Structural and mechanistic study of bacterial complex I and its cyanobacterial ortholog

by

Vladyslav Kravchuk

March, 2023

A thesis submitted to the
Graduate School
of the
Institute of Science and Technology Austria
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

Committee in charge:
Prof. Ryuichi Shigemoto, Chair
Prof. Leonid Sazanov
Prof. Florian Schur
Prof. Clemens Plaschka

The thesis of Vladyslav Kravchuk, titled "Structural and mechanistic study of bacterial complex I and its cyanobacterial ortholog", is approved by:

Supervisor: Leonid Sazanov, IST Austria, Klosterneuburg, Austria

Signature: _____

Committee Member: Florian Schur, IST Austria, Klosterneuburg, Austria

Signature: _____

Committee Member: Clemens Plaschka, Research Institute of Molecular Pathology, Vienna, Austria

Signature: _____

Defense Chair: Ryuichi Shigemoto, IST Austria, Klosterneuburg, Austria

Signature: _____

signed page is on file

ISTA Thesis, ISSN: 2663-337X

ISBN: 978-3-99078-029-9

I hereby declare that this thesis is my own work and that it does not contain other people's work without this being so stated; this thesis does not contain my previous work without this being stated, and the bibliography contains all the literature that I used in writing the dissertation.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee, and that this thesis has not been submitted for a higher degree to any other university or institution.

I certify that any republication of materials presented in this thesis has been approved by the relevant publishers and co-authors.

Signature: _____

Vladyslav Kravchuk

March, 2023

signed page is on file

Abstract

Most energy in humans is produced in form of ATP by the mitochondrial respiratory chain consisting of several protein assemblies embedded into lipid membrane (complexes I-V). Complex I is the first and the largest enzyme of the respiratory chain which is essential for energy production. It couples the transfer of two electrons from NADH to ubiquinone with proton translocation across bacterial or inner mitochondrial membrane. The coupling mechanism between electron transfer and proton translocation is one of the biggest enigma in bioenergetics and structural biology. Even though the enzyme has been studied for decades, only recent technological advances in cryo-EM allowed its extensive structural investigation.

Complex I from *E.coli* appears to be of special importance because it is a perfect model system with a rich mutant library, however the structure of the entire complex was unknown. In this thesis I have resolved structures of the minimal complex I version from *E. coli* in different states including reduced, inhibited, under reaction turnover and several others. Extensive structural analyses of these structures and comparison to structures from other species allowed to derive general features of conformational dynamics and propose a universal coupling mechanism. The mechanism is straightforward, robust and consistent with decades of experimental data available for complex I from different species.

Cyanobacterial NDH (cyanobacterial complex I) is a part of broad complex I superfamily and was studied as well in this thesis. It plays an important role in cyclic electron transfer (CET), during which electrons are cycled within PSI through ferredoxin and plastoquinone to generate proton gradient without NADPH production. Here, I solved structure of NDH and revealed additional state, which was not observed before. The novel “resting” state allowed to propose the mechanism of CET regulation. Moreover, conformational dynamics of NDH resembles one in complex I which suggest more broad universality of the proposed coupling mechanism.

In summary, results presented here helped to interpret decades of experimental data for complex I and contributed to fundamental mechanistic understanding of protein function.

Acknowledgments

I'm grateful to Prof Leonid Sazanov FRS for his excellent supervision and guidance during my PhD. I've learnt from him a lot and not only about structural biology and bioenergetics but also about life. Go big or go home – biggest takeaway for my future pathway.

I would also like to thank my colleagues from IST Austria, namely Kristina Lukić and Julia Steiner who helped me a lot in the lab especially at the beginning. Other members of the lab: Gergely Pinke, Irene Vercellino and Domen Kampjut for fun times and support during my stay at IST Austria. We became good friends and I'm happy to have you guys in my life.

I acknowledge IST scientific support facilities, especially EMF, for their excellent job in training and assistance. Special thanks to Dr Valentin Hodirnau who helped me to collect most of my data and Dr Alexej Charnagalov for his expert help with protein purification. I am a recipient of a DOC Fellowship of the Austrian Academy of Sciences, titled "Structural characterization of *E. coli* complex I: an important mechanistic model" under the grant number 25541. This work was also partially funded by the ERC Advanced Grant 101020697 RESPICHAIN.

I'm thankful to my first biology teacher Lydmila Oleksiivna for instilling love for life science and many other teachers and research supervisors I had along the way.

Last but not least, I'm grateful to my family and parents for the immense support in my life. Without them none of this would be possible.

About the Author

Vladyslav Kravchuk completed a BSc in Biotechnology at the National Aviation Univesity in Ukraine. During his studies he worked with Professor Alexander Kornelyuk on molecular dynamics and mechanism of mammalian cytoplasmic tyrosyl-tRNA synthetase. Vladyslav joined IST Austria in September 2017 to work with Professor Leonid Sazanov on the molecular structure and function of molecular proton pumps involved in energy production and to learn cryogenic electron microscopy. His results build the basis for a first-author publication in the journal Nature and presented at several conferences.

List of Publications Appearing in Thesis

Chapter 3, as well as parts of chapter 2 have been published here:

Kravchuk V., Petrova O., Kampjut D., Wojciechowska-Bason A., Breese Z. & Sazanov L. A universal coupling mechanism of respiratory complex I. *Nature* **609**, 808–814 (2022). https://doi.org/10.1038/s41586-022-05199-7

Table of Contents

List of figures

List of tables

List of abbreviations

Å	Ångstrom
ATP	Adenosine triphosphate
AU	Absorption units
CHAPS	3-[(3-cholamidopropyl)-dimethylammonio]-1-propane sulphonate
Cryo-EM	electron cryo-microscopy
CTF	Contrast transfer function
DDM	n-dodecyl-β-maltoside
DQ	Decylubiquinone
ETC	Electron transport chain
FeCN	Potassium hexacyanoferrate (III) K3[Fe(CN)6]
FMN	Flavin mononucleotide
FSC	Fourier shell correlation
IMS	Intermembrane space
kDa	Kilodalton
LMNG	Lauryl maltoside neopentylglycol
MA	Membrane arm
MES	2-(N-Morpholino)ethanesulfonic acid
NAD+	Nicotinamide adenine dinucleotide, oxidised
NADH	Nicotinamide adenine dinucleotide, reduced
NADP+	Nicotinamide adenine dinucleotide phosphate, oxidised
NADPH	Nicotinamide adenine dinucleotide phosphate, reduced
PA	Peripheral arm
PC	Phosphatidylcholine
PDB	Protein data bank
PE	Phosphatidylethanolamine
Q	Quinone
QH_2	Quinol
ROS	Reactive oxygen species

1 Introduction

1.1 Electron transport chain in different organisms

Adenosine triphosphate (ATP) is essential for cells to drive numerous processes during their lifecycle. ATP is primarily synthesized in the mitochondrial matrix by oxidative phosphorylation (OXPHOS), or cellular respiration. Electrons harvested mostly from the catabolic processes of tricarboxylic acid cycle (TCA), fatty acid oxidation and glycolysis enter the electron transport chain (ETC) on the inner mitochondrial membranes. According to the Chemiosmotic theory (postulated by Peter Mitchell in 1961) the free-energy released as the result of reactions performed in respiratory complexes is used for proton translocation from the mitochondrial matrix to the intermembrane space, in order to create a transmembrane difference in electrochemical potential (MITCHELL, 1961). The electrochemical potential is then used to drive ATP synthase for ATP synthesis.

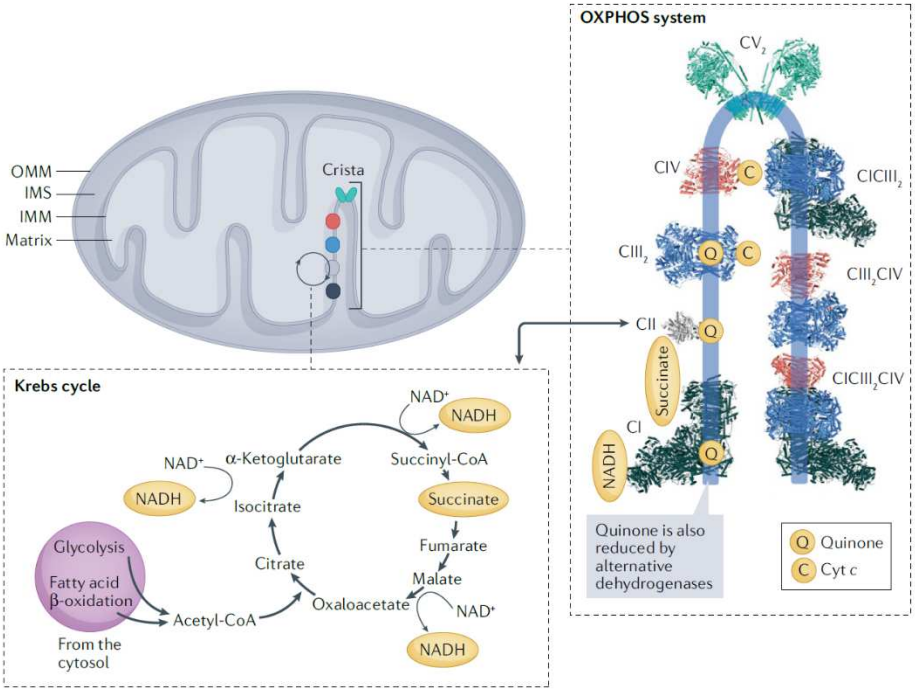

Figure 1.1 The mammalian mitochondrial electron transport chain.

Schematic representation of a mitochondrion and OXPHOX complexes. NADH, succinate, quinone (Q) and cytochrome c (C) are depicted as yellow ovals. NADH donates electrons to complex I (CI) and succinate donates electrons to complex II (CII), while quinone shuttles electrons from complexes I and II to complex III2 (CIII2) and cytochrome c (cyt c) shuttles electrons from complex III2 to complex IV (CIV). The OXPHOS components shown as shapes in the scheme on the left are shown as models in matching colours, together with their higher-order organization into supercomplexes: complex I (Protein Data Bank (PDB) ID 6ZKC) is coloured dark slate grey, complex II (PDB ID 1ZOY) is light grey, complex III2 (PDB ID 6Q9E) is blue, complex IV (PDB ID 5IY5)

is salmon, supercomplexes CICIII2CIV (PDB ID 5J4Z), CICIII2 (PDB ID 6QBX) and CIII2CIV (PDB ID 7O3C) are coloured according to the individual complexes and ATP synthase dimer (CV2; PDB ID 7AJF) is cyan. ATP synthase is found at the cristae edges, where its dimers bend and shape the membrane. Complexes I, II, III2 and IV and their supercomplexes are found in the flat areas of cristae. Figure adapted from (Vercellino & Sazanov, 2022).

Mitochondrial ETC consists of four protein complexes: NADH:ubiquinone oxidoreductase (Complex I), succinate dehydrogenase (Complex II), cytochrome bc1 complex (Complex III) and cytochrome c oxidase (Complex IV). The electron transfer through the ETC is coupled to the proton translocation out of the mitochondrial matrix by the proton pumps, Complex I, III and IV, while complex II does not pump protons. The complexes create an electrochemical gradient across the membrane (proton motive force, pmf) which drives the ATP synthase (complex V) (Figure 1.1 and 1.2a) (Sazanov, 2015; Vercellino & Sazanov, 2022). Electron transport between the complexes is mediated by the membrane-embedded ubiquinone (Q) and the soluble cytochrome c. Complex I in mitochondria and bacteria is the entry point for electrons from NADH, which are used to reduce Q to ubiquinol (QH2). QH2 is further used by complex III to reduce cytochrome c in the intermembrane space (IMS); complex IV uses cytochrome c to reduce molecular oxygen, which is the ultimate electron acceptor. For every electron pair transferred from NADH to O_2, 10 protons are transported out of the matrix to the IMS (Nicholls & Ferguson, 2002). Complex II provides additional electrons into the chain by catalysing the oxidation of succinate to fumarate with the reduction of ubiquinone to ubiquinol.

The respiratory chain of plant and fungal mitochondria is similar to the mammalian one and consist from the four core complexes I-IV (Figure 1.2b). Besides, plants' ETC also contains an alternative oxidase (AOX), which reduces oxygen to water directly from quinol oxidation without proton translocation, as well as dehydrogenases, that can directly transfer electrons to ubiquinone (Schertl & Braun, 2014; Millar *et al*, 2011). Some fungi like *Saccharomyces cerevisiae* lack complex I, which is compensated by presence of alternative NAD(P)H dehydrogenase (Nosek & Fukuhara, 1994).

Chloroplasts contain unique complexes such as photosystems which are essential part of its ETC (Figure 1.2d). Photosystem II (PSII) is present in membranes of photosynthetic organisms and oxidizes water to protons and oxygen. Then electrons from PSII are transferred through the quinone/quinol pool to cytochrome b6f complex and then to photosystem I (PSI).

PSI in turn reduces ferredoxin (Fd) which is a substrate for complex I homologue NDH-1 (Rochaix, 2011).

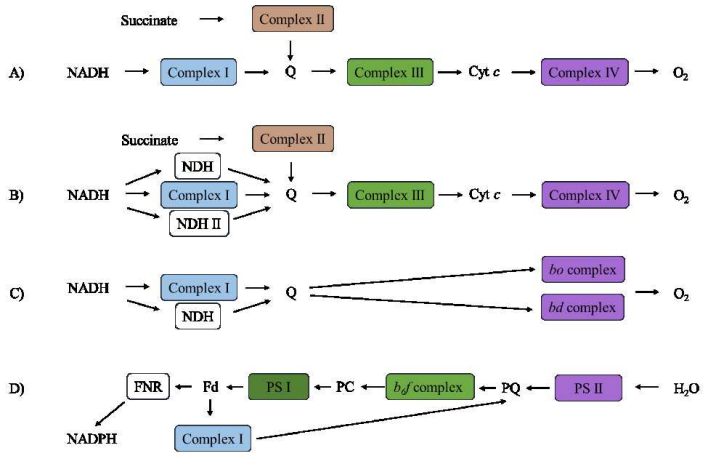

Figure 1.2 Schematic representations of the ETC in different organisms.

A) mammal mitochondria B) plant and fungal mitochondria C) E. coli cytoplasmic membrane D) plant chloroplast and cyanobacterial membrane. Figure adapted from (Friedrich *et al*, 1995).

In fact, the core physical principles of ETC are strongly conserved across all domains of life. However, bacteria, in contrast to eukaryotic mitochondria, have branched respiratory chains, which can use different routes of electron transfer depending on the growth conditions (Kaila & Wikström, 2021). An example of bacterial (*E. coli*) ETC is shown on Figure 1.2c and 1.3.

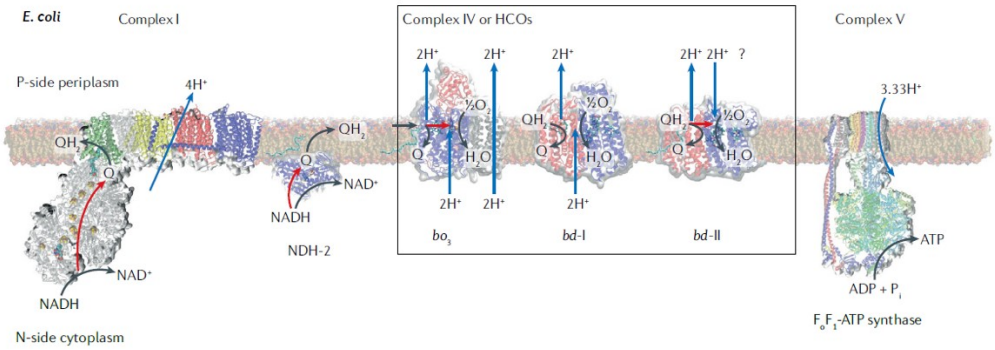

Figure 1.3 E.coli electron transport chain

The reported proton pumping is given for two-electron reduction steps. HCO, haem–copper oxidase, NDH-2, type II NADH:ubiquinone oxidoreductase; N-side, negatively charged cytoplasmic or stromal side; Pi, inorganic phosphate; P-side, positively charged periplasmic or luminal side; Q, quinone; QH2, quinol; black arrows, chemical reactions; blue arrows, protons; red arrows, electrons. Figure adapted from (Kaila & Wikström, 2021)

1.2 Complex I

1.2.1 Catalytic reaction and the core architecture of Complex I

Complex I is the first and the largest enzyme of the mitochondrial and bacterial respiratory chain. It catalyses NADH oxidation and quinone reduction coupled to the proton translocation across the mitochondrial or bacterial inner membrane. For the one NADH molecule oxidized, one molecule of quinone (Q) is reduced and four protons are translocated. Complex I itself accounts to nearly 40% of pmf generated (Sazanov, 2015). Complex I consist of two major arms: peripheral and membrane embedded. Peripheral arm (PA) catalyses electron transfer from NAHD to Q. It harbours eight conserved iron-sulphur (Fe-S) clusters and a non-covalently bound electron acceptor flavin mononucleotide (FMN). FMN accepts electrons form NADH which then tunnel through the iron-sulphur clusters to the "Q-cavity" where a Q molecule is reduced to QH2 (Verkhovskaya *et al*, 2008). Complex I has different subunit nomenclature for different species. Here the nomenclature from *E. coli* will be used (Figure 1.4).

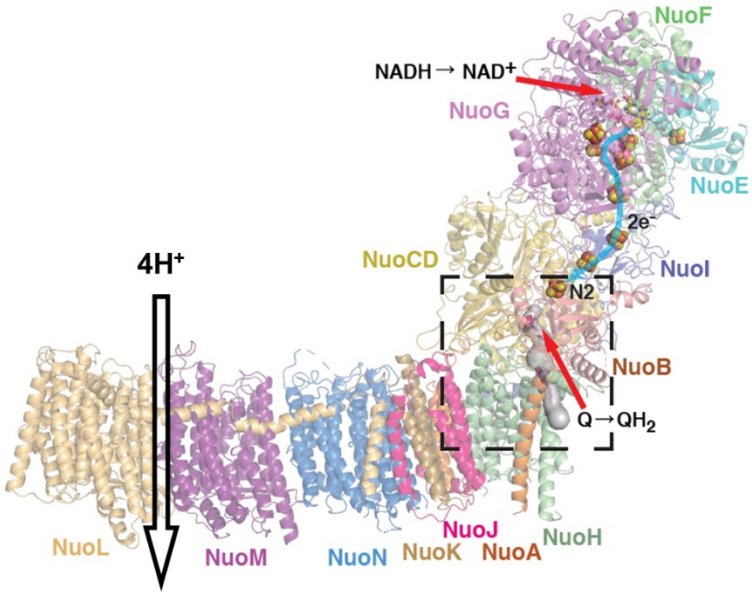

Figure 1.4 Overall structure and reaction centres of E.coli complex I

E. coli complex I consist of 13 core subunits, which are coloured in different colours. Grey surface shows Q-cavity. 4 protons are pumped across the membrane as a result of Q reduction and NADH oxidation. Figure adapted from (Kravchuk et al, 2022)

The core part of the peripheral arm contains 7 (6 in *E. coli*) subunits: NuoF, E, G, CD (genes C and D are fused), B, I. All the known cofactors (FMN and Fe-S clusters) are located within the PA. Seven Fe-S clusters are part of the main redox chain (Ohnishi, 1998) while the off-path non-conserved cluster N7 is believed to have solely a structural role i.e. maintaining complex stability (Birrell *et al*, 2013; Dorner *et al*, 2017). Another off-path cluster, N1a, is fully conserved and is likely to have a functional role, which is still not clear, one possibility being the reduction of reacting oxygen species (ROS) production (Sazanov & Hinchliffe, 2006). Substrates and cofactors of complex I have low reduction potentials, which lead to non-specific electron transfers like the reduction of O_2 to superoxide or H_2O_2. This makes complex I a source of ROS which cause DNA damage and may be related to aging (Hirst, 2013; Esterhazy *et al*, 2008).

The overall electron transfer pathway is the following: NADH→FMN→N3→N1b→N4→N5→N6a→ N6b→N2→Q (Figure 3.2c). The time of electron transfer from NADH to N2 is about 100 μs while the full catalytic cycle takes about 5 ms. Therefore, the overall electron transfer is thought to be limited by Q binding and release dynamics (Verkhovskaya *et al*, 2008). N2 has the highest potential (-100 to -150 mV) which is typical for a terminal cluster in a redox chain. It is believed that electron transfer from N2 to Q initiates a cascade of conformational and electrostatic changes propagating through the membrane arm.

The process of electron transfer is coupled to proton translocation across the membrane. Membrane antiporter-like subunits of the complex perform proton pumping (Efremov & Sazanov, 2011) (Figure 1.4). Membrane arm (MA) consists of 7 subunits, the three largest ones (NuoL, M and L) are antiporter-like subunits, homologous to each other and to cation/H^+ antiporters (Efremov & Sazanov, 2011). The three antiporter-like subunits are linked by a long traverse α-helix, which is part of NuoL and may be a coupling element of the complex. Each of the antiporter-like subunits contain two symmetry-related half-channels, one open to the matrix arm and the other to the intermembrane space (IMS). A series of protonatable residues in the middle of the membrane connects the half channels, perhaps enabling H^+ translocation. Subunits NuoA, J and K form a connection between the antiporter-like subunits and NuoH. Subunit NuoH is a unique complex I subunit and has no homologs in other enzymes. It forms most of the Q-binding cavity and links the membrane arm with the peripheral (matrix) arm.

The architecture of the membrane domain seems to suggest the translocation of one proton through each antiporter-like subunit and one proton through so-called E-channel (formed mainly by Glu residues of NuoA, H, J and K). Thus four protons are transferred per

NADH molecule oxidized. The mechanism of proton transfer is believed to involve conformational and electrostatic changes but specific details remain unclear (Kampjut & Sazanov, 2020).

1.2.2 Evolution and assembly of complex I

Complex I belongs to a family of membrane-bound oxidoreductases that is related to a class of membrane-bound [NiFe] hydrogenases. [NiFe] hydrogenases couple substrate oxidation and hydrogen reduction to active proton transport (Hedderich, 2004). Subunits of complex I share homology to related proteins, which suggests their origin from unification of other subcomplexes (modules) with distinct functions as the modules trace back to two unrelated protein families (Moparthi & Hägerhäll, 2011; Friedrich & Scheide, 2000).

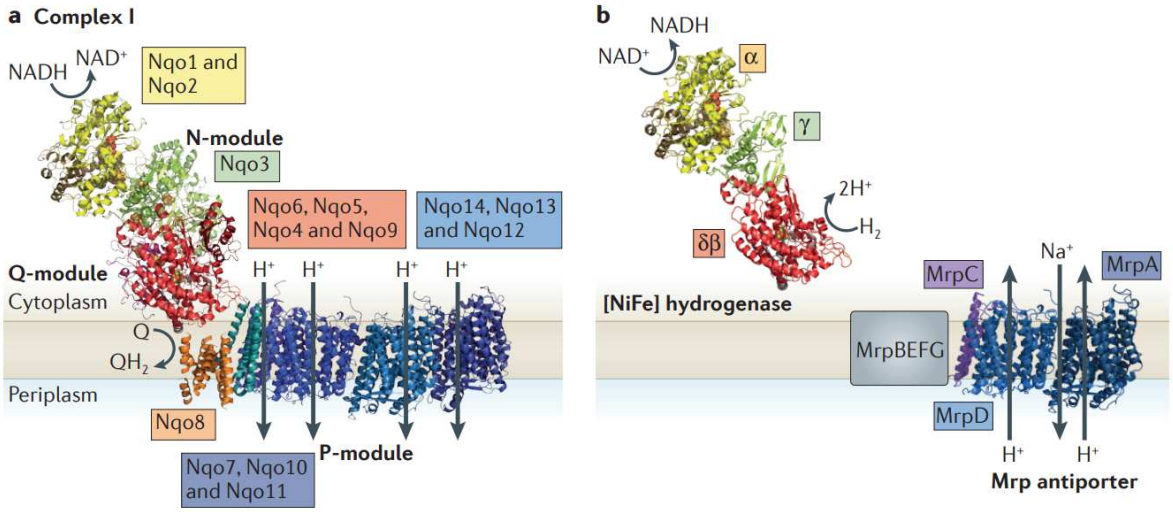

Figure 1.5 Evolution of complex I.

a. Complex I is made of N, Q and P modules. NuoH (homologue of ND1) was added later and does not have a clear evolutionary history. b. Mrp antiporter and [NiFe] hydrogenase are the evolutionary precursors of the P-module and N/Q-module respectively. Figure adapted from (Sazanov, 2015)

Peripheral arm (where NADH-Q oxido-reduction takes place) of complex I is originated from soluble [NiFe] hydrogenases, while proton translocating part was provided by Mrp cation/H^+ antiporters (Figure 1.5). The combination of soluble hydrogenase and antiporter was likely to have resulted in the emergence of several known types of membrane-bound hydrogenases, which later evolved into complex I. The figure above shows the structure of the entire Thermus thermophilus complex I (PDB ID 3M9S) and the evolutionary modules together with homology-based architectures of the NAD+-reducing [NiFe] hydrogenase and the Mrp antiporter.

