18	95.54	97.81
Allowed	2.89	2.82	2.82	4.46	2.19
Outliers	0.00	0.00	0.00	0.00	0.00

1128 **Table 2** Residues built in the models
1129

Subunit	Total number of residues	Built residues			Co-factors	Fragments build in entire complex only
		Entire complex	Membrane arm	Cytoplasmic arm		
NuoF	445	1-441		1-441	FMN, N3	
NuoE	166	11-166		11-166	N1a	
NouG	908	1-907		1-907	N1b, N4, N7	
NuoI	180	23-180		39-180	N6a, N6b	23-38
NuoB	220	43-76, 86-179, 190-220		53-71, 90-179, 190-220	N2	43-53, 71-76, 86-90:
NuoCD	596: C 1-172 D 212-596	9-596		9-205, 210-218, 224-233, 238-596	-	206-209, 219-223, 234-237
NuoH	325	52-321	52-214, 223-321		-	215-222
NuoA	147	15-38, 60-127	15-38, 66-127		-	61-66
NuoJ	184	1-164	1-164		-	
NuoK	100	1-100	1-100		-	
NuoN	485	1-191, 199-437, 447-483	1-191, 199-437, 447-483		-	
NuoM	509	1-504	1-504		-	
NuoL	613	1-612	1-612		-	

1130
1131

1132 **Table 3** Properties of *E. coli* peripheral arm extensions (analyzed in PIZA)
1133

Subunit	Extension residues numbers	Interacts with subunit	Interaction surface, [Å^2]	Secondary structure	Specific interactions	Spatial overlap with subunits in other species
NuoF	C-term 424-445	NuoI, NuoD NuoB	254 557 118	no	Nb 8 Sb 0 Hb 0 Sb 1 Hb 0 Sb 0	Nqo15 *T. Therophilus* NUIM,NUZM,**NUMM***, *Y.Lipolytica* **Ndufs6**, Ndufs8, mouse
NouG	Insertion 687-781	NuoCD NuoI	1080 585	2 helical turns	Hb 12 Sb 5 Hb 13 Sb 4	Nqo5 *T. Therophilus* NUGM, NUYM *Y.Lipolytica* Ndufs3, Ndufs4 mouse
NuoI	C-term 139-180	NuoG NuoB NuoF NuoD NuoE	805 348 292 122 421	1 helical turn	Nb 8 Sb 5 Nb 1 Sb 0 Nb 3 Sb 0 Nb 0 Sb 1 Nb 1 Sb 0	Nqo15 *T. Therophilus* **NUMM** *Y.Lipolytica* **Ndufs6 mouse**
NuoB	C-term 196-220	NuoI NuoD NuoF	1095 216 117	2 helical turns	Nb 16 Sb 4 Nb 1 Sb 3 Ng 0 Sb 0	NUIM, N7BM *Y.Lipolytica* Ndufs8, Ndufa12 mouse

1134 *Subunits in bold are equivalent to NuoI
1135
1136
1137
1138

1139 **Table 4** Comparison of hydrogen bond networks surrounding the N1a cluster in complex I
1140 structures solved at high resolution

1141

Cluster	Hb acceptor	Hb donor		
		***E. coli* this work**	***A. aeolicus* [pdb: 6hla]**	***O. aries* [pdb: 6zk9]**
N1a	N1a S1	NH Asn136 3.5Å NH Leu134 3.1Å NδH Asn142 3.6Å	NH Ala130 3.6Å NH Leu128 3.5Å	NH Ala147 3.3Å NH Leu145 3.3Å
	N1a S2	NH CyS87 4.0Å	NH Cys91 3.5Å	NH Cys108 3.7Å OγH Thr105 2.4Å
	Sγ, Cys92 (86,103)*	NH Ser94 3.5Å	NH Ser88 3.5Å	NH Thr 105 3.7Å
	Sγ, Cys97 (91,108)	Nδ2H Asn142 4.1Å NH Asn 142 3.7Å OH W74 3.2Å	NH Val136 3.4Å	NH Met153 4.3Å
	Sγ, Cys133 (127,144)	OH W127 3.3Å NH Gly97^{F} 4.0Å	OH W794 3.2Å N GLy99^{F} 3.7Å	OH W649 3.2Å NH GLy103^{F} 4.5Å
		NH Gly135 3.3Å	N Gly129 3.5Å	NH Gly146 3.3Å
	Sγ, Cys137 (131,148)	N GLy97F 3.4Å	N Gly99F 3.3Å	N Gly103 3.2Å

1142 *Numbering in parenthesis is given for *A. aeolicus* and *O. aries*, respectively
1143
1144

Table 5 Differences in the hydrogen bond network of iron-sulfur clusters in complex I structures solved at high resolution and water molecules in the immediate cluster environment. (Only the clusters for which such comparison could have been done and clusters displaying differences in the environment are listed)

Cluster, Subunit	**Organism**		
	E. coli **this work**	***A. aeolicus*** **[pdb: 6hla]**	***O. aries*** **[pdb: 6zk9]**
N3, NuoF	His400 (+)	Leu395(-)	Leu407(-)
	Trp363(-)	Glu349(+)	Gln361(+)
	Asn196(+)	His198(-)	Lys202(-)
N1b, NuoG	HOH386		HOH1070
N7, NuoG	HOH441		
	HOH577		
	Cys228		Asp229
	Cys231		Asp232
	Cys235		Ser236
	Cys263		Ser264
			HOH929*
			HOH933*
			HOH1005*
N4, NuoG	Thr203(+)		Val205(-)
N5, NuoG	conserved		
N6a, NuoI	Phe92(-)		Tyr109(+)
			HOH539
N6b, NuoI	Leu48(-)		His65(+)
	Cys74(+)		Ala91(-)
	Leu116(-)		Glu133(+)
N2, NouB	HOH438		HOH353
	HOH211		HOH579
	$Arg250^{D}$		Arg85 dimethylated
	Ser62		Ala53

* Water molecules replacing the N7 cluster

1194 **Table 6** Histidine residues with unassigned features extending from the imidazole ring
1195

Residue	Modelled atom	Comments
^{E}His87	None	Interaction with ^{E}D146
^{E}His150	HOH	Distance 2.1 Å
^{E}His152	HOH	Distance 2.4-2.6 Å
^{G}His5	HOH	Density on both sides
^{G}His101	HOH	Positive environment
^{G}His123	HOH	Distance 2.2Å
^{G}His427	HOH	Distance 2.85 Å
^{G}His653	HOH	Distance 2.5 Å very strong
^{CD}His163	HOH	Distance 2.5 Å
^{CD}His507	HOH	Distance 4.2 Å

1196

1197