Interestingly, the assembly of complex I is also modular as involves combination of pre-assembled modules. Since complex I includes many cofactors, its subunits have to be assembled together in a coordinated way for proper holoenzyme activity. However, there is not much known regarding the assembly of the bacterial complex I comparing to the mitochondrial one which is larger (Friedrich *et al*, 2016; Guerrero-Castillo *et al*, 2017). Tentative assembly scheme is shown on Figure 1.6.

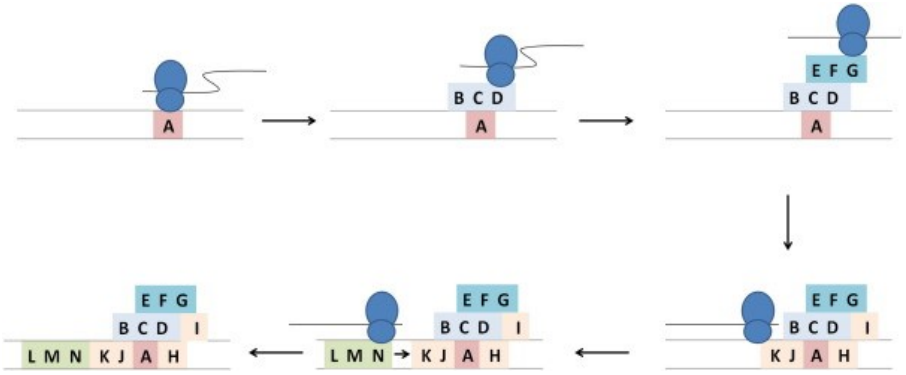

Figure 1.6 Scheme of assembly pathway of bacterial complex I

The ribosome (not scaled) is shown in blue; the mRNA is represented by the black line. Figure adapted from (Friedrich *et al*, 2016)

1.2.3 Complex I in different organisms

Complex I has characteristic L shape in all domains of life, however, its mass differs in different organisms depending on presence and absence of supplementary subunits.

Module	*Escherichia coli*	*Thermo-synechococcus elongatus*	*Thermus thermophilus*	*Yarrowia lipolytica*	*Bos taurus* (Bovine)	*Homo sapiens*	
Peripheral arm							Cofactors[1]
Dehydrogenase (N)	**NuoF**	-	**Nqo1**	**NUBM**	**51 kDa**	**NDUFV1**	FMN N3 (4Fe[51])
	NuoE	-	**Nqo2**	**NUHM**	**24 kDa**	**NDUFV2**	N1a (2Fe[24])
	NuoG	-	**Nqo3**	**NUAM**	**75 kDa**	**NDUFS1**	N1b (2Fe[75]) N4 (4Fe[75]C) N5 (4Fe[75]H) (N7)[2]
Connecting (Q)	**NuoD (NuoCD)**[3]	**NdhH**	**Nqo4**	**NUCM**	**49 kDa**	**NDUFS2**	No cofactor
	NuoC[3]	**NdhJ**	**Nqo5**	**NUGM**	**30 kDa**	**NDUFS3**	No cofactor
	NuoI	**NdhI**	**Nqo9**	**NUIM**	**TYKY**	**NDUFS8**	N6a (4Fe[TY]1) N6b (4Fe[TY]2)
	NuoB	**NdhK**	**Nqo6**	**NUKM**	**PSST**	**NDUFS7**	N2 (4Fe[PS])
Membrane arm							TMH[4]
-	**NuoH**	**NdhA**	**Nqo8**	**NU1M**	**ND1**	**ND1**	8-9
Pumping (P)	**NuoA**	**NdhC**	**Nqo7**	**NU3M**	**ND3**	**ND3**	3
	NuoJ	**NdhG**	**Nqo10**	**NU6M**	**ND6**	**ND6**	5
	NuoK	**NdhE**	**Nqo11**	**NULM**	**ND4L**	**ND4L**	3
	NuoN	**NdhB**	**Nqo14**	**NU2M**	**ND2**	**ND2**	11-14
	NuoM	**NdhD**	**Nqo13**	**NU4M**	**ND4**	**ND4**	14
	NuoL	**NdhF**	**Nqo12**	**NU5M**	**ND5**	**ND5**	16-17

Figure 1.7 Complex I subunits nomenclature in different organisms

E. coli complex I is the minimal complex I version which consists only from 13 (NuoCD is fused) core subunits, with a total mass of around 550 kDa. In contrast, mammalian enzyme acquired additional 31 subunits and the yeast one additional 27 subunits, making their mass up around 1 MDa (Figure 1.7) (Letts & Sazanov, 2015). Although supplementary subunits do not play role in the catalytic reaction, they are important for the complex stability, assembly and functioning (Stroud *et al*, 2016). Moreover, mutations in complex I supplementary subunits lead to human diseases, signifying their importance (Fiedorczuk & Sazanov, 2018). Structural conservation of the core part of the complex suggests conservation of the coupling mechanism.

1.2.4 Cyanobacterial NDH complex

Cyanobacterial NDH (cyanobacterial complex I) is a part of broad complex I superfamily. It plays an important role in cyclic electron transfer (CET), during which electrons are cycled within PSI through ferredoxin and plastoquinone to generate proton gradient without NADPH production. Balance between NADPH and ATP is essential for the Calvin-Besnon cycle and varies under physiological conditions (Kramer & Evans, 2011).

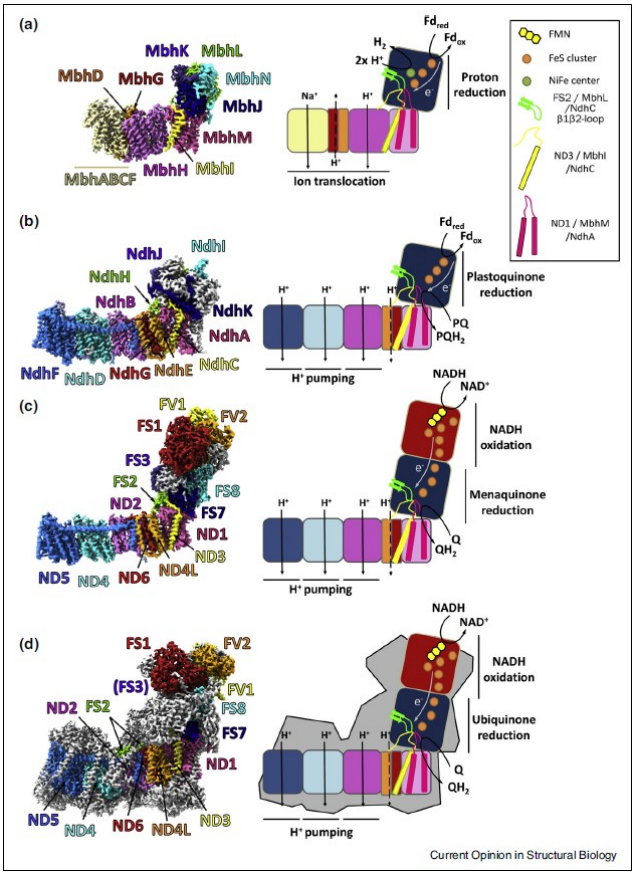

Figure 1.8 Conserved structural elements in complex I family proteins

Structures and schematic representations of (a) MBH complex from P. furiosus (PDB ID: 6CFW, EMD-7468) (b) NDH complex from T. elongatus (PDB ID: 6HUM, EMD-0281) (c) bacterial complex I from T. thermophilus (PDB ID: 4HEA) and (d) mitochondrial complex I from mouse (PDB ID: 6G2J, EMD-4345); subunits are coloured by homology (note that MbhI is related to ND3 and ND5). Accessory subunits are shown in gray. Conserved loops and structural elements are shown. Figure adapted from (Parey *et al*, 2020)

NDH has missing NADH-oxidizing N-module, however, the rest of the structure (11 subunits) is conserved (Figure 1.8). In addition, cyanobacterial NDH contains additional specific subunits, namely NdhL, NdhM, NdhN, NdhO, NdhP, NdhQ, NdhS, and NdhV (Pan *et al*, 2020). It binds ferredoxin (near subunits NdhV and NdhS) as electron donor to reduce plastoquinone (He *et al*, 2015). As Fd carries only one electron, two molecules are required for the reaction.

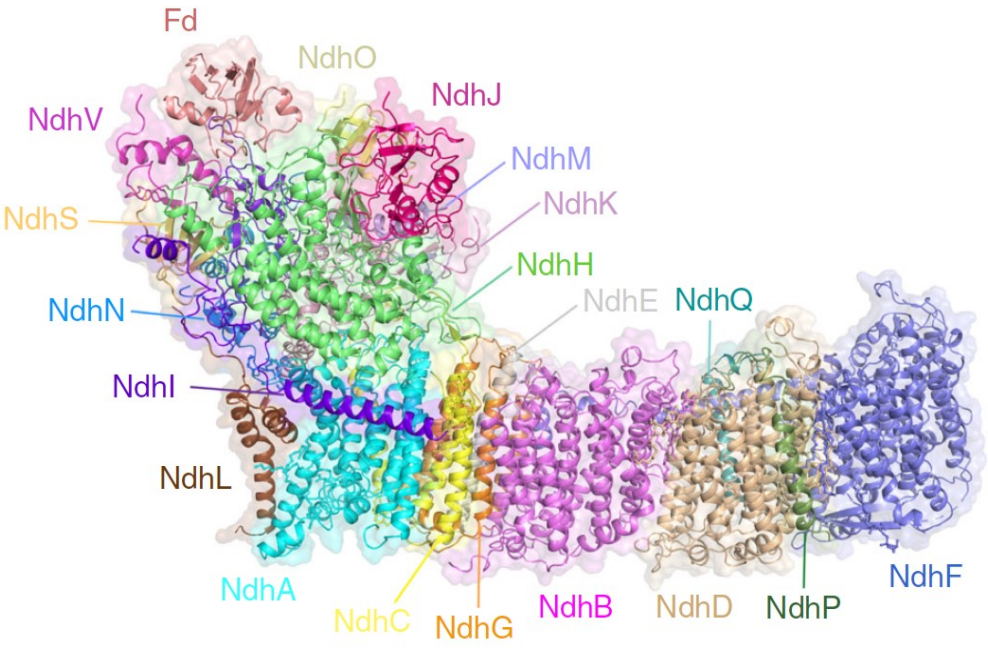

Figure 1.9 Structure of cyanobacterial NDH with bound Fd.

Subunits are coloured in different colours. Figure adapted from (Pan *et al*, 2020)*, PDB id: 6KHI.*

Several structures of NDH, including Fd bound state (Figure 1.9), have been resolved by Cryo-EM (Schuller *et al*, 2019; Pan *et al*, 2020; Zhang *et al*, 2020). However, there is still lack of mechanistic understanding of the enzyme.

1.2.5 Bacterial complex I as an important model

Despite the rapid progress made on complex I in the recent years, the mechanism of complex I is still hotly debated. *Ec*CI has been used for decades to reveal key features of complex I in general (Ohnishi *et al*, 2018). However, the structure of the entire complex is still unknown. So far, the gold standard model structure for interpreting experimental data was *Thermus thermophilus* complex I, which enabled remarkable progress for enzyme understanding (Baradaran *et al*, 2013). However, this model system has some disadvantages (compared to *E. coli*) which hinders further research. *T. thermophilus* is a thermophilic organism, enzymatic machinery of which is adapted to work at high temperatures and it uses exclusively menaquinone as a substrate. This makes it a somewhat specific model organism since adaptations might affect the structure. In this sense, *E. coli* is a better candidate for research, as it is mesophilic and uses both menaquinone and ubiquinone, and the last is the substrate of mammalian complex I.

Any mechanistic predictions should be ideally verified by site-directed mutagenesis. The availability of public libraries of mutants, as well as mutants prepared and purified in our lab, make *E. coli* the perfect model system. Solving structure of wild type *Ec*CI is the key to interpret already published mutations and the other experimental data. Moreover, it would allow to establish a robust protocol for solving structures of mutants and define new interesting key regions for mutations. Inactivating mutations may help to lock enzyme in particular intermediate steps of the catalytic cycle. Visualizing those "snapshots" would be crucial for understanding the general complex I mechanism.

Comparing complex I structures from different organisms is essential to reveal evolutionary traces and make general conclusions regarding the mechanism excluding species specific features. Structurally, the two bacterial systems (*Thermus* and *E. coli*) are quite different, particularly in the peripheral arm. The first and major difference is that NuoC and D subunits in are fused into the one single subunit (NuoCD) in *Ec*CI. Since NuoCD is part of the Q binding site, it is important to reveal the structural differences. Those differences might address the question why *Ec*CI is able to use both menaquinone and ubiquinone as substrates. Also, subunit NuoG in EcCI contains large C-terminal extension of unknown structure. Secondly, in contrast to *Tt*CI the distant N1a cluster can be reduced by NADH in *Ec*CI. Therefore, it would be interesting to investigate the cluster's environment. Moreover, unlike TtCI and mammalian enzyme, *Ec*CI upon purification retains one native bound Q molecule per complex (Anna Wojciechowska, Phd Thesis, MRC Cambridge, 2013), providing the possibility to visualize bound long-chain quinone. This has not been possible using *Tt*CI and externally added Q, as quinone is extremely hydrophobic in its native-like long-chain variants.

In summary, solving the EcCI structure and complementing the functional results with structural studies will provide a high-impact contribution to the field of bioenergetics.

1.3 Current structural and mechanistic understanding of complex I

1.3.1 History of structure determination

Complex I has been studied for decades, however, the majority of structural data started to appear relatively recently with the development of cryo-EM method. In early 2003, for the first time, using EM imaging of negatively stained sample, L-shape of the complex was revealed (Sazanov *et al*, 2003). Due to the enzyme size and complexity, its complete structure was solved in parts. Firstly, the X-ray structure of the peripheral arm including cofactors of *Tt*CI was solved in 2006 (Sazanov & Hinchliffe, 2006). It was followed by the X-ray structure

of the membrane domain of *Ec*CI in 2011, which revealed an unique fold and homology to antiporters (Efremov & Sazanov, 2011). The structure of membrane domain was incomplete due to the missing NuoH subunit, which is the key component of the PA-MA interface. Finally, in 2013 the structure of the entire *Tt*CI was solved, which gave insights into the interface interactions and quinone binding active site (Baradaran *et al*, 2013). This was followed by the crystal structure of eukaryotic yeast complex I at 3.6-3.9 Å resolution in 2015 (Zickermann *et al*, 2015).

Later, with the emerging direct electron detectors for cryo-EM, it became easier to resolve structures of big molecules (Kühlbrandt, 2014). The first high resolution structure of the entire mammalian complex was solved in 2016 (Fiedorczuk *et al*, 2016). This was followed by higher resolution structures including structures in different states (Kampjut & Sazanov, 2020; Grba & Hirst, 2020; Parey *et al*, 2019).

1.3.2 Structural states of complex I

Structures of ovine CI (*Ov*CI) were solved at high resolution, visualising water molecules, and in several redox states including turnover, which allowed authors to propose the mechanism of mammalian complex I involving cycling between the open and closed states (Kampjut & Sazanov, 2020).

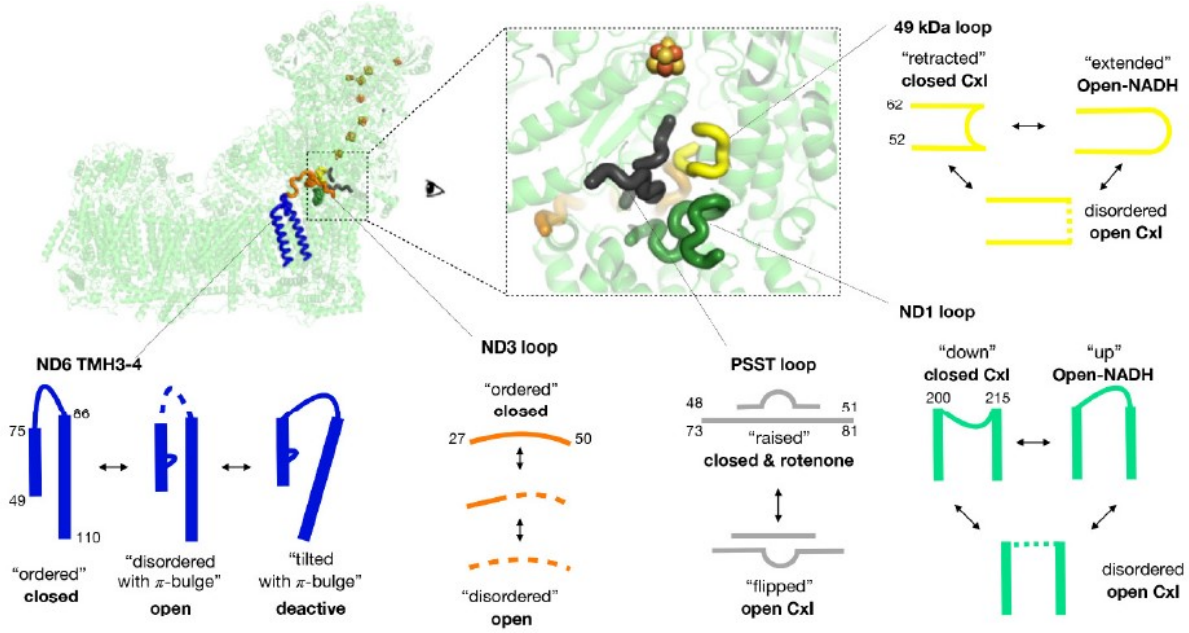

Figure 1.10 Conformational dynamics of open-to-closed transitions in mammalian complex I

Conformational dynamics of the catalytic site loops. Loops adopt different conformations upon complex I transition from open to closed state. Residue numbers indicate the start and end of the loops. The terminal N2 and the rest iron-sulfur clusters shown as orange and yellow spheres. Figure adapted from (Kampjut & Sazanov, 2020)

The opening comprises unfolding of key conserved loops around Q cavity, including NuoCD (49kDa) 220-230 β1-β2 loop, NuoB (PSST) 83-90 loop and NuoH (ND1) 208-230 TM5-6 loop, as well as PA/MA interface-flanking NuoA (ND3) 41-62 TM1-2 loop (Figure 1.10). Quinone binds in the open state, where expanded Q cavity facilitates quinone movement, but the reduction of quinone happens only in the closed state, when the cavity tightly engulfs the quinone. The transition to the closed state involves the striking rotation of the key ⱼTM3 (prefix indicates, for simplicity, *E. coli* nomenclature subunit), so that its π-bulge disappears. This leads to the formation of water wire from the Q site towards antiporter-like subunits (ALS) so that the protons for Q reduction come from the central axis. This brings redox "charge action" into ALS, initiating proton pumping, driven purely electrostatically since no conformational changes were observed within ALS. It was also tentatively suggested, on the basis of water molecules distribution, that a proton pathway to the periplasm is formed only in the distal ALS NuoL.

This radically novel mechanism raised debates and a proliferation of alternative proposals (Grba & Hirst, 2020; Kolata & Efremov, 2021; Mühlbauer *et al*, 2020; Galemou Yoga *et al*, 2020). Only mammalian enzyme was observed so far in closed and open states. All other CI structures – from *T. thermophilus* (Gutiérrez-Fernández *et al*, 2020), *E. coli* (Kolata & Efremov, 2021), *Y. lipolytica* (Parey *et al*, 2021), plant mitochondrial (Klusch *et al*, 2021), complex I homologue NDH (Schuller *et al*, 2019) show π-bulge conformation of ⱼTM3, i.e. open-like state, even though the degree of order of Q site loops varies. Therefore the validity of open-closed transition as part of catalytic cycle is questioned. Importantly, mammalian enzyme differs from other species by showing active/deactive (A/D) transition (Kotlyar & Vinogradov, 1990), suggesting the presence of high-energy barriers between different states, so that closed apo state can be observed. The deactive state resembles the open state by the conformation of ⱼTM3 and Q loops (Agip *et al*, 2018) , but is distinct due to relocation of ⱼTM4. Nevertheless, this similarity, along with the fact that closed/open ratio did not change under different conditions (Kampjut & Sazanov, 2020) , lead to suggestions that closed/open states of mammalian enzyme should be considered as active/deactive. Additionally, it is not clear how functionally important for the open state is the disorder of Q loops and how to explain the putative ND5-only proton ejection.

1.3.3 Coupling mechanism in complex I

The coupling mechanism between electron transfer and proton translocation is one of the biggest enigma in bioenergetics and structural biology (Sazanov, 2015; Zickermann *et al*, 2015; Hirst, 2013; Kaila, 2018; Kampjut & Sazanov, 2020, 2022).

Recently, the coupling mechanism for the mammalian complex I has been proposed (Kampjut & Sazanov, 2020). The mechanism comprises conformational changes between open and closed states and long-range electrostatic interactions (Figure 1.11). The key to the mechanism is the negative charge on quinone molecule, which forms after accepting the electrons; the negative charge pulls protons from the MA for the complete reduction and henceforth shifting the entire charge along MA. This allows protons to redistribute and eventually leave MA on the other side. Other possibilities include two-state stabilization mechanism and electrostatically driven "wave-spring" model. The two-state stabilization-change mechanism suggests that stabilization of negatively charged quinone intermediates drives a conformational changes in two "strokes" (quinone/semiquinone and semiquinone/quinol transitions), thereby transmitting energy to the membrane arm to drive proton pumping (Brandt, 2011; Zickermann *et al*, 2015). The electrostatic "wave-spring" model suggests that the negative charge on the quinone molecule drives an electrostatic wave of interactions between charged residues within the membrane, which results in one (or two) protons release. Subsequently, neutralization and relaxation of the molecule drives the reverse wave with two proton translocation (Verkhovskaya & Bloch, 2013).

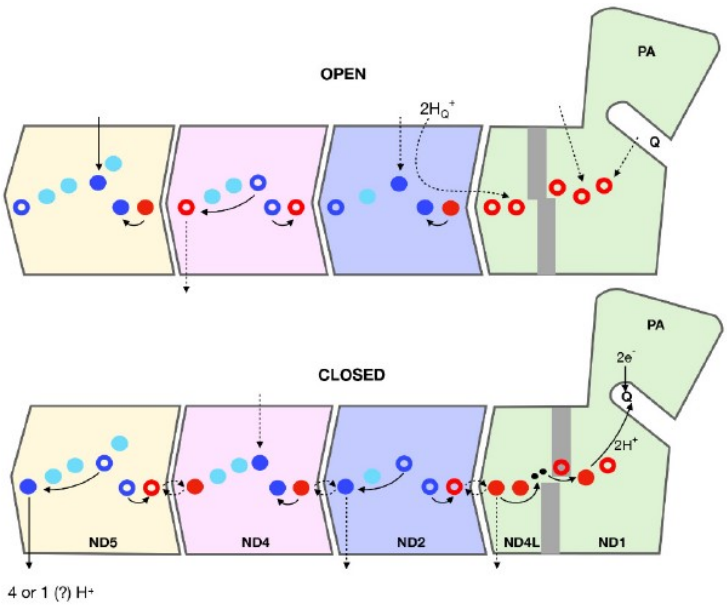

Figure 1.11 Schematic representation of the coupling mechanism proposed by (Kampjut & Sazanov, 2020)

Charge distribution of the proton pumping residues differs between the open and the closed states. Colored circles correspond to glutamate/aspartate (red), lysine (blue) or histidine (cyan). Full circles represent charged residues

and empty circles are neutral (state of histidines is not clear), black dots in the closed state represent water molecules. Arrows show proton movements to achieve the charge distribution as depicted in each state. Figure adapted from (Kampjut & Sazanov, 2022)

The mechanism proposed in (Kampjut & Sazanov, 2020) has been the most detailed and complete by the date and is therefore selected as a basis and the main working hypothesis for this work. Although it has been enthusiastically received by the scientific community, at the same time it also raised a further debate with a recent proliferation of publications with alternative proposals (Grba & Hirst, 2020; Kolata & Efremov, 2021; Mühlbauer *et al*, 2020; Galemou Yoga *et al*, 2020). Indeed, the mechanism was quite radical and poised unresolved questions: are some of the observed features specific for mammalian enzyme? What are the actual proton translocation pathways: through all the antiporters or only the terminal one? What is the role of the open state with the disordered quinone site? What is the exact sequence of protonation events within the antiporters? Thus, the main objective of this work is to resolve these issues and further elaborate the coupling mechanism.

1.4 Goals of the project

1. **To understand general features of the coupling mechanism**

The first goal of this project is to resolve high resolution structure of the entire *E. coli* complex I. This will be followed by other structures in different states: inhibited, reduced, with bound quinone and under catalytic turnover. These structures are necessary to observe conformational dynamics of the complex and derive general, non-specific features of the coupling mechanism. Besides, the structures will also help to explain features specific for *Ec*CI, i.e. unusual N1a cluster, rotenone insensitivity, etc.

2. **Investigate conformational dynamics of NDH complex**

The structure of NDH has been solved, however, not much is known about its mechanism and conformational dynamics. The second part of this work aims to study conformational dynamics and structure of NDH using extensive cryo-EM and structural data analysis. It would be interesting to see whether the two evolutionary related but distant structures (*Ec*CI and NDH) share the same mechanistic features, given that most of the core subunits conserved.

2 Methods

2.1 Sample Preparation

2.1.1 Sample preparation and purification

*Ec*CI was purified from BL21 (DDM and DDM/LMNG datasets) or MC4100 (LMNG datasets) *E. coli* cells by anion exchange and size exclusion chromatography as described before (Sazanov *et al*, 2003). This procedure gave a pure and active native protein preparation without any added purification tags. Concentrated *Ec*CI stocks were stored under liquid nitrogen in small aliquots in 20 mM Bis-Tris pH 6.0, 50 mM NaCl, 2 mM $CaCl_2$, 18% glycerol and ~2% DDM in case of DDM and DDM/LMNG datasets while LMNG purified protein was stored in 20 mM MES-HCl pH 6.0, 200 mM NaCl, 30% glycerol and ~0.7% LMNG. The amount of detergent in protein stocks was high because of the final concentrating procedure.

Purification protocol for sample in LMNG was slightly adjusted from (Sazanov *et al*, 2003). In detail, 30 ml of MC4100 membranes were thawed (from -80 °C) with protease inhibitor added (-EDTA, PMSF 0.002%). Then the membranes were solubilized with LMNG (2% final) for 1 hour. After solubilisation the membranes were spun at 220000 *g* using 70Ti rotor for 1 hour. For the following anion exchange chromatography, two buffers were made: AFG and BFG. AFG buffer includes 20 mM MES-HCl pH 6, 0.05 LMNG, 10% glycerol, 0.002% PMSF, while BFG is AFG + 1M NaCl. Before application on HiLoad 26/10 Q-sepharose column, the sample was diluted in 15% final BFG buffer. The column was equilibrated in 90% AFG and 10% BFG buffer. Fraction elution gradient was following:

1st run Vol, ml	1st run % BFG	2st run Vol, ml	2st run % BFG
0	15	0	5
60	20	10	15
400	35	300	30
30	100	50	100
100	100	100	100

Fraction collection was started at 21% BGF. NADH:FeCy activity of each fraction was checked to determine the protein content. For the second run, fractions with the highest activity were pooled and diluted 1:1 with AFG buffer with further application on BioRad Bio-Scale DEAE20 column. The column was pre-equilibrated with 95% AFG and 5% BFG. Collection was started at 11% BGF and fractions with the highest activity were pooled and concentrated ~ 50 times. Finally, concentrated sample was applied on Superose 6 10/300 GL column. The

column was equilibrated with 20mM MES-HCl pH 6, 0.05% LMNG, 10% glycerol, 0.002% PMSF and 200 mM NaCl. After election, fractions were checked for NADH/DQ activity and subunit composition using SDS-page electrophoresis.

NDH was purified as described in (Nowaczyk *et al*, 2011) and stored at -80^{0}C until use.

2.2 Analytical methods

2.2.1 Complex I activity measurements

Enzyme activity assays for *Ec*CI were done using a Shimadzu UV-2600 UV-VIS spectrophotometer at 30 °C or 18 °C (lower limit) and with magnetic stirring (320 rpm). NADH:DQ and NADH:FeCy (ferricyanide) oxidoreduction activities were measured by following NADH (ε = 6.1 mM^{-1} cm^{-1}) oxidation at 340 nm. Assay buffer was very similar to the one used for the preparation of grids (20 mM MES pH6.0, 2mM $CaCl_2$, 250mM NaCl) except for the presence or absence of *E.coli* total lipid extract (ETL) (0.25 mg/ml, with 0.1% 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) carry-over from ETL stock) and different detergent concentrations (Figure 3.6). *Ec*CI was equilibrated for 3 min with 100 μM DQ before NADH addition and NADH:DQ activity measurement. In all cases, the reaction was started by NADH addition. Control experiments were performed in the presence of 30 μM piericidin A inhibitor.

2.2.2 NDH activity measurements attempts

As to this date, there is no activity measurement protocol published. It is connected with difficulties of using Fd as an electron donor (comparing to NADH in complex I). Thus, we tried to develop the assay for NDH.

Since NDH has exposed Fe-S cluster, it may be possible to reduce the complex with dithionite. It was assumed that the reduction of DQ by dithionite is not rapid, and would be faster with NDH added (and activity could be derived from the difference). However, dithionite reduced DQ rapidly. Another attempt was to create the system where the limiting substrate would not be Fd. For that, Ferredoxin-NADP+ oxidoreductase (FNR), which catalyzes electron transfer between Fd and NADPH was used. However, no change in the spectrum was observed. It is possible that using non-native DQ (instead of plastoquinone) and FNR (spinach instead of cyanobacterial) affected the result of the assay.

Both assays were carried out in a quarts curette using OceanOptics UV-VIS spectrophotometer. DQ (ε = 13.55 mM^{-1} cm^{-1}) reduction was observed at 278 nm wavelength. Spinach FNR (F0628) was obtained from Sigma-Aldrich.

2.3 Electron cryo-microscopy

Electron cryo-microscopy (cryo-EM) is an imaging technique used to observe the three-dimensional structure of biological molecules, such as proteins, at near-atomic resolution. In cryo-EM, a sample is frozen in a thin layer of ice, and then observed using an electron microscope. The technique has been used to obtain high-resolution images of a variety of biological molecules, including viruses, proteins, and DNA. In recent years, cryo-EM has become one of the most powerful tools in structural biology, due to its ability to obtain near-atomic resolution images without the need for crystals (Kühlbrandt, 2014). The main advantage of cryo-EM over other imaging techniques is that it does not require the sample to be in a crystalline state. This is particularly important for biological molecules, which are often difficult or impossible to crystallize. In addition, cryo-EM can be used to image molecules in their native state, which is often important for understanding their function. Moreover, the technique allows observation of several conformational states of a molecule. The main disadvantage of cryo-EM is that it is a low-throughput technique, meaning that only a small number of samples can be imaged at a time. In addition, cryo-EM requires specialized equipment and expertise, which can make it expensive and difficult to access.

2.4 Grid preparation and data collection

2.4.1 Grid preparation for complex I

In the case of *Ec*CI datasets in DDM, the protein was run through Superose 6 SEC column immediately before grid preparation in 20 mM MES pH 6.0, 2 mM $CaCl_2$, 250 mM NaCl, 0.02% DDM and concentrated to ~10 mg/ml (~0.4% DDM final). The high protein concentration was required to achieve adequate coverage of the holes with particles. Additionally, an increased NaCl concentration in the buffer was required to prevent protein aggregation on the grid. For DDM_NADH dataset, 5 mM NADH was added immediately before freezing (20 seconds between NADH addition and vitrification). Quantifoil Cu/carbon 0.6/1 grids were used. Before using, grids were glow discharged (0.7 mbar and 30 mA for 2 min in the ELMO Glow Discharge unit, Agar Scientific, Stansted, UK). Blotting was done at

4°C, 100% humidity with 25 blotting force and 5s blotting time. All blotting and freezing procedures were done with 2.7 µL sample applied on a grid using FEI Vitrobot Mark IV.

In the case of datasets in DDM/LMNG, protein stocks were not subject to SEC but rather diluted (from ~17 mg/ml to ~0.5 mg/ml) in dilution buffer DB (20 mM MES-HCl pH 6.0, 250 mM NaCl, 2 mM $CaCl_2$ 0.01% LMNG) and incubated for at least 4 hours on ice. In case of LMNG-only preparation, the protein was used right after purification, diluted in DB. *E. coli* total lipid extract (ETL; Avanti Polar Lipids Inc., Alabama, USA) was added to the protein sample before blotting and freezing. ETL stock was prepared by drying the initial stock (25 mg/ml in chloroform) under nitrogen gas stream, washing with diethyl ether, drying under N_2 and dissolving at 7.5 mg/ml in 10 mM MES pH 6, 3% CHAPS buffer.

For DDM/LMNG_Apo dataset, *Ec*CI (0.2 mg/ml) was mixed with ETL (0.25 mg/ml, 0.1% CHAPS carry-over with ETL stock) and frozen without substrates. For DDM/LMNG_PieA dataset, *Ec*CI (0.16 mg/ml) was mixed with ETL (0.25 mg/ml, 0.1% CHAPS) and piericidin A (50 µM) followed by 2 min incubation at room temperature, then NADH (1.2 mM) was added, mixed and applied on a grid. For DDM/LMNG_DQ dataset, the protein (0.16 mg/ml) was mixed with ETL (0.25 mg/ml, 0.1% CHAPS) and DQ (500 µM) followed by 2 min incubation at room temperature, then the dilution buffer was added with subsequent mixing and application on a grid. For DDM/LMNG_FMN+NADH dataset the protein (0.16 mg/ml) was mixed with ETL (0.25 mg/ml, 0.1% CHAPS), FMN (50 µM) and NADH (1.2 mM) with subsequent mixing and application on a grid. For LMNG_Apo dataset, *Ec*CI (0.25 mg/ml) was mixed with ETL (0.25 mg/ml, 0.1% CHAPS) and frozen without substrates All concentrations indicated are final concentrations. For all datasets in DDM/LMNG, we used Quantifoil Cu/carbon 0.6/1 grid with a home-made 1.3 nm (0.9 nm in case of LMNG-purified sample) amorphous carbon support layer (produced using Leica EM ACE600 sputter coater). Before using, grids were glow discharged (0.7 mbar and 30 mA for 10 s in the ELMO Glow Discharge unit, Agar Scientific, Stansted, UK). Blotting was done at 15°C, 100% humidity with 25 blotting force and 2 s blotting time. Using carbon-coated grids led to several preferable orientations of the complex but did not prevent from achieving high-resolution good quality maps.

Since the use of carbon-coated grids allowed us to minimise protein concentration, under turnover substrate depletion was avoided, ensuring full turnover conditions at the time of snap-freezing. For the turnover datasets *Ec*CI (0.16 mg/ml for DDM/LMNG pH 6 and pH 8 and 0.25 mg/ml for LMNG) was mixed with ETL (0.25 mg/ml, 0.1% CHAPS) and DQ (410 µM for DDM/LMNG pH 6 and 750 µM for DDM/LMNG pH 8 and for LMNG) and incubated

for 2 min at room temperature (22°C). Then NADH (1.5 mM) was rapidly added with mixing by aspiration, followed by application on a grid. It took ~20 s from NADH addition to protein freezing and most of this time the sample was in the Vitrobot chamber kept at 15°C. Since the enzymatic activity at 18ºC in lipids/DDM/LMNG is 5.3 μmol NADH/min/mg prot, the maximal amount of substrates to be used within 20 s (1/3 of min) would be ~ 5.3 x 0.16 x 1000 / 3 ~ 280 μM, ensuring that neither DQ or NADH are used up by the time of plunge-freezing. In addition, the exact same sample mixture and in the same conditions as used for grid application, was applied to a NanoDrop (Denovix DS-11) spectrophotometer kept at 22°C and the reaction kinetics was followed at 340 nm, confirming that it proceeded well beyond 20 s, significantly slowing only at about 40 s. The turnover of *Ec*CI in all cases at the time of freeze plunging is clearly confirmed by the presence of strong cryoEM density for FMN, NADH, NuoF and NuoE, which would otherwise be absent in the presence of NADH and absence of turnover. Finally, only under turnover we observe closed *Ec*CI and Q_d-bound quinone (in three independent datasets).

2.4.2 Grid preparation NDH

100 μl NDH stock (~2.5 mg/ml, stored under liquid N_2 in 20mM BisTris pH 6.0, 100 mM NaCl, 25% glycerol, ~0.2% DDM) was diluted in 400 μl 20mM BisTris pH 6.0, 100 mM NaCl, 0.01% dilution buffer and concentrated to ~30 μl. The procedure was done twice in order to reduce glycerol concentration. 0.1% CHAPS (final concentration) was added to the sample to achieve better spread of particles. Afterwards, concentrated sample was applied on a Quantifoil Cu/carbon 0.6/1 grid (glow discharged the same way as in *Ec*CI case). The blotting was done using 4 s blotting time and 25 blotting force. We found that NDH concentration of 6-8 mg/ml is the most optimal. NDH_Fd was prepared in the same manner, except with addition of ferredoxin (11 μm final) and dithionite (5 mM final) together with CHAPS before freezing.

2.4.3 Cryo-EM data collection

Around 3000-3500 images were collected for most *Ec*CI datasets, while for DDM-LMNG pH8 and LMNG turnover datasets around 8000 and 11000 images were collected, respectively. All *Ec*CI datasets in DDM and NDH_Apo were collected with TF Krios TEM at CEITEC electron microscope facility in Brno, while all *Ec*CI datasets in DDM/LMNG and LMNG and NDH_Fd were collected with TF Krios and Glacios TEM at IST Austria electron

microscope facility. Image collection settings and equipment are summarized in Appendix 1-4.

2.5 Image processing: Complex I

All data processing was done in RELION 3.0 and 3.1 (Zivanov *et al*, 2018) . All datasets had a similar processing strategy for the sake of consistency. Movies were aligned and dose-weighted using MotionCor2 (Zheng *et al*, 2017). CtfFind4 was used for CTF estimation on non-dose-weighted micrographs (Rohou & Grigorieff, 2015) .

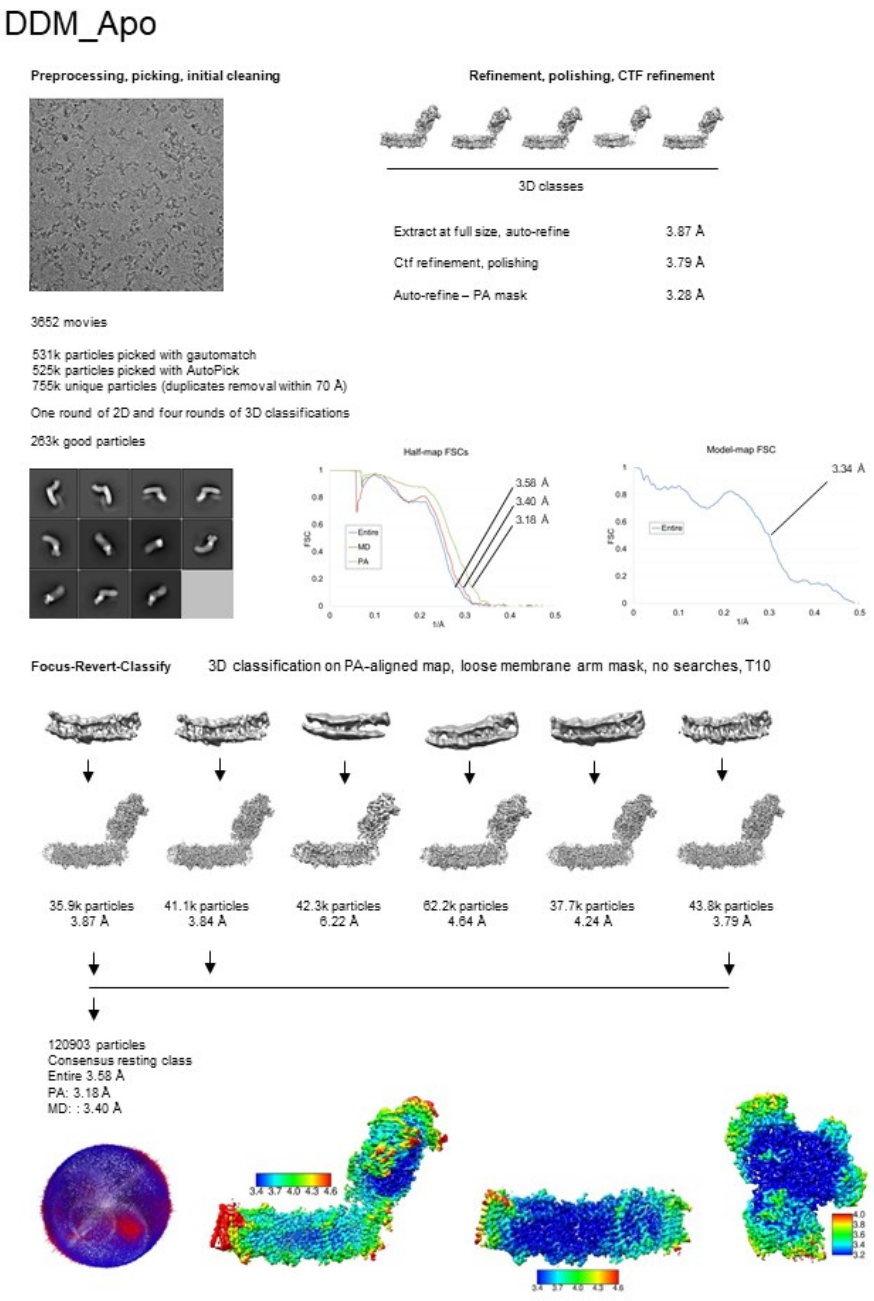

Figure 2.1 Processing scheme of DDM_APO dataset

Figure adapted from (Kravchuk *et al*, 2022)

About 3k movies were collected per dataset (Appendix 1-4). Poor micrographs were filtered based on AccumMotionTotal, CtfFigOfMerit and CtfMaxResolution parameters and manually by CTF image (e.g., removing micrographs with the ice ring). Our general processing procedure implemented four steps: picking, cleaning (using 2D and 3D classification), refinement (using CtfRefine and Polishing tools in RELION (Zivanov *et al*, 2019) and class separation using focus-revert-classify (FRC) strategy. The FRC strategy comprises initial focused alignment with a mask around the peripheral arm, followed by classification (with no angular or transitional searches) with a mask around the membrane arm (Letts *et al*, 2019). Given the diversity in hinge angles between the peripheral and membrane arms, aligning all the particles by the peripheral arm maximizes the differences during classification focused on membrane arm. The peripheral arm was chosen for initial alignment because it contains heavy FeS clusters, helping the alignment. 2D classifications were done with k = 100 and T = 2 unless otherwise stated.

For DDM_Apo (Figure 2.1), the first dataset collected, 2D classes from manually picked 3k particles were used as a template to initially pick 500k particles using RELION's Autopick. Particles were extracted 2X binned and classified using 2D and 3D classifications.

Low-pass filtered (40 Å) *Tt*CI (PDB 4HEA (Baradaran *et al*, 2013)) structure was used as the initial model at the very first 3D classification. Best 3D class was used as a template to pick 531k particles with Gautomatch v0.56 (K. Zhang software) and 525k particles with RELION's Autopick, leaving 755k unique particles after duplicates removal. Particles were extracted 2X binned and cleaned with one round of 2D and four rounds of 3D classifications. The first 3D classification was performed with k = 6 and T = 4 in two steps: 50 iterations with 7.5° global angular search and 10 iterations with 3.7° global angular search (569k particles remained). Further we performed 2D classification (560k particles remained). The second 3D classification was performed with k = 4 and T = 4 in two steps: 50 iterations with 7.5° global angular search and 20 iterations with 3.7° global angular search. Next, the duplicates were again removed and 467k particles remained. The third 3D classification was performed with k = 4 and T = 8 in two steps: 50 iterations with 7.5° global angular search, 10 iterations with 3.7° global angular search and 30 iterations with 1.8° local angular search (371k particles remained). The fourth 3D classification was done without angular searches with k = 5 and T = 6 (264k particles remained). Before the fourth 3D classification, the particles were refined with the loose mask around the molecule. After classifications, clean particles were refined into a single "consensus" structure, followed by per-particle defocus and per-particle trajectory refinement using CtfRefine and Polishing tools (CtfRefine – Bayesian polishing – CtfRefine). Refined

particles were subjected to FRC classification with k = 6 and T = 10. This resulted in six resting state classes that differed from each other only by the degree of openness. The three best classes were combined into one consensus resting class. The final density map was formed from focus-refined PA and MA maps. The PA and MA maps were aligned on the most populated class and merged into the final composite map using the "vop max" command in Chimera software (Pettersen *et al*, 2004).

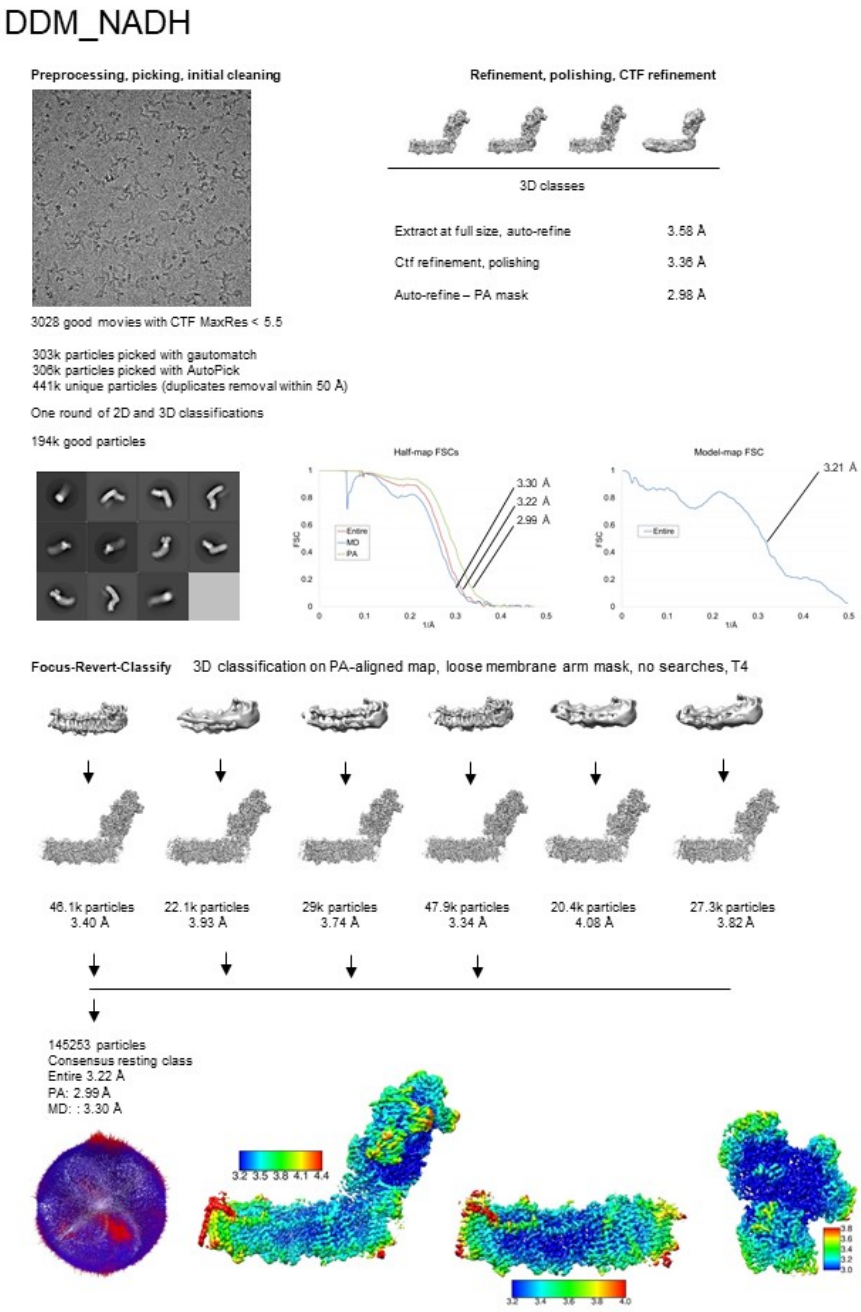

Figure 2.2 Processing scheme of DDM_NADH dataset

Figure adapted from (Kravchuk *et al*, 2022)

For DDM_NADH dataset (Figure 2.2), 303k particles were picked with Gautomatch and 306k particles with RELION's Autopick using *Ec*CI structure as a 3D template and joined with duplicates removal (441k unique particles). Particles were extracted 2X binned and cleaned using one round of 2D (290k particles remained) and one round of 3D classifications (226k particles remained). The 3D classification was done with k = 4 and T = 8 in four steps: 45 iterations with 7.5° global angular search, 10 iterations with 3.7° global angular search, 25 iterations with 1.8° local angular search and 10 iterations with 0.9° local angular search. Good classes were re-extracted at the full pixel size, followed by duplicates removal (194k particles remained). Next, particles were refined into a single "consensus" structure, followed by per-particle defocus and per-particle trajectory refinement using CtfRefine and Polishing tools. Refined particles were subject to FRC classification with k = 6 and T = 4. This resulted in six resting state classes that differed from each other only by the degree of openness. The four best classes were combined into one consensus resting class. The final density map was formed from focus-refined PA and MA maps. The PA and MA maps were aligned on the most populated class (class 4) and merged using "vop max" command in Chimera software.

For DDM/LMNG_Turnover_pH6 dataset (Figure 2.3), 773k particles were picked with RELION's Autopick using *Ec*CI structure as a 3D template. Particles were extracted 2X binned and cleaned using one round of 2D (629k particles remained) and one round of 3D classifications (269k particles remained). The 3D classification was done with k = 6 and T = 4 in three steps: 25 iterations with 7.5° global angular search, 5 iterations with 3.7° global angular search and 10 iterations with 1.8° local angular search. Good classes were re-extracted at the full pixel size. Next, particles were refined into a single "consensus" structure, followed by per-micrograph aberrations (two rounds), per-particle defocus (two rounds) and per-particle trajectory refinement using CtfRefine and Bayesian polishing tools. Another cleaning step was performed: particles were aligned on the PA and classified without searches with a loose mask around PA; then same was done for MD (204k particles remained).

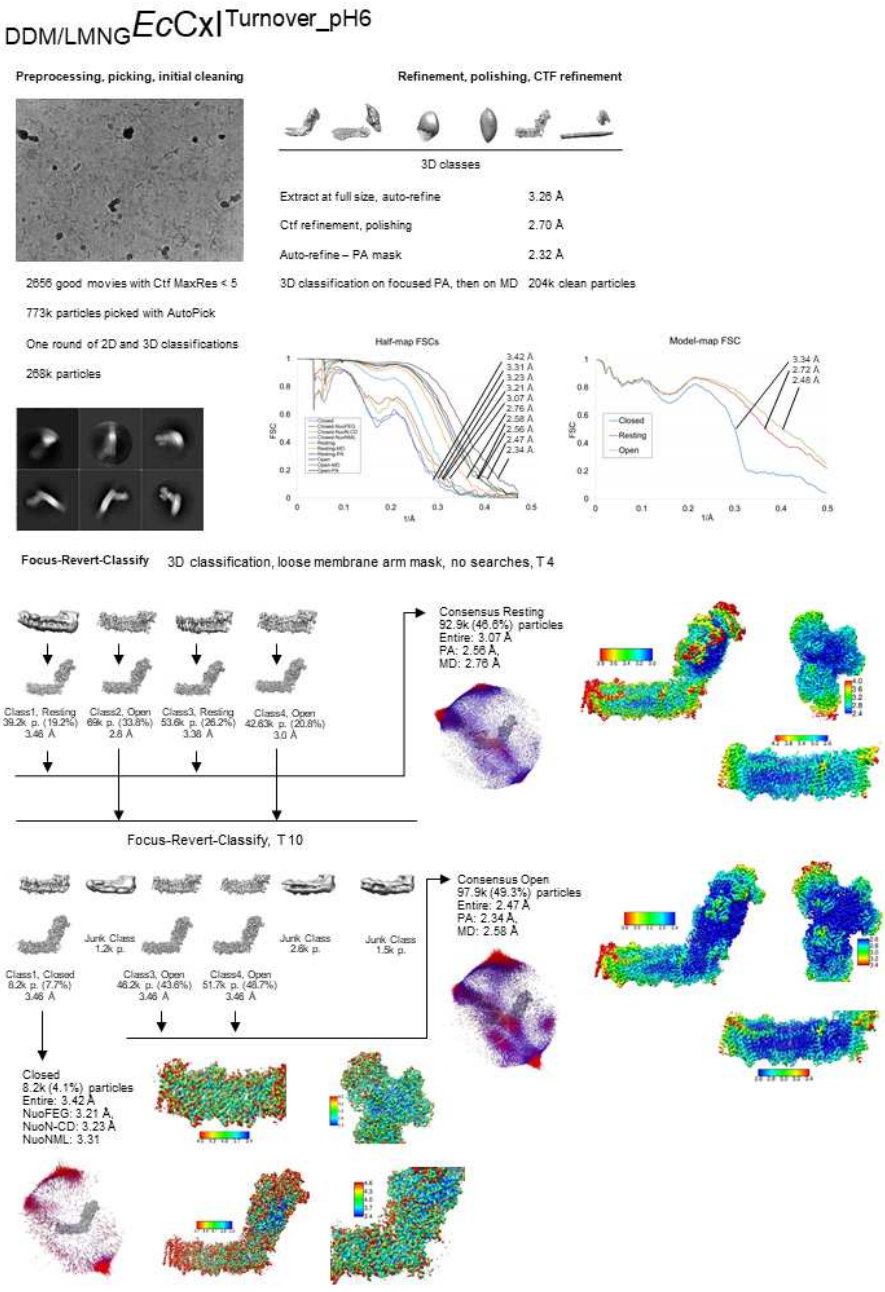

Figure 2.3 Processing scheme of DDM/LMNG_Turnover_pH6 dataset

Figure adapted from (Kravchuk *et al*, 2022)

Afterwards, we performed FRC classification with k = 4 and T = 4. This resulted in two resting classes and two open classes that differed from each other by degree of openness. The two resting classes were combined into one consensus resting class. The final density map was formed from focus-refined PA and MA maps. The PA and MA maps were aligned on the most populated resting class (class 3) and merged using “vop max” command in Chimera software. The particles from the two open classes were joined and FRC classified with k = 6 and T = 10.

This resulted in three good classes: one closed and two open. The two open classes were combined into one consensus open class. The final density map for the open class was formed from focus-refined PA and MA maps filtered by local resolution. The PA and MA maps were aligned on the most populated open class (class 4) and merged using "vop max" command in chimera software. The final density map for the closed class was formed from globally filtered (to the value of 0.143 FSC cut-off) focus-refined maps of NuoFEG, NuoAJKHCD (the PA-MA junction subunits) and NuoNML combined using "vop max" command in chimera software. These three focus areas were selected to achieve the improved density throughout MA, in view of limited resolution of the closed class.

For DDM/LMNG_Apo dataset (Figure 2.4), 918k particles were picked with RELION's Autopick using *Ec*CI structure as a 3D template. Particles were extracted at 2x the physical pixel size and cleaned using two rounds of 2D (487k particles remained) and one round of 3D classifications (387k particles remained). The first round of 2D classification was done with k =100 and T = 2 and the second round was done with k = 20 and T = 2.5 on each good class (obtained from the first round) separately. Afterwards, all good 2D classes were joined together. The 3D classification was done with k = 4 and T = 4 in three steps: 20 iterations with 7.5° global angular search, 10 iterations with 3.7° global angular search and 15 iterations with 1.8° local angular search. Good classes were re-extracted at the full pixel size. Next, particles were refined into a single "consensus" structure, followed by per-micrograph aberrations (two rounds), per-particle defocus (two rounds) and per-particle trajectory refinement using CtfRefine and Bayesian polishing tools. Another cleaning step was performed: particles were aligned on the PA and classified without searches with a loose mask around PA; then same was done for MA (366k particles remained). Afterwards, we performed FRC classification with k = 4 and T = 4. This resulted in two resting state classes and two open state classes that differed from each other by the degree of openness. The two resting classes were combined into one consensus resting class. The final density map was formed from focus-refined PA and MA maps. The PA and MA maps were aligned on the most populated resting state class and merged using "vop max" command in chimera software. The same procedure was done to obtain the final composite open class map.

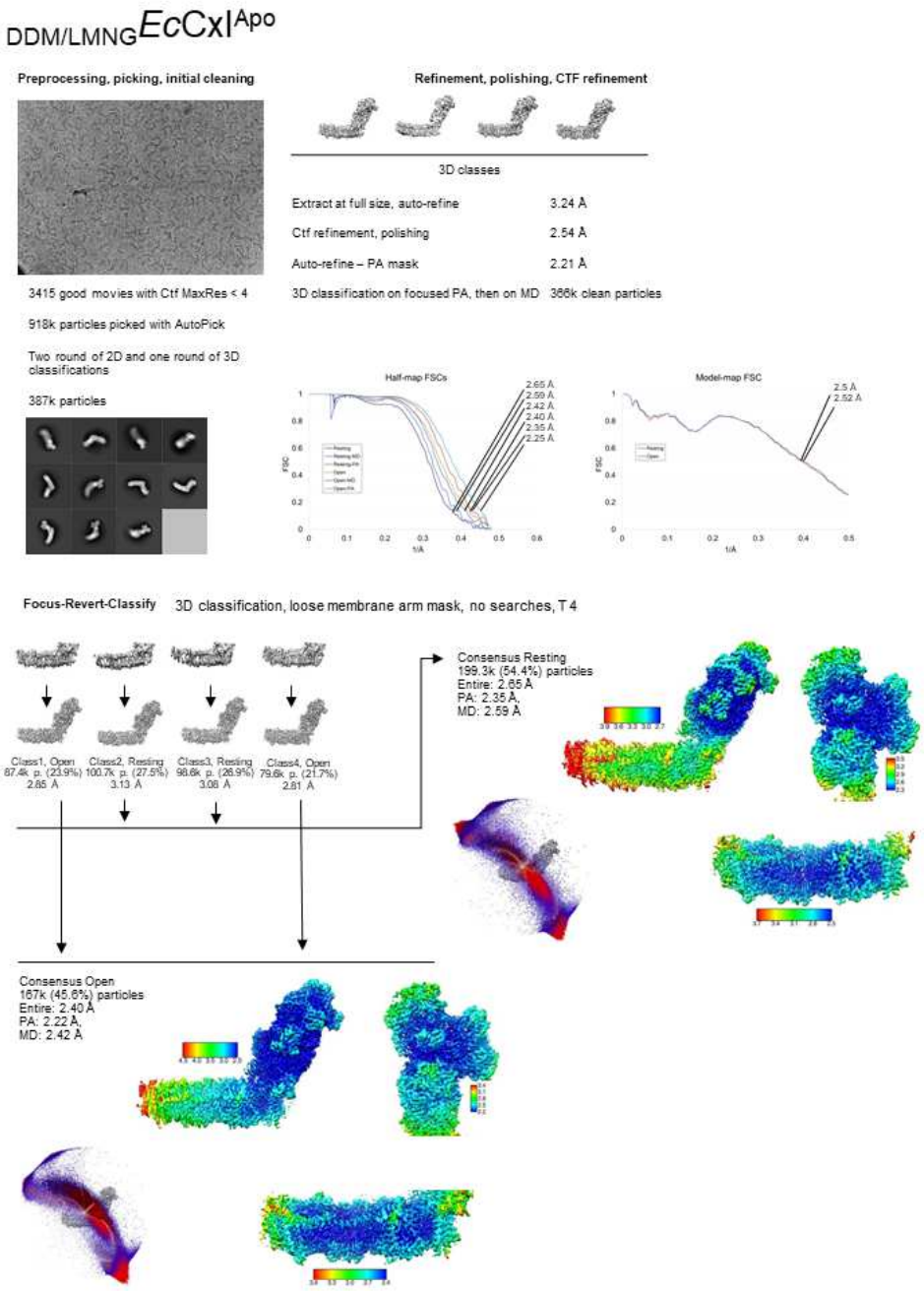

Figure 2.4 Processing scheme of DDM/LMNG_APO dataset

Figure adapted from (Kravchuk *et al*, 2022)

For DDM/LMNG_PieA dataset (Figure 2.5), 463k particles were picked with Gautomatch using 2D projections of 3D *Ec*CI structure (obtained from previous datasets) as a template. Particles were extracted 2X binned and cleaned using two rounds of 2D (256k particles remained) and one round of 3D classifications (233k particles remained). The first round of 2D classification was done with k =100 and T = 2 and the second round was done

with k = 20 and T = 2.5 on each good class (obtained from the first round) separately. Afterwards, all good 2D classes were joined together and duplicates removed.

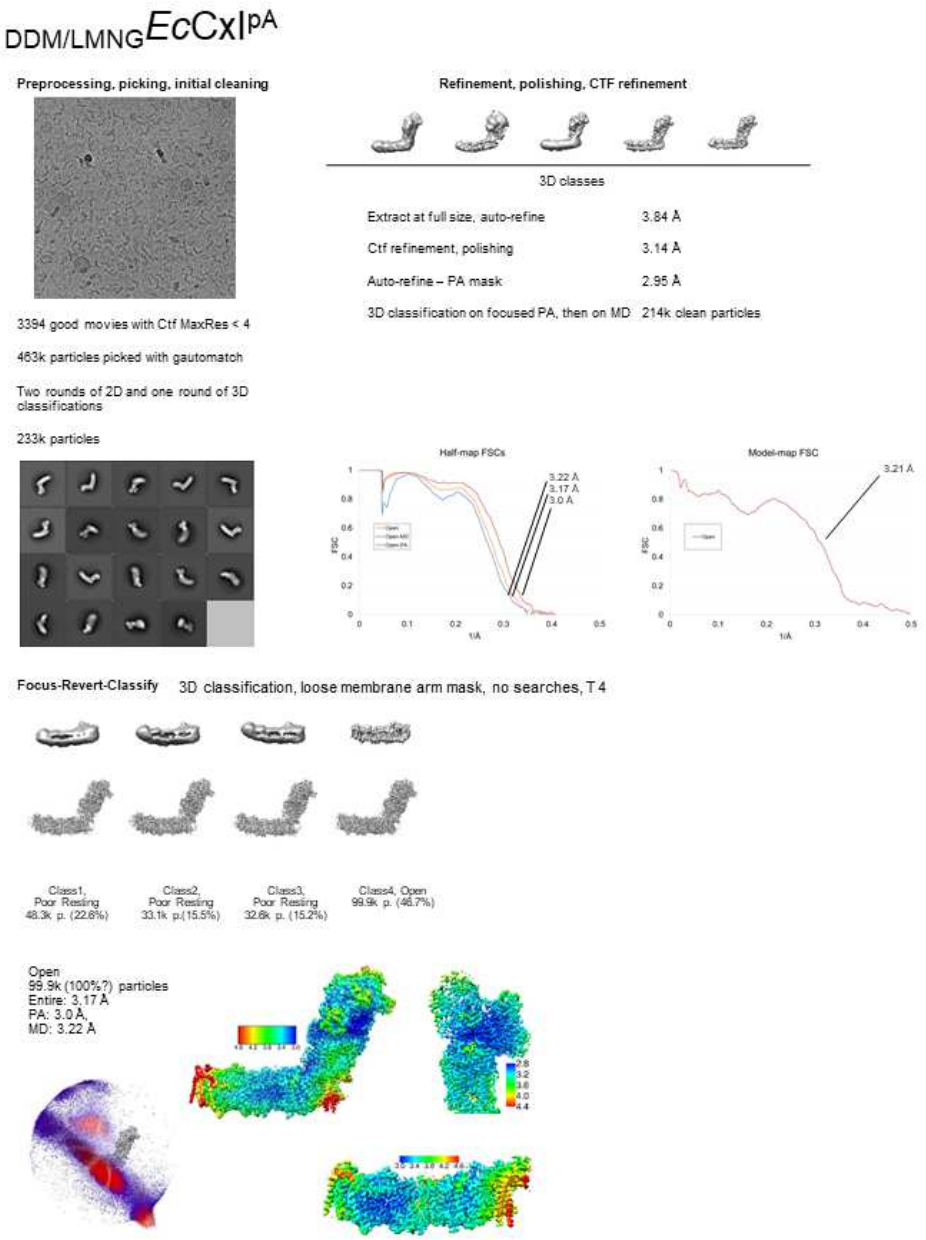

Figure 2.5 Processing scheme of DDM/LMNG_PieA dataset

Figure adapted from (Kravchuk *et al*, 2022)

The 3D classification was done with k = 5 and T = 4 in three steps: 25 iterations with 7.5° global angular search, 12 iterations with 3.7° global angular search and 7 iterations with 1.8° local angular search. Good classes were re-extracted at the full pixel size. Next, particles were refined into a single “consensus” structure, followed by per-micrograph aberrations (two

rounds), per-particle defocus (two rounds) and per-particle trajectory refinement using CtfRefine and Bayesian polishing tools. Another cleaning step was performed: particles were aligned on the PA and classified without searches with a loose mask around PA; then same was done for MA (214k particles remained). Afterwards, we performed FRC classification with k = 4 and T = 4. This resulted in three poor (low resolution) resting state classes and one open state class. Because of low resolution (> 4 Å) and absence of differences between the other resting classes, we did not model DDM/LMNG_PieA resting class. The final density map for the open class was formed from focus-refined PA and MA maps. The PA and MA maps were aligned on the entire initial map and merged using "vop max" command in Chimera software.

For DDM/LMNG_NADH+FMN dataset (Figure 2.6), 563k particles were picked with AutoPick using *Ec*CI structure as a 3D template.

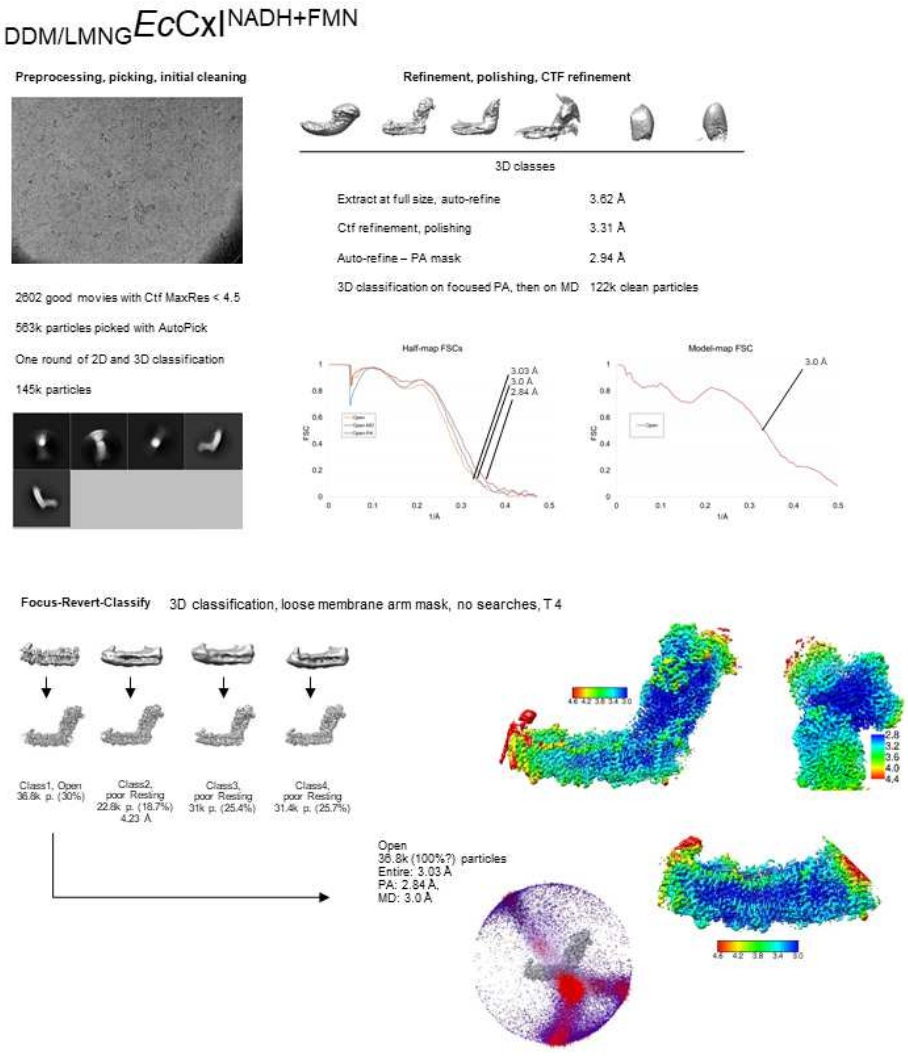

Figure 2.6 Processing scheme of DDM/LMNG_NADH+FMN dataset

Figure adapted from (Kravchuk *et al*, 2022)

Particles were extracted 2X binned and cleaned using one round of 2D (498k particles remained) and one round of 3D classifications (145k particles remained). The 3D classification was done with k = 6 and T = 4 in three steps: 20 iterations with 7.5° global angular search, 5 iterations with 3.7° global angular search and 15 iterations with 1.8° local angular search. Good classes were re-extracted at the full pixel size. Next, particles were refined into a single "consensus" structure, followed by per-micrograph aberrations (two rounds), per-particle defocus (two rounds) and per-particle trajectory refinement using CtfRefine and Bayesian polishing tools. Another cleaning step was performed: particles were aligned on the PA and classified without searches with a loose mask around PA; then same was done for MA (122k particles remained). Afterwards, we performed FRC classification with k = 4 and T = 4. This resulted in three poor (low resolution) resting classes and one open class. Because of low resolution (> 4 Å) and absence of differences between the other resting classes, we did not model DDM/LMNG_NADH+FMN resting class. The final density map for the open class was formed from focus-refined PA and MA maps. The PA and MA maps were aligned on the entire initial map and merged using "vop max" command in Chimera software.

For DDM/LMNG_DQ dataset (Figure 2.7), 825k particles were picked with RELION's Autopick using *Ec*CI structure as a 3D template. Particles were extracted 2X binned and cleaned using two rounds of 3D classifications. The first round of 3D classification was done with k = 6 and T = 4, 20 iterations with 7.5° global angular search (417k particles remained). The second round of 3D classification was done with k = 4 and T = 4 in three steps: 15 iterations with 7.5° global angular search, 10 iterations with 3.7° global angular search and 15 iterations with 1.8° local angular search (137k particles remained). Good classes were re-extracted at the full pixel size. Next, particles were refined into a single "consensus" structure, followed by per-micrograph aberrations (two rounds), per-particle defocus (two rounds) and per-particle trajectory refinement using CtfRefine and Bayesian polishing tools. Another cleaning step was performed: particles were aligned on the PA and classified without searches with a loose mask around PA; then same was done for MD (123k particles remained). Afterwards, we performed FRC classification with k = 4 and T = 4. This resulted in three resting state classes and one open state class that differed from each other by the degree of openness. The three resting state classes were combined into one consensus resting state class. The final density map was formed from focus-refined PA and MA maps. The PA and MA maps were aligned on the most populated resting state class and merged using "vop max" command in Chimera software. The same procedure was done to obtain the final composite open state class map.

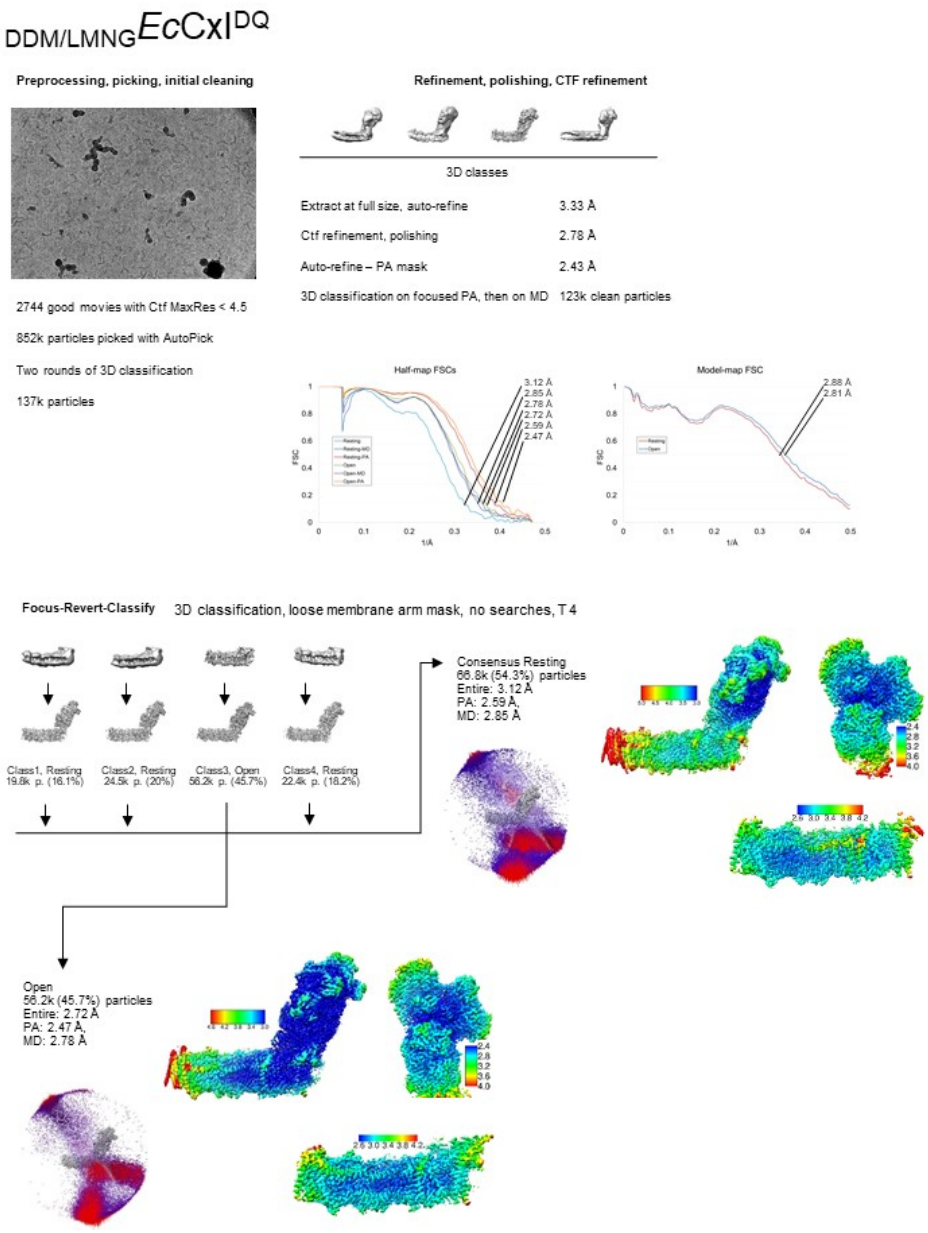

Figure 2.7 Processing scheme of DDM/LMNG_DQ dataset

Figure adapted from (Kravchuk *et al*, 2022)

For DDM/LMNG_Turnover_pH8 dataset (Figure 2.8), 1.65 million particles were picked with RELION's Autopick using *Ec*CI structure as a 3D template. Particles were extracted 2X binned and cleaned using one round of 2D (870k particles remained) and one round of 3D classifications (325k particles remained). The 3D classification was done with k = 4 and T = 4 in three steps: 20 iterations with 7.5° global angular search, 10 iterations with 3.7° global angular search and 10 iterations with 1.8° local angular search. Good classes were re-extracted at the full pixel size. Next, particles were refined into a single "consensus" structure, followed by per-micrograph aberrations (two rounds), per-particle defocus (two rounds) and per-particle trajectory refinement using CtfRefine and Bayesian polishing tools.

Afterwards, we performed FRC classification with k = 6 and T = 10. This resulted in two resting classes, two open classes, one closed and one junk class. The two resting classes were combined into one consensus resting class.

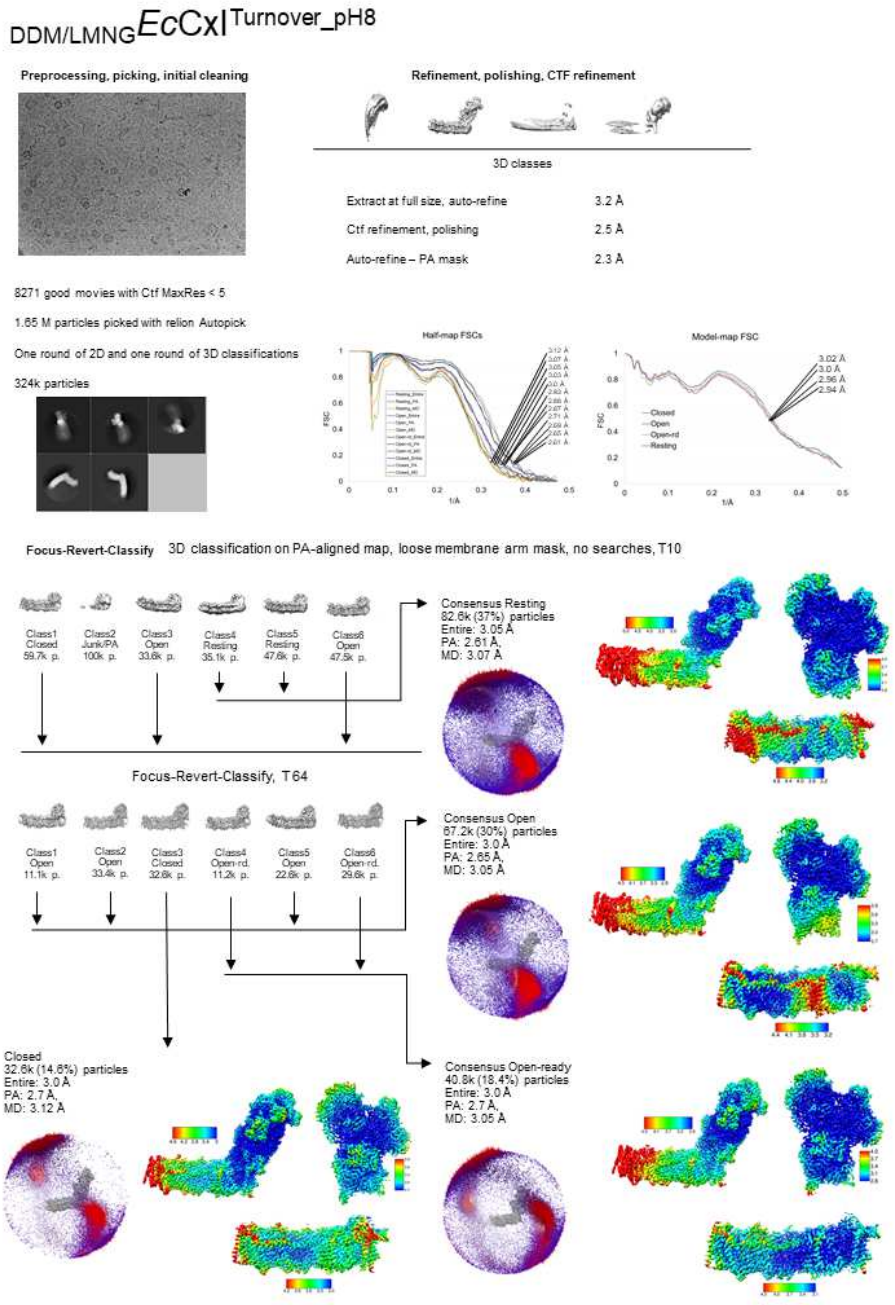

Figure 2.8 Processing scheme of DDM/LMNG_Turnover_pH8 dataset

Figure adapted from (Kravchuk *et al*, 2022)

The final density map was formed from focus-refined PA and MA maps. The PA and MA maps were aligned on the most populated resting class and merged using "vop max" command in Chimera software. The particles from the two open and the closed classes were

joined and FRC classified with k = 6 and T = 64. This resulted in three open, one closed and two open-ready classes. Same type classes were combined into consensus classes. The final density map for each class was formed from focus-refined PA and MA maps filtered by local resolution. The PA and MD maps were aligned on the most populated class and merged using "vop max" command in chimera software.

For LMNG_Apo dataset (Figure 2.9), ~1 million particles were picked with RELION's Autopick using *Ec*CI structure as a 3D template.

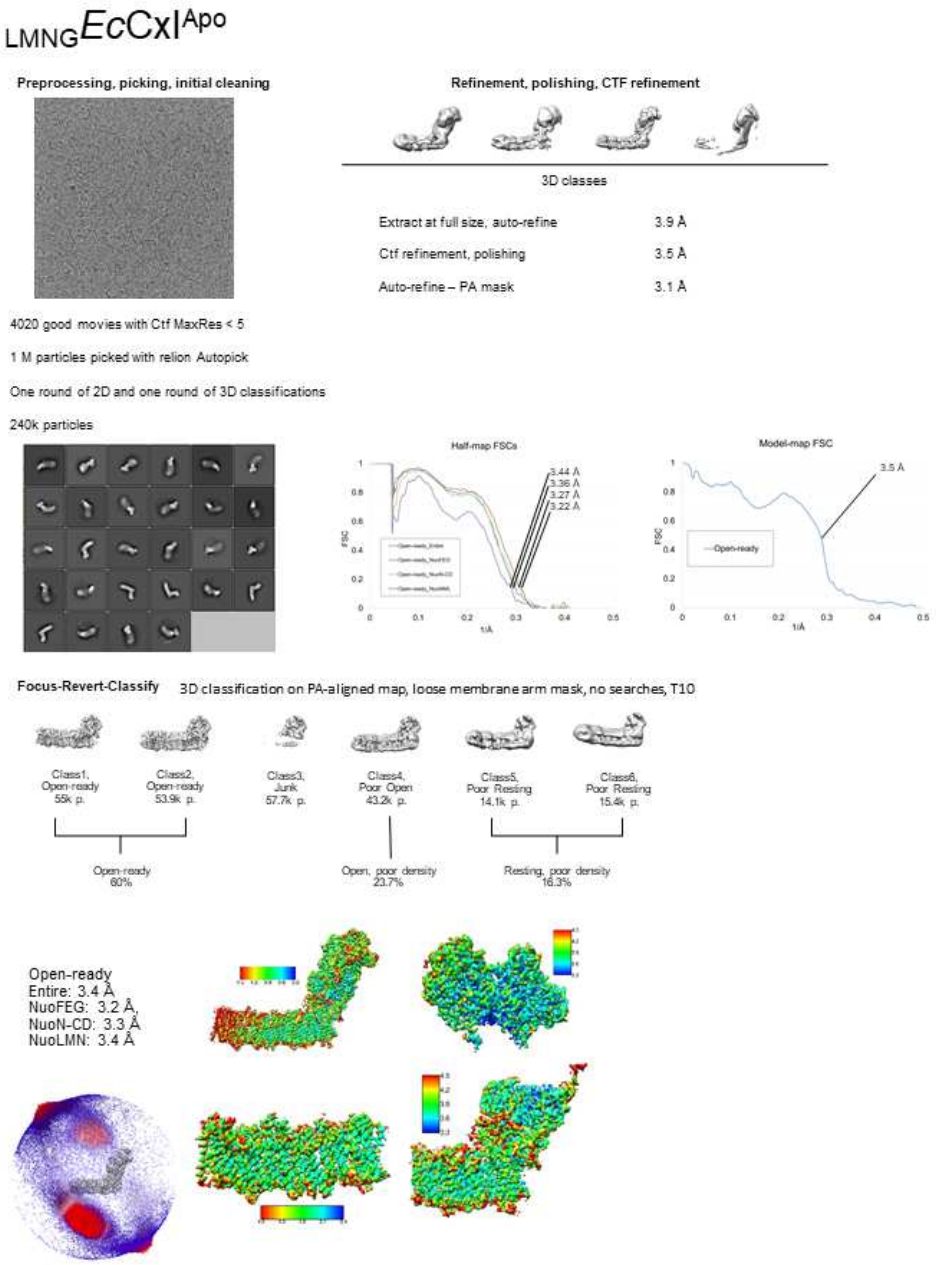

Figure 2.9 Processing scheme of LMNG_APO dataset

Figure adapted from (Kravchuk *et al*, 2022)

Particles were extracted 2X binned and cleaned using one round of 2D (550k particles remained) and one round of 3D classifications (240k particles remained). The 3D classification was done with k = 4 and T = 4 in three steps: 20 iterations with 7.5° global angular search, 10 iterations with 3.7° global angular search and 10 iterations with 1.8° local angular search. Good classes were re-extracted at the full pixel size. Next, particles were refined into a single "consensus" structure, followed by per-micrograph aberrations (two rounds), per-particle defocus (two rounds) and per-particle trajectory refinement using CtfRefine and Bayesian polishing tools. Afterwards, we performed FRC classification with k = 6 and T = 10. This resulted in two resting classes, two open-ready classes, one open and one junk class. Structures of the resting and open classes were not built because of poor density (4.5-5 Å resolution). The two open-ready classes were combined into one consensus class. The final density map was formed from globally filtered (to the value of 0.143 FSC cut-off) focus-refined maps of NuoFEG, NuoAJKHCD (the PA-MA junction subunits) and NuoNML combined using "vop max" command in chimera software.

For LMNG_Turnover dataset (Figure 2.10), ~4.5 million particles were picked with RELION's Autopick using *Ec*CI structure as a 3D template. Particles were extracted 2X binned and cleaned using one round of 2D (1.6 million particles remained) and one round of 3D classifications (708k particles remained). The 3D classification was done with k = 4 and T = 4 in three steps: 20 iterations with 7.5° global angular search, 10 iterations with 3.7° global angular search and 10 iterations with 1.8° local angular search. Good classes were re-extracted at the full pixel size. Next, particles were refined into a single "consensus" structure, followed by per-micrograph aberrations (two rounds), per-particle defocus (two rounds) and per-particle trajectory refinement using CtfRefine and Bayesian polishing tools. Afterwards, we performed FRC classification with k = 6 and T = 10. This resulted in 5 mixed (open, open-ready, resting) classes, and one closed class. The final density map for the closed class was formed from focus-refined PA and MA maps. The particles from the five classes were joined and FRC classified with k = 4 and T = 16. This resulted in two resting and two open-ready classes. The resting classes were combined into one consensus class. Open-ready classes were merged and subjected of another FRC classification with k = 6 and T = 64. This resulted in 5 open-ready classes with slightly different PA-MA angle and one open class. Open-ready classes were merged into one consensus class. The final density maps for each class were formed from focus-refined PA and MA maps filtered by local resolution. The PA and MA maps were aligned on the most populated class and merged using "vop max" command in chimera software.

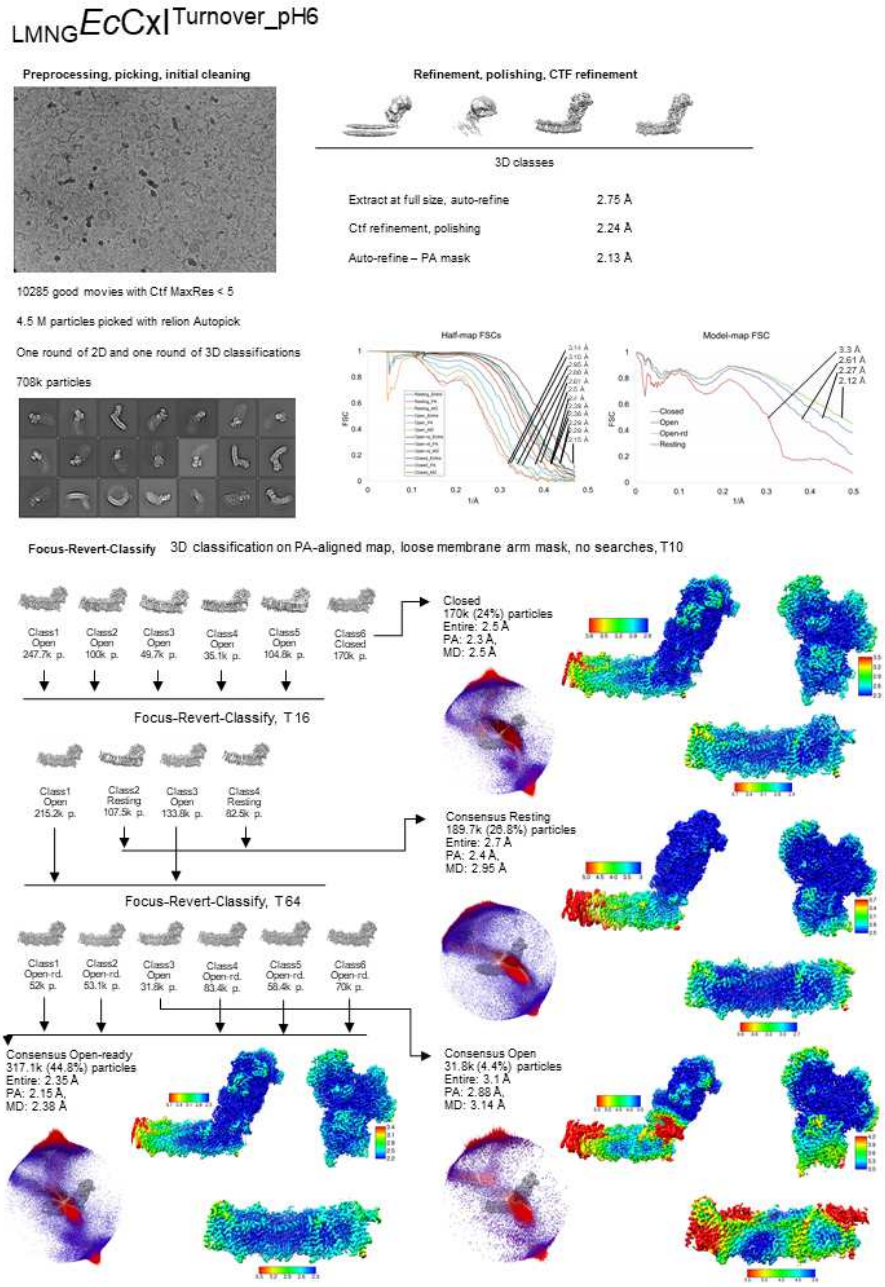

Figure 2.10 Processing scheme of LMNG_Turnover dataset

Figure adapted from (Kravchuk *et al*, 2022)

All resolutions are based on the gold-standard (two halves of data refined independently) FSC = 0.143 criterion. Local masks used for focused refinement correspond to regions shown as local resolution maps of PA and MA in Supplementary Figures S1-S13. Masks were created in RELION with extend 7 and soft-edge 10 pixels command. All maps were post-processed, B-factor sharpened and filtered by local resolution in RELION. The

density for the weaker features, such as some of bound quinones, is better defined in non-sharpened maps – these can be reproduced, if needed, using B-factor "Blur" feature in Coot.

2.6 Image processing: NDH

For NDH_APO 2721 good micrographs (after filtering by CTF and motion) were collected.

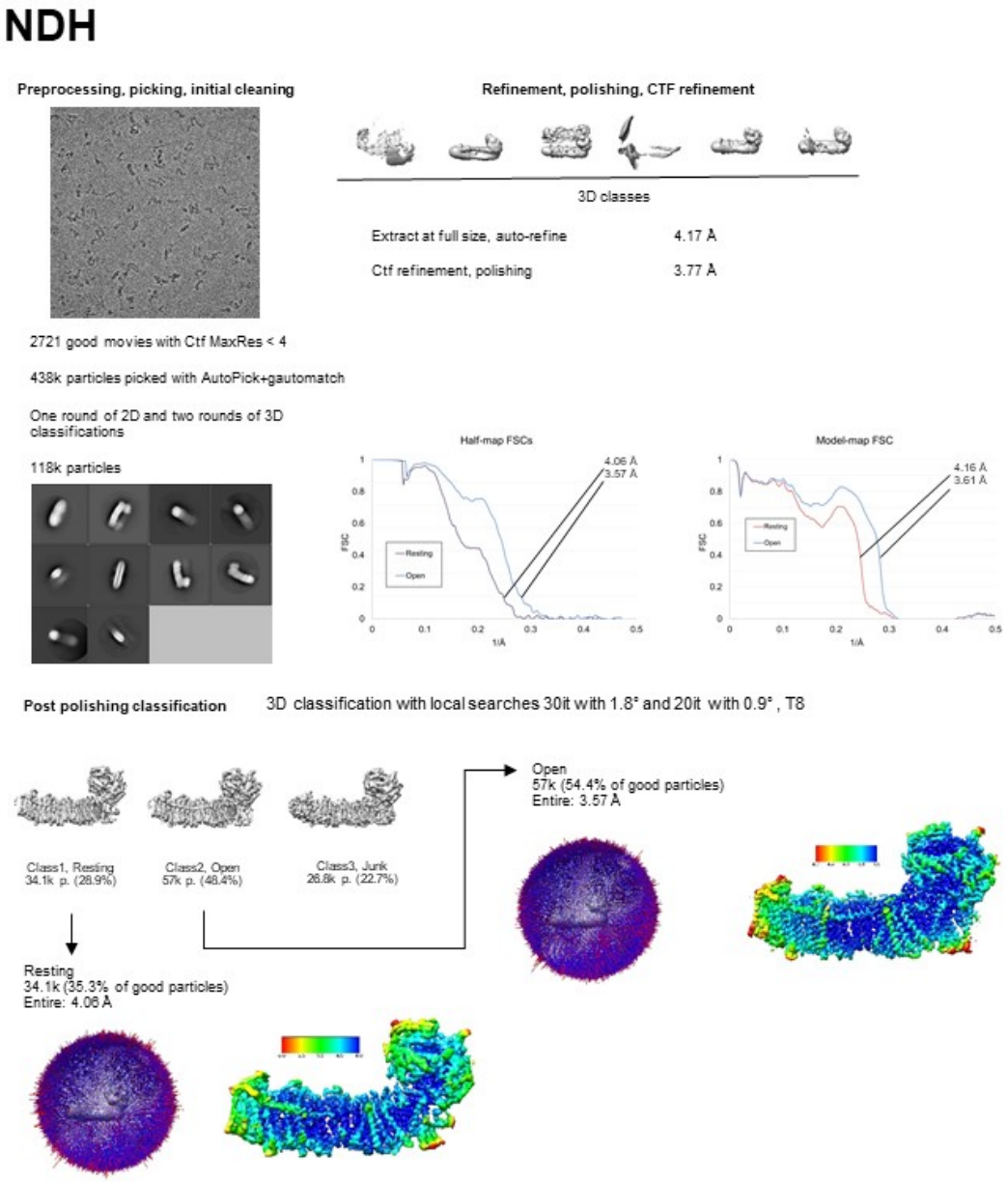

Figure 2.11 NDH Processing

2D classes from automatically picked particles (Laplacian blob (Zivanov *et al*, 2018)) were used as a template to pick 438k particles using RELION's Autopick and Gautomatch

v0.56 (K. Zhang software) combined. Particles were extracted 2X binned and cleaned using one round of 2D (300k particles remained) and two rounds of 3D classifications (118k particles remained). The first 3D classification was done with k = 6 and T = 4 and the second with k = 4 and T = 5. Low-pass filtered (40 Å) *Tt*CI (PDB 4HEA (Baradaran *et al*, 2013)) structure with manually removed N-module was used as the initial model at the very first 3D classification. Good classes were re-extracted at the full pixel size. Next, particles were refined into a single "consensus" structure, followed by per-micrograph aberrations (two rounds), per-particle defocus (two rounds) and per-particle trajectory refinement using CtfRefine and Bayesian polishing tools. Afterwards, another local 3D classification (with 1.8° and 0.9° angular searches) with k = 3 and T = 8 was performed. This resulted in a resting state, open state and a junk classes. Open and Resting classes where then separately refined.

2.7 Atomic model building and refinement

2.7.1 Complex I model building

Initial models for the subunits of the peripheral arm, NuoH and NuoA were generated using homology modeling implemented in Phyre2 server (Kelley *et al*, 2015), using *Tt*CI (PDB 4HEA) as a template. Coordinates for the rest of the subunits were from the X-Ray structure of *Ec*CI membrane arm (PDB ID 3RKO (Efremov & Sazanov, 2011)). All the subunits were fitted as rigid bodies with UCSF Chimera into the best resolved focus-refined maps of the peripheral and membrane arms. One round of five cycles of a real space refinement workflow in PHENIX (Adams *et al*, 2010) software was applied to resolve atom clashes and create the initial structure. Further, the initial structure was manually corrected in Coot (Emsley *et al*, 2010) with *de-novo* rebuilding of incorrectly homology-modelled parts (mainly in NuoG, NuoCD and NuoB). Secondary structure identification and modelling of poor density regions were assisted by data from PredictProtein server (Bernhofer *et al*, 2021). The initial structure was further improved by iterating manual adjustment in Coot and automated real space refinement in PHENIX, using our script which performs two rounds of a single cycle of ADP refinement with subsequent three cycles of global energy minimization to optimize B-factors so that electron radiation-damaged carboxylate side-chains acquire high B-factors and do not lead to main-chain distortions (Letts *et al*, 2019). This way we built template structures of the peripheral and membrane arms against the best resolved focus-refined maps. These template structures had been fit as rigid bodies into each class, followed by adjustment and correction with the iterative use of Coot and PHENIX. Topologies for lipids and ligands for refinement

in Coot and PHENIX were generated using the grade web-server (http://grade.globalphasing.org). In order to reliably build experimental water molecules, the maps were filtered by local resolution and resampled at 0.5 Å per pixel using relion_image_handler. After this procedure, water molecules displayed strong signal (>2σ), had nearly spherical densities, were not clashing with other atoms, and participated in hydrogen bonds, which are all strongly indicative of real water molecules. This allowed automatic placement of water molecules in Coot, which were then all checked and corrected manually guided by "undowse" feature of Molprobity web-server (Williams *et al*, 2018). Using this protocol, we could reliably build water molecules in the maps of higher than 2.8 Å resolution. The same resampled maps were used for depiction purposes.

2.7.2 NDH model building

Previously resolved NDH model (PDB ID 6HUM (Schuller *et al*, 2019)) was fit (membrane and peripheral domains separately) into density and refined using phenix and coot routine mentioned above. NdhL subunit in the resting state was almost de-novo rebuilt as the conformation was not observed before. Ligand topologies and quality control routine was similar to the described above.

2.8 Model analysis and figure preparation

Cavities and channels inside the protein were predicted using MOLE web-server. The overall quality of the models was assessed using Molprobity (Chen *et al*, 2010), Q-scores (Pintilie *et al*, 2020) and EMRinger (Barad *et al*, 2015). Visualization and analysis of protein density and structure, as well as figures preparation were done using PyMol, UCSF Chimera and ChimeraX (Goddard *et al*, 2018).

3 E. coli complex I – structure and mechanism

3.1 Catalytic activity and sample preparation challenges

At first, we purified *Ec*CI using DDM as a main detergent as established earlier (Sazanov *et al*, 2003) and after extensive grid optimization collected two datasets: native *Ec*CI (DDM_Apo) and *Ec*CI with NADH (DDM_NADH). Image processing and classifications have shown that the complex was quite flexible and adopted different conformational states (Figure 3.1b).

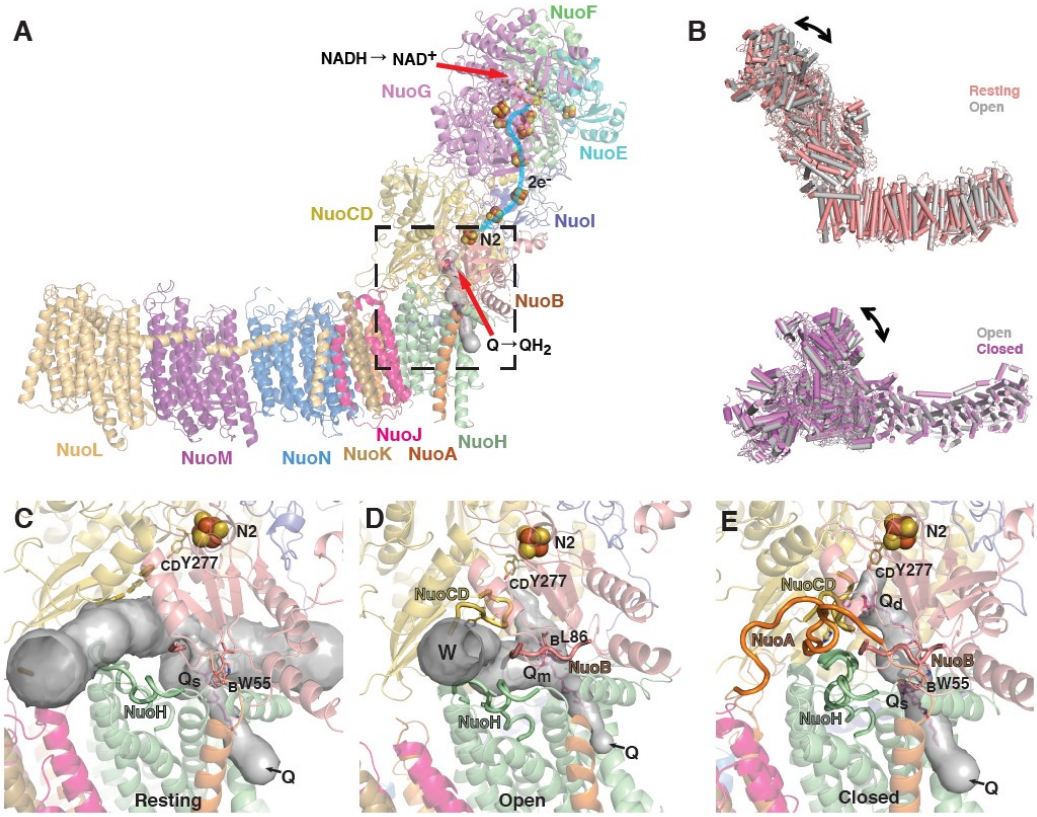

Figure 3.1 The overall structure of the E. coli complex I in different states

A) An overview of the structure. Different subunits are coloured and labelled in corresponding colours. NADH and quinone binding sites are indicated. Fe-S clusters are shown as spheres. The area around Q cavity, shown in C-E, is squared. B) E. coli complex I exists in three conformational states: resting, open (with open-ready state as its subtype) and closed. The structures of different states were aligned on MA to show conformational dynamics of PA. C-E) The environment of Q cavity in different states. The cavity calculated in MOLE is shown as grey surface and key loops lining the cavity are shown in orange (NuoA), yellow (NuoCD) and salmon (NuoB). C) The resting state has fully exposed Q cavity with most key loops disordered, except for NuoH, which is in the extended conformation. In some cases quinone can be found bound in Q_s (shallow) site. D) The open state has re-formed Q cavity with NuoCD and NuoB ordered. NuoH and NuoA loops are partly disordered, therefore the cavity is not enclosed and solvent can penetrate inside through the indicated W site. Quinone enters via the indicated Q entry and binds in the Q_m (median) site (magenta sticks). E) In the closed state all the key loops are ordered and therefore the cavity is enclosed. Quinone is bound in the Q_d (deep) site, while short-chain DQ can also bind in the Q_s site. The closing of the complex is accompanied by the change in conformation of NuoCD and NuoH loops. Figure adapted from (Kravchuk *et al*, 2022).

The nature of these conformational changes was similar to reported previously rotating-hinge-like motions between the peripheral and membrane arms which are likely to be typical for complex I (Fiedorczuk *et al*, 2016; Letts *et al*, 2019). However, after finishing *Ec*CI models and comparing them to the other known CI structures we revealed that in all conformational states from these datasets *Ec*CI had quite disordered and exposed Q cavity and relatively wide angle between PA and MD (Figure 3.1b,c). It was shown that DDM inhibits *Ec*CI activity (Sazanov *et al*, 2003), but as we found out, at high DDM concentration the activity is almost absent (Figure 3.6). Thus, in order to achieve active preparation and catalytic turnover during the freezing at as high as possible activity rate we had to completely rethink the sample preparation conditions.

As a result of trials we found optimal conditions. First, instead of DDM as a main detergent we switched to milder LMNG. Secondly, we added *E. coli* total lipid extract (ETL) to the solution buffer. Thirdly, the blotting temperature was increased from 4 °C to 15 °C. Finally, carbon-coated grids were used to significantly decrease amount of protein and detergent in the solution. In order to control for DDM affect, we also purified *Ec*CI exclusively in LMNG without any DDM exposure. All these optimizations allowed us to obtain active sample and collect high resolution data.

3.2 Overall structure and specific features of the minimal complex I version

3.2.1 Peripheral arm peculiarities

High resolution allowed us to improve the previous crystal structure of the MA (Efremov & Sazanov, 2011), while the structures of the NuoH and the PA are new (Figure 3.2), solved independently before recently published structures (Kolata & Efremov, 2021). The general structure resembles the known core complex I architecture, with a few key differences. *Ec*CI has the largest NuoG subunit among known CI structures, with a long (~100 residues) insertion loop with a unique structure (Figure 3.2a). The loop stabilizes the complex by increasing the interaction surface area mainly with the NuoCD subunit, effectively replacing such accessory subunits as the Nqo16 from *Tt*CI or the mammalian 18 kDa subunit. Calcium is essential for *Ec*CI activity and stability (Sazanov *et al*, 2003) and NuoG subunit contains a Ca^{2+} ion bound at the site unique for *E. coli*, coordinated by acidic residues from the insertion loop (Figure 3.2a). NuoC and D subunits are fused in *E. coli* by a species-specific loop-helix-loop ($_{CD}$LHL) element, which interacts with NuoG, B and A subunits, stabilising the interface region (Figure 3.2a). Additional inter-subunit interactions are provided by the *E. coli*-specific

C-terminal extensions of about 20 to 40 residues in subunits NuoB, I, and F, stabilising the "back" of the PA (Figure 3.2b). It is likely that all these extensions are necessary to maintain the stability of the minimal version of complex I, as their sequences are conserved in Enterobacteria. The enzymes lacking such extensions instead contain additional subunits (*T. thermophilus* (Baradaran *et al*, 2013), *Paracoccus denitrificans* (Yip *et al*, 2011) and mitochondrial).

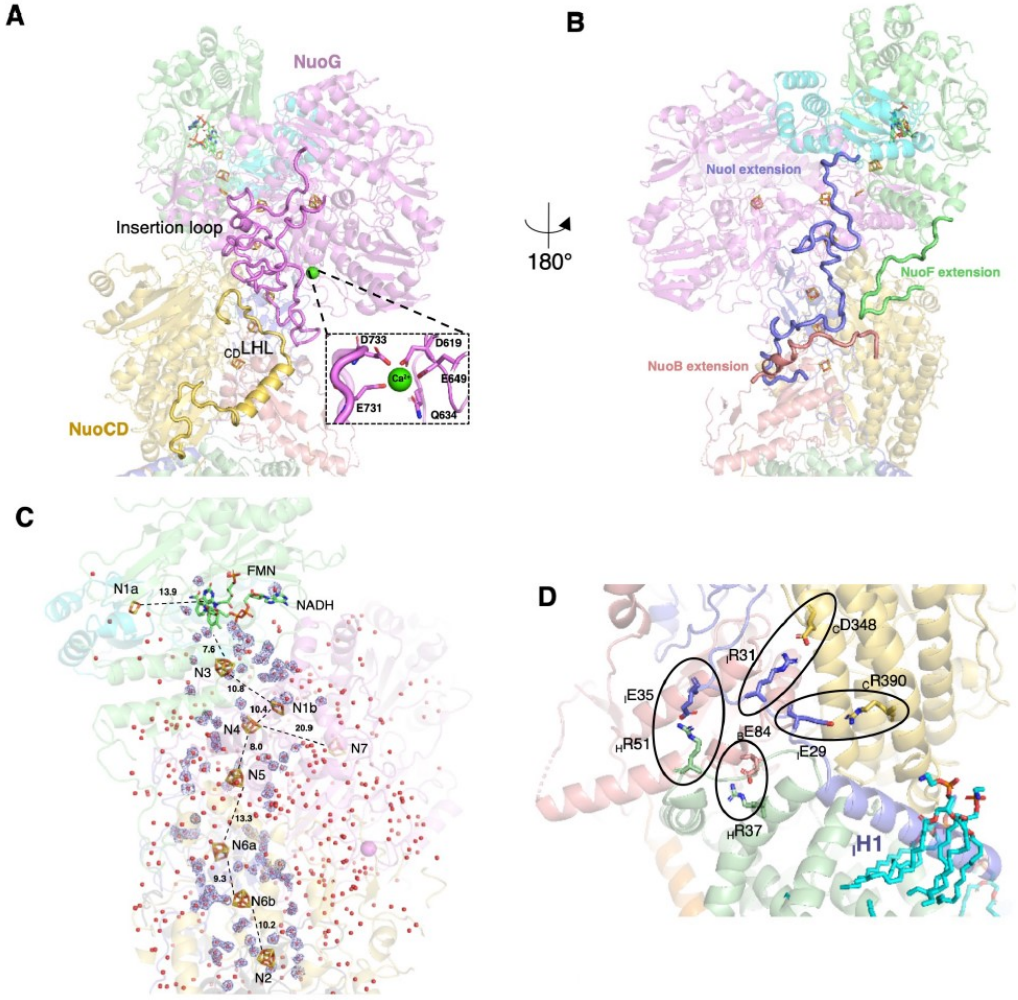

Figure 3.2 Features of the Peripheral Arm structure

A) *Novel elements stabilizing the PA. The NuoG insertion loop (magenta) interacts with NuoG and the rest of the complex, increasing the interaction surface area. Ca^{2+} binds at the interface of the insertion loop and core NuoG structure. Loop-helix-loop connecting element of NuoCD subunit ($_{CD}$LHL, yellow) is located on the surface and interacts with the NuoG insertion loop.* ***B)*** *Unique C-terminal extensions in EcCI increase surface interacting area and stabilize the minimal CI version.* ***C)*** *Position of Fe-S clusters along the peripheral arm. Electrons are transferred as a hydride from NADH to FMN then one by one via eight Fe-S clusters to quinone. The edge-to-edge distance (Å) between the clusters is indicated. N1a and N7 are off-path clusters. The density of the water molecules within 10 Å from Fe-S clusters is shown in blue, and all experimentally identified waters are shown as red spheres.* ***D)*** *Multiple salt bridges stabilize the Q cavity and PA-MA interface in open and closed states. NuoI*

helix H1 is surrounded on all sides by tightly bound lipid molecules, which strengthen its binding to the rest of the complex. This interaction is likely essential for the overall stability of the Q cavity. Figure adapted from (Kravchuk et al, 2022)

3.2.2 Unusual redox potential of N1a FeS cluster

In contrast to other species, in *Ec*CI the N1a cluster has an unusually high redox potential (~ -235 mV in *E. coli vs* < - 400 mV in *P. denitrificans* or bovine) and can be reduced by NADH (de Vries *et al*, 2015; Zu *et al*, 2002). The structure reveals that $_{E}$N142 forms a strong hydrogen bond to the N1a S atom. This polar residue is unique to *E. coli*, as it is replaced by a hydrophobic Met or Val in other species (Birrell *et al*, 2013). The potential of FeS cluster can be increased by the polar interactions, including solvent waters (Langen *et al*, 1992). One of the two waters resolved near the cluster N1a in *E. coli* forms a hydrogen bond to the Sγ atom of $_{E}$C97. This water is absent in ovine enzyme, while the second water, bonded to $_{E}$C133, is present (PDB 6zk9). Therefore, the higher N1a redox potential in *E. coli* can be attributed to the unique $_{E}$N142, consistent with the effects of mutations of this residue (Birrell *et al*, 2013; Holt *et al*, 2016), as well as to the additional water molecule.

Most other clusters in the redox chain lack waters in their immediate environment, both in *E. coli* and in ovine, and are mostly equipotential at about -250/-300 mV (Ohnishi, 1998; Euro *et al*, 2008b). The terminal cluster N2, which has the highest potential in the chain (~ -200 mV in *E. coli* (Euro *et al*, 2008b) and -140 mV in bovine (Ingledew & Ohnishi, 1980)), is a clear exception as it has two waters forming hydrogen bonds to the cluster. The waters' position is conserved between *E. coli* and ovine enzymes. Additionally, this cluster interacts with the conserved NuoCD residues R254, R274 and H359. This highly polar environment explains the high potential of the cluster. Overall, up to 860 waters are observed within the highly hydrated PA, and some waters closer to the core of the protein are conserved with the ovine enzyme, including the two waters situated roughly between the N3 and N4 (Figure 3.2c) but located too far from the clusters for any direct interactions.

Complex I is a major source of ROS with the fully reduced FMN being the main producer (Esterhazy *et al*, 2008). In the bovine enzyme, FMN predominantly donates an electron to O_2 forming superoxide, which is detoxified to H_2O_2 by superoxide dismutase (Grivennikova & Vinogradov, 2013). In contrast, *Ec*CI forms H_2O_2 directly (Esterházy *et al*, 2008), which is not due to the high N1a potential (Birrell *et al*, 2013). We compared the NADH-binding sites of *Tt*, *Aquifex aeolicus* (*Aa*), *Yarrowia lipolytica (Yl)* and mammalian NuoF structures with *Ec*CI. About 8 Å from FMN we identified an arginine residue ($_{F}$R320), which is not present in other structures (Figure 3.3c), being replaced by Gly in the mitochondrial and

by Met in the *Tt* enzymes. It is thus possible that the extra positive charge on $_F$R320 retains the nascent superoxide long enough for the spontaneous degradation into hydrogen peroxide to happen. Additionally, when NADH is bound, this arginine forms a tight hydrogen bond to the ribose moiety of NADH. This possibly changes the kinetics of NADH/NAD^+ binding/release and, as a consequence, ROS production. Mutagenesis on R320 would be necessary to verify these hypotheses.

Recently, from the analysis of structures of the *Aa*NuoFE subcomplex, it was suggested that the peptide bond between E95 and S96 ($_F$E93 and $_F$M94 in *Ec*CI) flips away from FMN in the reduced state and towards FMN in oxidized, as a mechanism to reduce ROS production (Schulte *et al*, 2019). This contradicts our high-resolution structures - the peptide is clearly always oriented away from FMN (Figure 3.3ab), similarly to *Oa*CI (Kampjut & Sazanov, 2020). The difference with *Aa* is likely due to the different properties of the protein and the FeS clusters in a two-subunit subcomplex as compared to the intact enzyme (Schulte *et al*, 2019; Peng *et al*, 2018).

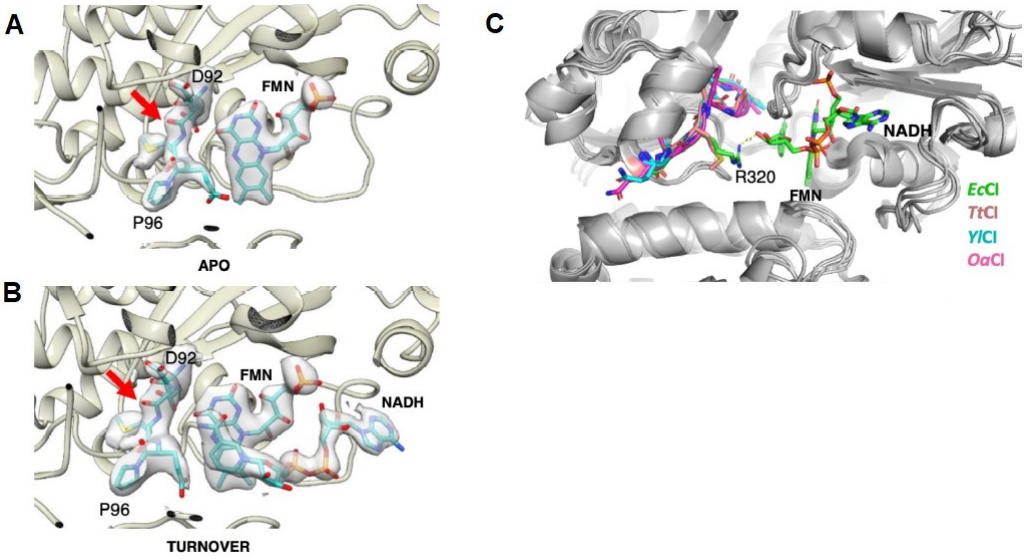

Figure 3.3 Environment of N1a cluster

A-B) The conformation of the backbone between $_F$D92 and $_F$P96 is unchanged in EcCI regardless of NADH presence. Red arrows point to the backbone oxygen atom, which was suggested to change conformation (peptide bond flip) in Aquifex aeolicus studies. C) Comparison of NuoF subunits from different CI species. EcCI contains unique $_F$R320, which points into the active site and interacts with NADH. Figure adapted from (Kravchuk et al, 2022)

When the *E. coli* (Holt *et al*, 2016) or bovine (Gostimskaya *et al*, 2007) enzyme is reduced by NADH in the absence of electron acceptors (DQ or FeCy), FMN reversibly dissociates from the complex, as a possible mechanism for the prevention of ROS production (Holt *et al*, 2016). The dissociation is prevented at high protein concentrations when protein-

bound FMN concentration exceeds the K_d for FMN dissociation (Holt *et al*, 2016). Our DDM datasets were obtained with high *Ec*CI concentration on a grid, 8-10 mg/ml, as compared to ~0.2 mg/ml in DDM/LMNG, due to the use of the carbon support in the latter case. Consistently, in the structure of the reduced enzyme in DDM (DDM_NADH) we did not observe FMN dissociation and saw clear NADH density alongside FMN (Figure 3.4a). However, in DDM/LMNG_PieA+NADH dataset substantial parts of NuoF and NuoE subunits (including FMN) were disordered and had a weak density (Figure 3.4c). In the DDM/LMNG_NADH+FMN dataset, external FMN was added, which according to the experimental data recovers the electron transfer activity (Holt *et al*, 2016). Indeed, we observed stronger FMN density, although NuoF and E subunits were still substantially disordered (Figure 3.4d). These data unequivocally confirm our original proposal that FMN dissociates under these conditions, in contrast to alternative explanations (Gnandt *et al*, 2017).

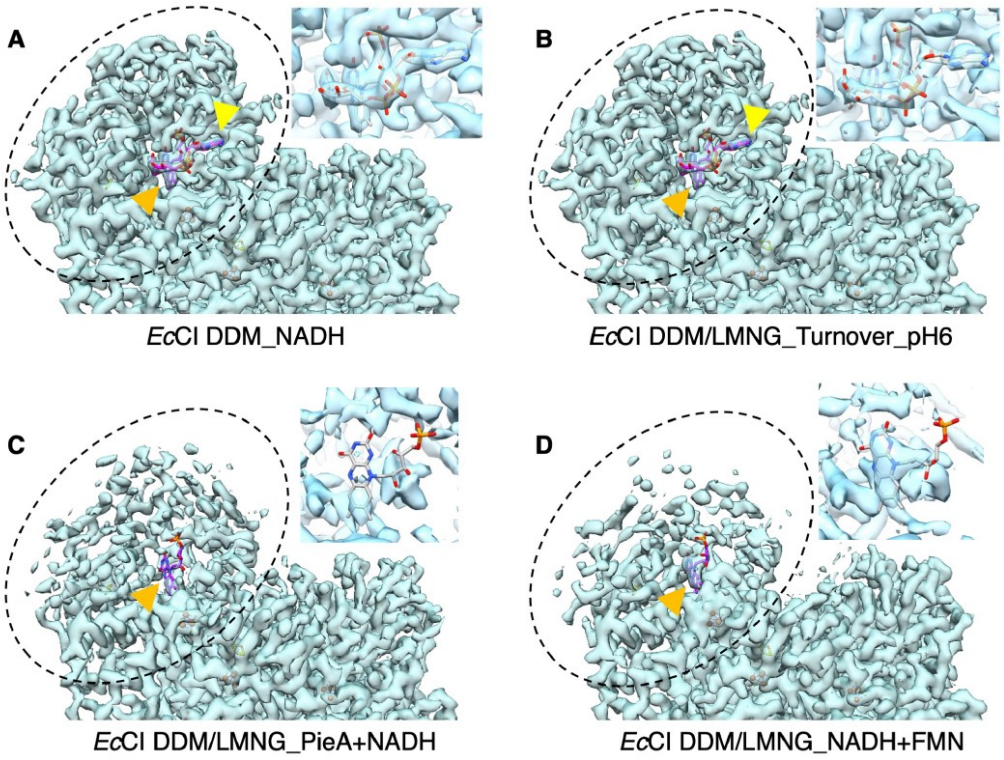

Figure 3.4 Cryo-EM density of the FMN binding site at different conditions in the presence of NADH

NuoFE subunits are highlighted with a dashed circle, FMN and NADH are indicated by the orange and yellow arrows, respectively. The concentration of CI applied to EM grids was ~10 mg/ml in condition A and ~0.2 mg/ml in conditions B-D. ***A)*** *When the concentration of the holoenzyme is high (above the K_d for FMN dissociation) NuoFE, FMN and NADH all have clear densities.* ***B)*** *The same is true when the protein concentration is low but a constant electron flow from FMN to DQ occurs during turnover.* ***C)*** *When the protein concentration is low and the complex is reduced without electron acceptor present, NuoFE subunits get disordered and most of FMN completely dissociates from the active site.* ***D)*** *However, FMN remains bound when external excess of FMN is added, even though NuoEF subunits still get disordered. Inserts show zoom-in into FMN/NADH (A, B) and FMN*

(C, D) density. Taken together, this data provides additional confirmation of the true turnover condition in B. Figure adapted from (Kravchuk et al, 2022)

Importantly, in the turnover dataset, FMN, NADH, NuoF and NuoE all have strong densities (Figure 3.4b), confirming that the constant electron transfer from NADH to DQ was happening at the time of flash-freezing (otherwise NuoF and NuoE would be disordered and FMN would dissociate as in Figure 3.4c).

3.3 Conformational dynamics of complex I

3.3.1 Three conformations of the *E.coli* complex I

DDM datasets may show the "resting" state, where activity is uncoupled from proton pumping (Belevich *et al*, 2017) due to disruption of Q cavity. The resting state has similarities to the deactive state of mammalian CI, but the disorder around Q cavity encompasses a larger area in *Ec*CI. However, in contrast to deactive *Ovis aries* complex I (*Oa*CI), $_{H}$TM5-6 loop is ordered in resting *Ec*CI, adopting "up" conformation (Figure 3.1c) and pushing NuoCD so that PA is shifted about 8 Å away from MA (Figure 3.1b).

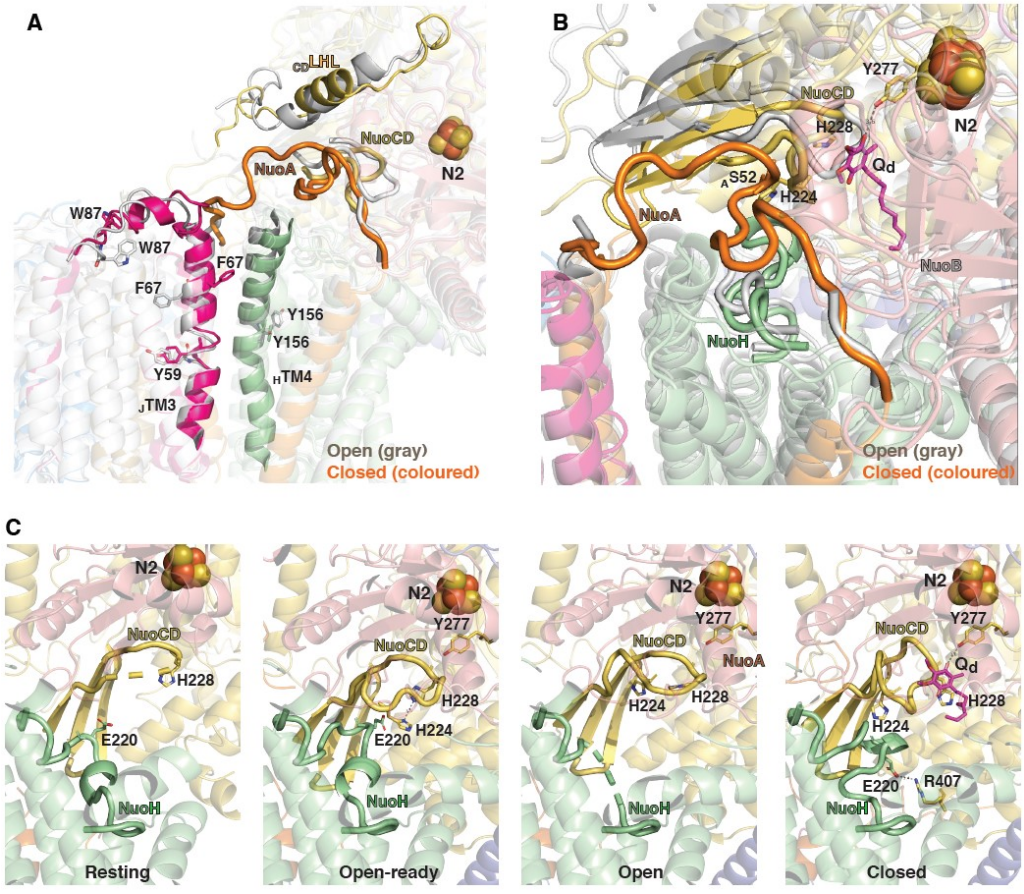

Figure 3.5 Conformational changes induced upon EcCI closing.

A) Global conformational changes upon open (grey) to closed (coloured) state transition. Key areas with changing conformation are indicated. NuoCD LHL element radically changes conformation, guiding NuoA loop to order and close the Q cavity. This process is assisted by the re-arrangement of NuoJ TM3-4 loop (note different W87 positions) and the rotation of $_{J}$TM3 and $_{H}$TM4 (note different $_{J}$F67 and $_{H}$Y156 positions). B) Conformational changes inside the Q cavity upon open to closed transition. NuoCD loop adopts retracted conformation, allowing quinone (magenta sticks) to bind in the Q_d site. Here quinone would clash with the extended NuoCD loop (grey), so the retraction is necessary. NuoH and NuoA loops get ordered, enclosing the Q cavity. C) Comparison of key NuoCD and NuoH loops in different states. In resting state NuoCD loop is disordered and NuoH is ordered in the "up" conformation resembling, but distinct from that in open-ready state. In open-ready state both loops are ordered such that conserved $_{H}$E220 can compensate for $_{C}$H224 charge, allowing key $_{CD}$H224 and $_{CD}$H228 to interact strongly, stabilizing extended conformation of NuoCD loop. In open state NuoCD loop keeps this conformation whilst NuoH loop is disordered, facilitating quinone movement. In closed state NuoH loop changes its conformation to "down" such that $_{H}$E220 flips to form a salt bridge with conserved $_{CD}$R407. This helps $_{CD}$H224 and $_{CD}$H228 to separate, allowing retracted conformation of NuoCD loop and quinone binding in Q_d site. Figure adapted from (Kravchuk et al, 2022)

As mentioned above, to obtain more native preparation, we diluted DDM-solubilised *Ec*CI stock in in a milder detergent LMNG, with added *E. coli* lipids, which showed optimal activity (Figure 3.6).

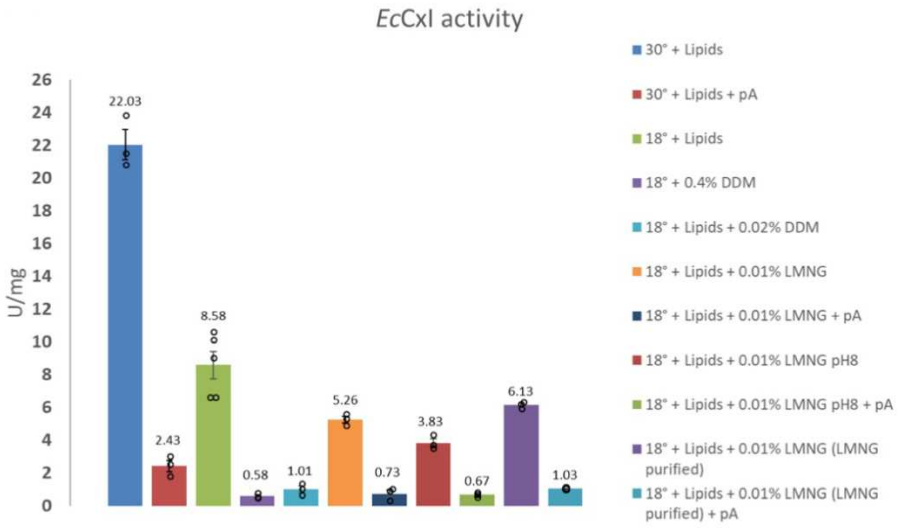

Figure 3.6 NADH:DQ oxidoreduction activity assays.

Results are represented in μmol NADH min^{-1} mg^{-1} protein, as the mean ± standard error with values for individual measurements shown as circles. E) EcCI. When present, lipids were added as 0.25 mg/ml ETL, and piericidin A (pA) inhibitor was added to 30 μmol. Figure adapted from (Kravchuk et al, 2022)

We collected six such DDM/LMNG datasets in order to identify any specific features induced only by the turnover (or by reduction/Q/inhibitor) (Table 1). In all DDM/LMNG datasets the "resting" and "open" (not reported previously) states were observed. The resting state was similar to the DDM structures, while in the "open" state, the PA has joined back to the MA and the Q cavity was mostly reformed. The key loops ($_{H}$TM5-6 and $_{A}$TM1-2) were disordered and $_{J}$TM3 had a π-bulge, hence we termed this an "open" *Ec*CI state (Figure 3.1d).

Exclusively under turnover (confirmed by several lines of experiments) we observed a third *Ec*CI state. We termed it "closed" as it contained an enclosed Q cavity with all the key loops ordered, and a rotated ⱼTM3 without π-bulge (Figure 3.1e). Notably, the fact that closed state was not observed in NADH, DQ or PieA *Ec*CI datasets suggests that neither reduction nor quinone/inhibitor binding alone can induce it, thus in *E. coli* the closed state is a higher energy intermediate and energy input during turnover is required for it to be observed.

Dataset (external ligands added)	**State** # particles (% total); resolution (Å), entire complex and PA/MA-focused refinement; the occupied Q binding sites (Qs, Qm or Qd).			
	Resting	Open-ready	Open	Closed
*Ec*CI DDM Apo (none)	121k (**100**%) 3.6 Entire 3.2 PA / 3.4 MA			
*Ec*CI DDM NADH (NADH)	145.2k (**100**%) 3.2 Entire 3.0 PA / 3.3 MA Qs			
*Ec*CI DDM/LMNG Apo (none)	199.3k (**54**%) 2.7 Entire 2.4 PA /2.6 MA		167k (**46**%) 2.4 Entire 2.2 PA / 2.4 MA Qm	
*Ec*CI DDM/LMNG DQ (DQ)	66.8k (**54**%) 3.1 Entire 2.6 PA / 2.8 MA		56.2k (**46**%) 2.7 Entire 2.5 PA / 2.8 MA Qm	
*Ec*CI DDM/LMNG NADH+FMN (NADH, FMN)	85.2k (**70**%) 4.2 – 5.0 Entire Poor map		36.8k (**30**%) 3.0 Entire 2.8 PA / 3.0 MA Qm	
*Ec*CI DDM/LMNG PieA (NADH, Piericidin A)	114k (**53**%) 4.1 – 4.2 Entire Poor map		99.9k (**47**%) 3.2 Entire 3.0 PA / 3.2 MA Qm	
*Ec*CI DDM/LMNG Turnover pH 6 (NADH, DQ)	92.9k (**47**%) 3.1 Entire 2.6 PA / 2.8 MA		97.9k (**49**%) 2.5 Entire 2.3 PA / 2.6 MA Qm	8.2k (**4**%) 3.4 Entire 3.2 PA / 3.3 MA Qd
*Ec*CI DDM/LMNG Turnover pH 8 (NADH, DQ)	82.6k (**37**%) 3.05 Entire 2.6 PA / 3.1 MA	40.8k (**18**%) 3.0 Entire 2.7 PA / 3.05 MA	67.2k (**30**%) 3.0 Entire 2.6 PA / 3.1 MA	32.6k (**15**%) 3.05 Entire 2.7 PA / 3.1 MA Qd, Qs
*Ec*CI LMNG Apo (none)	19.5k (**16**%) 4.7 Entire Poor map	108.9k (**60**%) 3.4 Entire 3.2 NuoFEG 3.3 NuoAJKHCD 3.4 NuoLMN	43.2k (**24**%) 4.6 Entire Poor map	
*Ec*CI LMNG Turnover pH 6 (NADH, DQ)	189.7k (**27**%) 2.7 Entire 2.4 PA / 2.9 MA	317.1k (**45**%) 2.35 Entire 2.15 PA / 2.3 MA	31.8k (**4**%) 3.1 Entire 2.9 PA / 3.2 MA	170k (**24**%) 2.5 Entire 2.3 PA / 2.5 MA Qd, Qs

Table 1. Summary of datasets

In contrast to the mammalian enzyme, where the open state reflects the larger angle between the PA and MA due to the PA tilt, in *Ec*CI open and closed states differ mainly by the rotation of the PA (Figure 3.1b). This suggests that similar conformational changes around the Q site do not necessarily lead to similar PA movements in different species. The "open" and "closed" states terms can still be applied for *Ec*CI and other species, referring mainly to the open and closed, respectively, Q cavity.

To verify whether the resting state could be a DDM-induced artefact, we solubilised and purified *Ec*CI entirely in LMNG. In apo and turnover datasets the proportion of the resting

state dropped but remained significant (Table 1). Its structure was unchanged, suggesting that the resting state is not an artefact of DDM exposure, but can be partially (DDM/LMNG datasets) or strongly (DDM datasets) promoted by DDM and associated de-lipidation. LMNG turnover structures also revealed a higher proportion of closed state than in DDM/LMNG, suggesting overall stabilisation of *Ec*CI in milder detergent. As all non-mammalian species studied so far show only open state in the absence of turnover, the appearance of closed *Ec*CI state only under turnover, in three independent datasets, is a definite proof that closed/open states are true catalytic intermediates. High similarity of open-to-closed transition between *Ec*CI and *Oa*CI also confirms that open/closed states of mammalian enzyme should be considered as catalytic intermediates.

3.3.2 Quinone binding cavity

Quinone binds within an elongated cavity at the PA/MA interface. CI-bound quinone previously was observed at the deep (Q_d) and shallow (Q_s) sites within the cavity, ~12 Å and ~24 Å from cluster N2, respectively (Baradaran *et al*, 2013; Parey *et al*, 2021). In *Oa*CI both sites were occupied in the closed state under turnover (Kampjut & Sazanov, 2020), which is possible with short-tailed DQ, as the native quinone occupies the entire length of the cavity (Gu *et al*, 2022). Notably, in the open state *Ec*CI (DDM/LMNG datasets) the only observed quinone was bound in the additional site ~ 16 Å from N2, with headgroup interacting with $_{CD}$Q328 and stacked against $_{B}$L86 (Figure 3.7b). DQ, native UQ8 or piericidin A were bound in the same site, which we termed median, or Q_m. The Q_d binding is prevented in all *Ec*CI open state structures by the extended conformation of the NuoCD β1-β2 loop, blocking the deep end of the cavity (Figure 3.7a). In *Oa*CI open states this loop is disordered, with extended conformation observed only in the NADH-reduced state (Kampjut & Sazanov, 2020). An extended loop is observed in the NDH complex (Pan *et al*, 2020), with plastoquinone (PQ) bound in Q_m position (Figure 3.7c). A conserved NuoB alanine is replaced by $_{B}$Y65 in *E. coli* or by F54 in NDH, facing the cavity, which may create a bottleneck responsible for Q_m site, not present in other species (in open states of *Oa*CI site Q_s is occupied). This may explain why *Ec*CI shows lower affinities to many inhibitors than the mitochondrial enzyme (Friedrich *et al*, 1994).

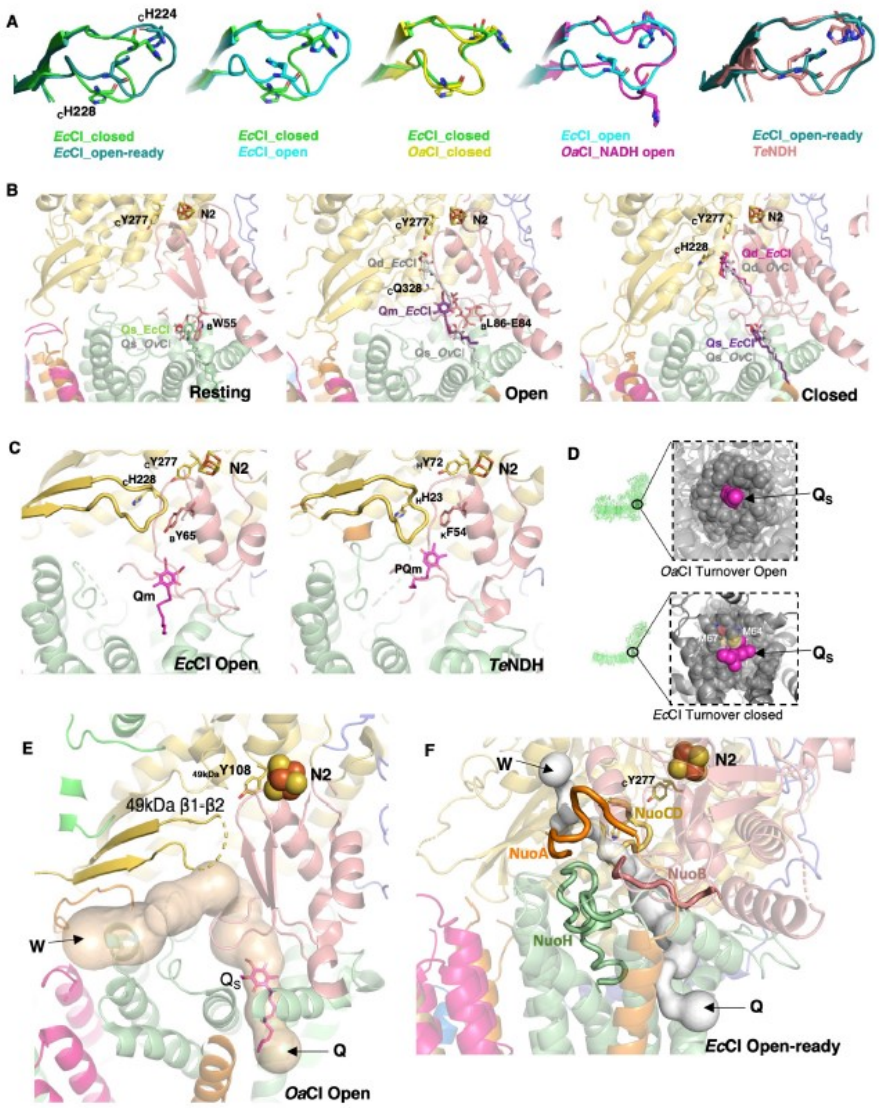

Figure 3.7 Quinone-binding site comparisons

A) *Comparison of key NuoCD loop in different CIs. Ovine CI (OaCI) is from PDBs 6ZKC for closed and 6ZKE for open state. TeNDH (Thermosynechococcus elongatus NDH complex) is from PDB 6NBY. Key conserved histidines are shown as sticks.* ***B)*** *Q binding sites. Left: in the resting state Q binds at the entrance to the Q cavity, consistent with the mammalian Q_s binding site. Middle: in the open state Q binds in the Q_m site, in between Q_d and Q_s sites (Q_d and Q_s quinones from the aligned structures are shown for comparison). Right: in the closed state DQ binds deep inside of the cavity, consistent with the mammalian Q_d binding site, and also in the Q_s site (LMNG datasets). Key residues interacting with quinone headgroup in each site are indicated. Quinone molecules from the aligned OaCI structures are shown as grey sticks. Q_d site is narrow with a tight Q coordination, while Q_s is looser, with some variability in the mode of binding.* ***C)*** *Extended NuoCD loop and side-chains of NuoB Helix3 block access to the Q_d binding site in EcCI open state (left) and in TeNDH (PDB ID 6KHJ) (right), with plastoquinone (PQ) bound in the same site as Q_m in E. coli.* ***D)*** *Q hydrophobic tail seals the Q entrance. Top: OaCI, DQ bound in the Q_s site (PDB 6ZKE) is depicted as magenta spheres and protein atoms within 8 Å as transparent gray spheres. Bottom: EcCI, model of UQ8 fitted into Q_d site of the closed state structure is depicted as magenta spheres and protein atoms within 8 Å as transparent grey spheres, except for $_H$M64 and $_H$M67*

(yellow), framing the entry. ***E)*** *Q cavity in mammalian open complex I (PDB 6ZKE) is exposed to the matrix via W site, consistent with EcCI.* ***F)*** *In the open-ready state of EcCI, although NuoA loop is partly ordered, the Q cavity is still exposed to the matrix via W site. . Figure adapted from* (Kravchuk *et al*, 2022)

In the closed *Ec*CI under turnover, NuoCD loop is retracted, allowing for DQ to bind in Q_d site. A second DQ molecule is observed close to Q_s site, probably shifted from Q_m by the tail of Q_d-bound DQ (Figure 3.7b). At the Q_d site DQ accepts electrons from cluster N2, forms H-bond with the conserved $_{CD}$Y277 and stacks against $_{CD}$H228 from NuoCD loop, similar to *Oa*CI (Kampjut & Sazanov, 2020) and *Tt*CI (Baradaran *et al*, 2013). Along with NuoCD loop retraction, open-to-closed state transition includes ordering of $_{H}$TM5-6 and $_{A}$TM1-2 loops, rotation of $_{J}$TM3 and $_{H}$TM4 with flip of Y156 (Figure 3.5ab), and tilting of NuoH helices. These features faithfully reproduce the open-to-closed state transition in *Oa*CI (Kampjut & Sazanov, 2020).

*Ec*CI shows additional re-arrangements not observed in *Oa*CI (Figure 3.5a). $_{J}$TM3-4 loop completely changes conformation, with W87 flip, while in *Oa*CI it was disordered in the open state. $_{CD}$LHL element, linking NuoC and NuoD, thus absent in other species, is also almost completely rebuilt, resulting in the rotation and shift of its helix. Together these re-arrangements help to "push in" the NuoA loop into the crevice between NuoCD and NuoB in the closed state, ordering the loop and closing Q cavity. The conformation of the flexible $_{H}$TM5-6 loop, containing many conserved charged residues, is exactly the same in the closed states of *Ec*CI and *Oa*CI, consistent with its essential mechanistic role.

The open state, uniquely so far to *E. coli*, could be separated by 3D classification into two states - one "open" (described above) and another we termed "open-ready", which likely represents an additional, previously not resolved, intermediate in the catalytic cycle. The "open-ready" state differs from "open" mainly by conformations of $_{H}$TM5-6 and NuoCD loops. The $_{H}$TM5-6 loop, disordered in open state, adopts ordered conformation differing from closed state by the reorganisation bringing invariant $_{H}$E220 close to $_{CD}$H224. These concerted loop movements suggest that open-ready state may represent the enzyme ready to bind external Q (since no Q is observed in the cavity), while the open state represents the stage when Q has entered and bound in the Q_m site (observed in open states from most datasets), followed by transition to the closed state.

3.4 Proton translocation pathways

High resolution of the structures allowed us to identify a large number of water molecules (~860 in PA and ~430 in MA) and to reveal for the first time the MA hydration pattern in the bacterial CI. Apart from the protein surface, waters clustered around the central axis, connecting the key charged residues (Figure 3.8a). The MA structures for resting, open and open-ready states did not show significant differences in the overall structure, therefore we discuss the highest resolution (2.3 Å in MA) open-ready LMNG turnover structure and compare it to the closed LMNG turnover state (2.5 Å in MA).

Analysis of proton translocation pathways affirms (Kampjut & Sazanov, 2020) that the E-channel and NuoN/M ALS lack connections to the periplasm, blocked by large hydrophobic residues. Subunit NuoL is unique in having a highly hydrated exit to the periplasm, connecting LysTM12 (K399) to $_{L}$D400 and polar residues nearby. We also do not observe any conformational changes in ALS between open/closed turnover states. Therefore, in all states of the complex, proton exit pathway into the periplasm is formed only in the distal ALS NuoL. This counter-intuitive feature is thus conserved from bacterial to mitochondrial (Kampjut & Sazanov, 2020; Parey *et al*, 2021) enzyme.

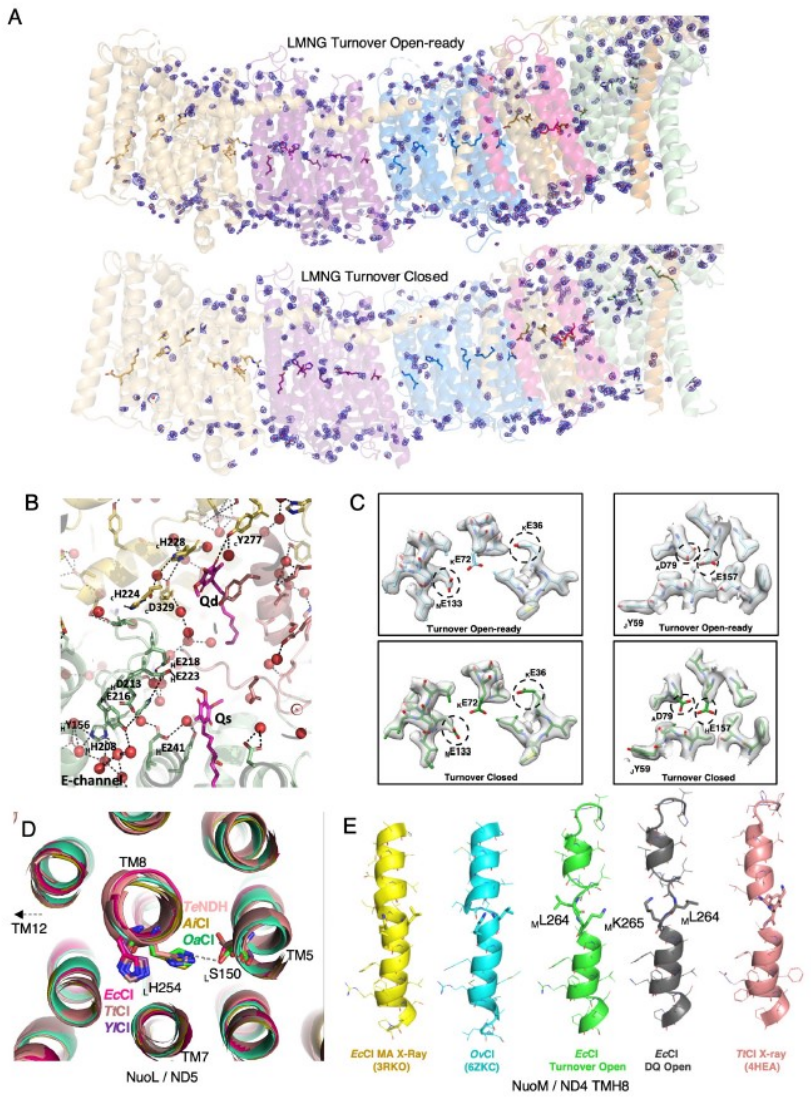

Figure 3.8 Waters and proton translocation pathways

A) *Cryo-EM densities for the experimental waters in the MA-focus-refined maps of EcCI LMNG_Turnover_pH6 open-ready (top) and closed (bottom) states. To allow clear visualization, the density is carved around modelled waters (red spheres) and is shown in light blue. The model is coloured by subunit as in Fig. 3.1a. Key residues from the central hydrophilic axis of EcCI are shown as sticks.* ***B)*** *A putative proton transfer pathway between the E-channel and the key* $_{C}$*H228/*$_{C}$*D329 residues, likely proton donors for quinone. Key protonatable residues, experimentally resolved waters and quinones from LMNG_Turnover_pH6 closed state are shown. Potential H-bonds are indicated by black dashes.* ***C)*** *Detailed analysis of cryo-EM density reveals charge of Glu and Asp residues in MA. Carboxyl side-chain densities of some key residues are absent (circled) in the closed state, suggesting their negative charge. In contrast, the same residues in the open state preserve densities suggesting their neutral charge.* ***D)*** *Comparison of NuoL TM8 helices from different CI species. Structures were aligned on EcCI NuoL subunit. Key* $_{L}$*H254 residue and* $_{L}$*S150 with which it can interact are shown as sticks. Due to flexibility of TMH8 key histidine can be preferentially linked either to key TM12 residue as in EcCI, TtCI and YlCI, or to key TM7 residue and the rest of the central axis as in OaCI, TeNDH (PDB 6KHJ) and AiCI (Arabidopsis italiana mitochondrial CI, PDB 7AR8).* ***E)*** *Comparison of NuoM TM8 helices from different CI species. In EcCI* $_{M}$*TM8 is*

flexible and adopts different conformations. It is "linked" (green) in DDM/LMNG datasets PieA, Apo, Turnover in open states, and in resting states in Apo, Turnover and DDM_NADH. In DDM/LMNG datasets NADH+FMN and DQ in the open states, and resting states in DQ and DDM_Apo it is "flipped" (grey). However, both of these conformations are consistent with other CI structures as shown. . Figure adapted from (Kravchuk *et al*, 2022)

Following proton pathways along the central axis we see a highly hydrated connection all the way from key $_{L}$K399 to $_{N}$E133 via repeating Lys(Glu)TM12-Lys(His)TM8-LysTM7-GluTM5 series, linked by additional charged residues (labelled on Figure 3.9a). Only NuoL and NuoM have a link to the cytoplasm via branching residues sitting on broken TM8. These residues ($_{L}$H254 and $_{M}$K265) could be able to switch the conformation (not linked to the redox state) and thus help to re-distribute incoming protons along the central axis. The connection continues till the essential (Efremov & Sazanov, 2011) NuoK residues: $_{N}$GluTM5(E133)-$_{K}$E72-$_{K}$E36. Then in all *Ec*CI open states there is a long (~13 Å) break in hydration, with hydrophobic residues from $_{J}$TM3 blocking the path from $_{K}$E36 to $_{A}$D79. However, in the closed state the cytoplasmic half of the $_{J}$TM3 rotates, π-bulge disappears and small residues G61-A62 replace A62-I63, which opens the cavity for waters to fill in the break (Figure 3.9bc), creating the connection to $_{A}$D79. $_{H}$TM4 helix also rotates, so that the invariant Y156, sitting on another π-bulge, flips over ~180^{o} from facing the lipids in the open state directly into the created water path, helping to establish a firm Grotthuss connection (similarly to Y142 in *Oa*CI (Kampjut & Sazanov, 2020)). The conformation of this tyrosine is a distinct feature helping to easily recognise open or closed state (Figs. 3.5a and 3.9bc). From $_{A}$D79 a pathway continues via NuoH residues E157-Y156-H208-E216-D213-E218 and some ordered waters in the Q cavity (Figure 3.8b and 3.9a) towards the $_{CD}$D329/H228 pair, a likely source of two substrate protons for quinone (Kampjut & Sazanov, 2020).

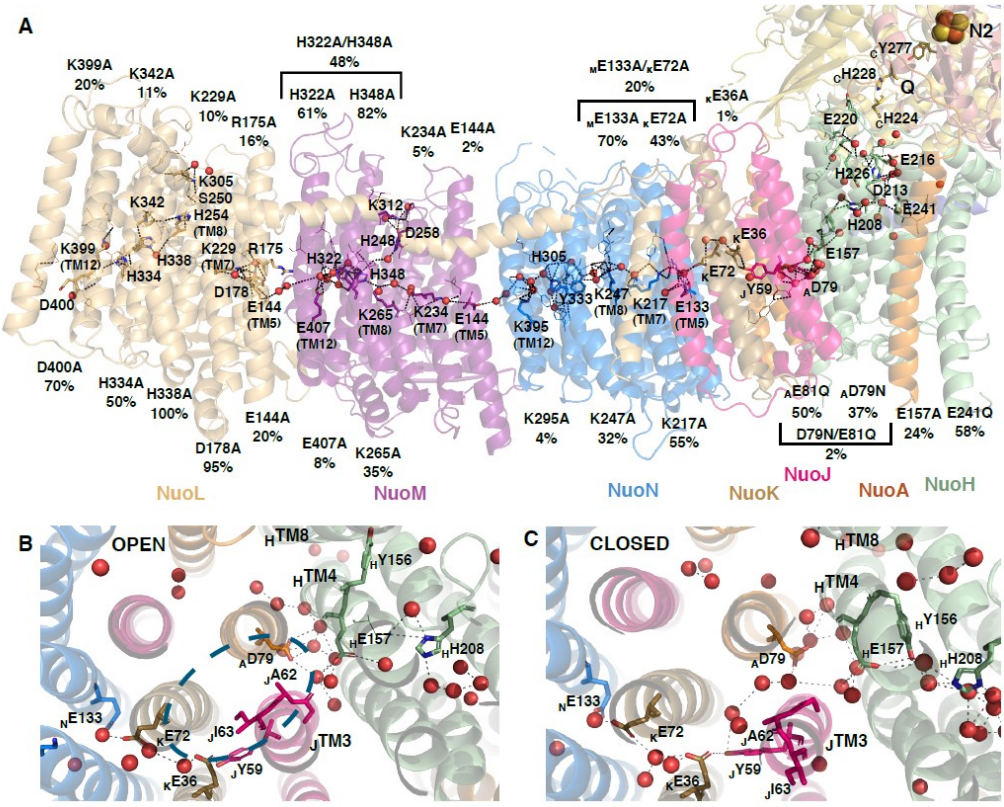

Figure 3.9 Proton translocation pathways.

A) Membrane arm contains the central axis of charged residues, essential for the proton transfer and the coupling. Structure of the LMNG turnover open-ready state is shown coloured by subunit, with essential residues shown as sticks. Key ALS residues are also identified by their TM helix. Experimentally observed waters are shown as red spheres (waters beyond 5 Å from essential residues are omitted for clarity). Putative proton pathways through Grotthus-competent residues (shown as lines unless key residue) and waters are shown as black dashes. The activity of mutant variants of EcCI is shown as NADH:DQ oxidoreduction in % of WT activity. B) Left: In all open, open-ready and resting states $_{J}$TM3 I63-A62 residues impose a hydrophobic block between $_{K}$E36 and $_{A}$D79. Right: In the closed state $_{J}$TM3 rotates anticlockwise, which removes the hydrophobic block and allows waters to come in and connect $_{K}$E36 and $_{A}$D79. $_{H}$Y156 also comes in helping to establish robust connection further in the E-channel. Structures are of LMNG_Tunover_pH6 open-ready (left) and closed (right) states with experimentally observed waters (red spheres). Figure adapted from (Kravchuk *et al*, 2022).

Since we did not observe any conformational changes within ALS under turnover conditions, electrostatics are likely to drive proton transfer within these subunits. Consistently, judging from density for carboxylate side-chains (almost disappearing in charged state), $_{H}$E157, $_{A}$D79, $_{K}$E36 and $_{N}$E133 were all neutral in open state but charged in closed turnover state (Figure 3.8c), as also observed in *Oa*CI (Kampjut & Sazanov, 2020).

Because key TM12 residue in NuoM is a glutamate (E407) instead of lysine, we suggested previously that NuoM might pump protons in anti-phase with NuoL/N (Kampjut & Sazanov, 2020). To test this hypothesis $_{M}$E407 was mutated to lysine and it was found that *Ec*CI was still active, which suggests that in fact all three ALS act in a similar fashion (Kravchuk *et al*, 2022). Residues around the main Q entry site were also mutated (Kravchuk *et*

al, 2022), confirming that it is the only entry used by quinone, refuting recent proposals that Q may also enter from the cytosol (Masuya *et al*, 2021).

3.5 Open-to-closed state transition

All the hallmarks of the open-to-closed state transition are conserved between the evolutionarily distant bacterial and mammalian enzymes (Kampjut & Sazanov, 2020), suggesting that they are the key to the universal mechanism of complex I. In the open state, due to disorder of key loops, the Q site is open to the cytoplasm via the wide opening of the cavity (or branching tunnel) emanating roughly from the Q_m site, both in bacterial and mammalian enzymes (W site in Figure 3.1d). As a key novel feature of the mechanism that explains experimental observations, we propose that this additional W site is essential for exit or entry of water molecules accompanying the entry or exit, respectively, of quinone via the main entry point from the lipid bilayer (Q in Figure 3.1de). Before quinone entry the cavity is pre-filled with waters, therefore, without the extra "hole" (W) it will be difficult for quinone to get an access into the cavity, since its tail would be blocking the Q entry like a cork in the bottle (Figure 3.4d). Similarly, when quinol exits the cavity, because of the blocking tail the waters must come in through the W site to fill in the vacated space within the cavity. In stark contrast, in the closed state the key loops get ordered, sealing off the cavity and tightly engulfing the bound quinone (Figure 3.1e and 3.7d), so that waters cannot get into the cavity and protons for quinone protonation have to come from the central MA axis. This process would be optimal with the native long-tailed quinone but will also work with DQ, as the limited number of waters in the cavity (sealed by the lipids at the Q entry) will not be able to provide two protons.

Another defining feature of the open-to-closed transition is $_J$TM3 rotation. It is likely caused by a combination of tilting of NuoH TM helices, freeing up space for rotation, and a large shift of NuoCD β-sheet with β1-β2 loop (Figure 3.5b). This sheet pulls along with it the tightly interacting $_A$TM1-2 and $_J$TM3-4 loops, which probably causes re-winding of $_J$TM3. Importantly, the pattern of two small residues followed by a large hydrophobic residue in NuoJ (Gly61-Ala62-Ile63) is fully conserved in all species. The π-bulges on $_H$TM4 and neighbouring $_H$TM8, allowing $_H$Y156 flipping, are also conserved (Figure 3.9bc). As this pattern is responsible for the creation of the water wire, this suggests that the transfer of "charge action" of quinone oxidoreduction towards ALS is conserved.

Since key glutamates in the E-channel are unprotonated in the closed state, the proportion of closed state could be expected to increase with increasing pH. Therefore we

collected *Ec*CI turnover dataset at pH 8, for comparison with the initial pH 6 data (Table 1). Strikingly, the proportion of closed state indeed increased dramatically. Importantly, the activity of *Ec*CI and *Oa*CI has actually decreased with pH (Figure 3.6), firmly establishing that closed state is not equivalent to active state (since its proportion does not follow the activity), but instead is a part of catalytic cycle along with open state. Same phenomenon was observed for the mammalian enzyme (Kravchuk *et al*, 2022). The pattern of changes in Table 1 suggests that the pKa of key residues involved in open-to-closed transition is probably close to 8. The pH in mitochondrial matrix is about 8.0, and in *E. coli* cytoplasm about 7.6-7.8 (Zilberstein *et al*, 1984), therefore the increased proportion of closed state at higher pH likely reflects on the *in vivo* situation.

3.6 A universal coupling mechanism

On the basis of a compendium of our bacterial and mammalian enzyme structures and mutagenesis data (Appendix 5) we propose the universal, applicable to all species, "domino effect" mechanism of complex I, depicted in Figure 3.10. In essence, quinone/quinol binding and release happen in the open state (Steps 1, 4-5), enabled by waters coming via site W. The cycle starts with Step 1, where quinone binds and initiates the transition to the closed state (Step 2). Quinone is reduced and two protons are taken from the central MA axis to complete the reaction, which results in the re-distribution of protons, so that key TM12 residues are protonated and the charge is switched between TM5 and TM7 residues (Step 3). This is a highly energised state, akin to stacked dominos ready to fall. In the transition to open state, TM8 residues are protonated from the cytoplasm (Step 4), and $_{L}$TM12 proton is ejected into the periplasm due to electrostatic interactions. This initiates a series of proton transfers along the central axis due to appearance of a "vacancy" on the "left" of the chain and the electrostatic "pressure" of the incoming proton from the "right" (Step 5), akin to stacked dominoes falling. This results in four protons in total ejected from NuoL and the cycle re-starting.

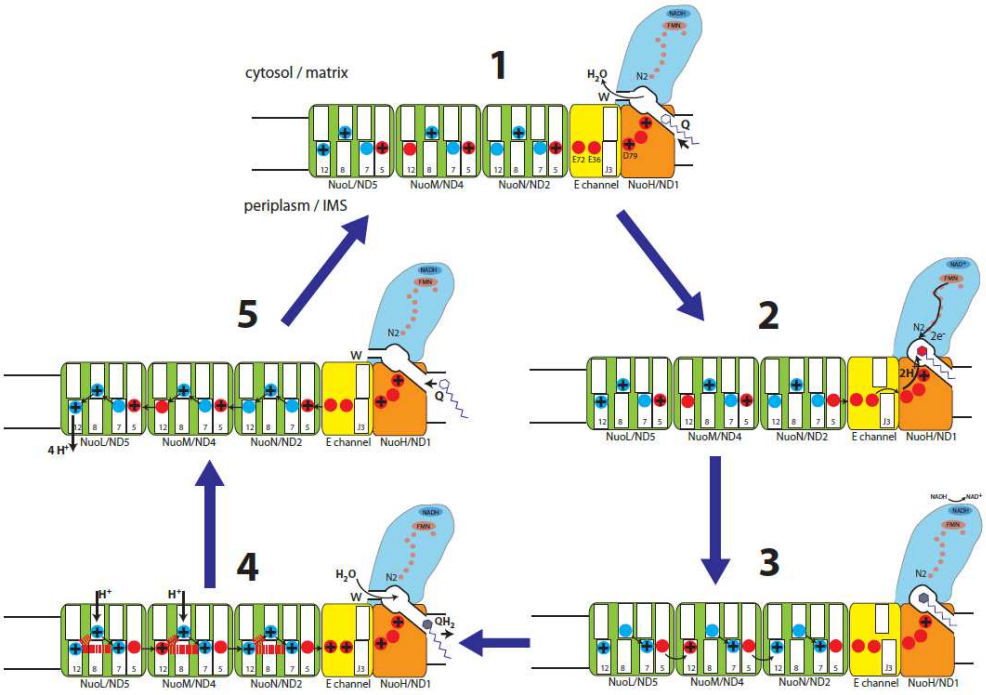

Figure 3.10 A "domino effect" coupling mechanism of complex I.

An overview of the proposed complex I mechanism. Individual steps involve conformational changes around Q cavity / E-channel and electrostatic interactions in antiporters NuoL/M/N, as described in the text. Complex I cycles between the open state (Steps 1, 4 and 5), where the Q cavity is widened and opened both to the lipid bilayer (Q) and to the cytosol (W), and the closed state (Steps 2 and 3), where the Q cavity is enclosed and tightly engulfs the bound quinone. NADH oxidation and electron transfer in PA in Step 2 are fast and not rate-limiting. Charged quinone intermediate is indicated by the red headgroup, and quinol by the grey-filled headgroup. The charged residues on the MA central axis are indicated in blue for lysines and in red for glutamates/aspartates. For clarity, the protonated forms are shown with a + sign and un-protonated are empty (although the actual charge would be +/0 for lysine and 0/- for glutamate/aspartate). The key helices in antiporters are indicated by their numbers. In the open state the water wire between the Q cavity and the central axis in the E-channel is broken at ȷTM3 (indicated as J3). The connection is established in the closed state due to ȷTM3 rotation. Black arrows indicate proton transfer, including re-distribution along the central axis. Access from the cytosol happens only via NuoL/M and the exit into the periplasm only via NuoL. Electrostatic interactions, resulting in the ejection of four protons into periplasm in Step 5, are indicated as red dashes in Step 4. Figure adapted from (Kravchuk *et al*, 2022).

The detailed sequence of events in the proposed mechanism is as follows (Figure 3.10). **(Step 1)** Much of the time the enzyme spends in the open (or open-ready) state, waiting for quinone from the membrane pool to enter and temporarily bind in the Q_m (bacterial) or Q_s (mammalian) site near the entrance. Site W is open for waters to flow out of the cavity and give space to the incoming quinone, while the Q cavity is expanded to allow for the unimpeded quinone entry. We propose that in this state the ALS are maximally protonated at the key *TM8*

sites (Lys/HisTM8) by protons coming into NuoL/NuoM from the cytoplasm (or mitochondrial matrix) and redistributed along the central axis into NuoN and part of the E-channel harbouring $_{K}$E72 and $_{K}$E36 (for simplicity we indicate any protonated residues just by a "+" sign). The rest of the E-channel is disconnected from this chain at $_{J}$TM3, preventing proton leak into the Q cavity and back to the cytosol. This may be important because protons are quite scarce in the cytosol. For the closely interacting *TM7/TM5* sites (LysTM7/GluTM5), the proton is proposed to reside in *TM5* site, as suggested by our observations in *Ec*CI (Figure 3.8c) and *Oa*CI (if only the density of glutamates is considered (Kampjut & Sazanov, 2020)). The *TM12* (Lys/GluTM12) sites are proposed to have a lower *pK*a and tuned to remain unprotonated in this state due to electrostatic interactions with protonated *TM8* and *TM5* sites. The exception is NuoL *TM12* site, proposed to be protonated as it is distal and so does not have a *TM5* partner from a neighbouring subunit. **(Step 2)** Bound quinone traverses into the Q_d site, triggering the open-to-closed transition, so that the W site is closed off and the Q cavity tightly engulfs quinone. $_{J}$TM3 rotates, establishing the uninterrupted proton path from the Q cavity all the way to MA tip. Quinone accepts two electrons from cluster N2 in quick succession (Verkhovskaya *et al*, 2008), and the unstable charged quinone intermediate is immediately protonated by the coordinating $_{CD}$H228/$_{CD}$D329, creating a double negative charge in the area. Since the Q cavity is sealed, the protons for the re-protonation of $_{CD}$H228/$_{CD}$D329 come from the central axis. The available redox energy from quinone chemistry appears to be sufficient to displace at least four ($_{H}$E157, $_{A}$D79, $_{K}$E36 and $_{N}$E133 (Figure 3.8c)) protons along the central axis, as they are not transferred against the pmf. Two of these are substrate protons and the rest may be shifted towards the Q site to enhance the charge action signal near NuoN. **(Step 3)** In a "minimal" interpretation (Occam's razor) of the subsequent events, de-protonation of the E-channel residues first triggers proton transfer from *TM8* to *TM7* site in NuoN, due to the removal of large positive charge around *TM5* area. In a series of "domino effect" events, the removal of $_{N}$*TM8* charge allows $_{M}$*TM5* proton to hop on $_{N}$*TM12* site, repeated in NuoM/L by $_{M}$*TM8* to $_{M}$*TM7*, $_{L}$*TM5* to $_{M}$*TM12* and $_{L}$*TM8* to $_{L}$*TM7* hops. De-wetting of the *TM8* area due to de-protonation, as observed in MD simulations (Mühlbauer *et al*, 2020), would prevent the back-flow of protons. Effectively, due to the "forcefully" protonated *TM12* sites and a shift of proton from *TM5* to *TM7* sites the enzyme will now be in a highly energised state, akin to a loaded spring or stacked dominos ready to fall. **(Step 4)** The presence of the freshly produced quinol in the Q_d site along with the re-protonated state of coordinating residues triggers the transition from the closed to the open state, so that the Q cavity widens and the W site opens, allowing waters to come in and help quinol on its way out. The *TM8* sites (and $_{K}$E72/E36) can be fully

protonated from the cytosol, blocked off from the Q cavity by $_{J}$TM3. In total at this stage six protons (four to be pumped and two substrate) will enter the central axis. Five of them can enter via NuoL/M, re-distributed along the central axis, while $_{H}$D79 can be protonated via open Q cavity. In this scenario the re-protonation of *TM8* sites (and $_{K}$E72/E36) is rather "passive", the key to coupling being that *TM12* and *TM7/TM5* sites state is fully controlled by quinone reactions. **(Step 5)** Electrostatic interactions with the protonated *TM8* and *TM7* sites (red dashes in Step 4) lead to a large decrease of *pK*a's of *TM12* residues, forcing them to lose their protons. In NuoL the *TM12* proton would be ejected directly into the periplasm (or mitochondrial IMS). In the reverse wave of the "domino effect" (exact sequence of events is given at the end of this paragraph) this will initiate a sequence of proton hops from $_{L}$*TM8* to $_{L}$*TM12*, $_{L}$*TM7* to $_{L}$*TM8* and $_{M}$*TM12* to $_{L}$*TM5*, repeating twice more in NuoM/N and ending with $_{K}$E72 donating proton to $_{N}$*TM5*. The simple natural basis for this transfer of protons along the central axis is the appearance of a "vacancy" on the "left" of the chain and the electrostatic "pressure" of the incoming proton from the "right" (or reverse in Step 3). Effectively, after the cycle is repeated three times, each time ending closer to NuoL, in the end the $_{M}$*TM12*, $_{N}$*TM12* and $_{K}$E72 sites would transfer their protons along the central axis towards NuoL, adding up with the initial $_{L}$*TM12* to the four protons ejected into the periplasm. This brings the system back to Step 1, with *TM8* protonated and a proton in *TM7/TM5* "switch" sitting again on *TM5*, thus the cycle re-starts. Crucially, for the mechanism to work, a *TM12* proton from NuoN/M must be transferred to the neighbouring NuoM/L *TM5* and not directly to the periplasm, as otherwise the process will not be initiated in the next subunit (i.e. a domino will fall without tripping the next one). Similarly, in Step 3 it is essential for protons to hop across subunit interfaces from *TM5* to *TM12* sites. Therefore, our mechanism naturally explains initially counter-intuitive NuoL-only exit. The pump works with protons moving along the entire central axis either towards Q cavity (Steps 2-3) or in the reverse wave (Step 5), thus the periplasm/IMS side must be shielded from the solvent everywhere except the NuoL exit.

After a prolonged absence of turnover the enzyme enters a resting (bacterial) or a deactive (mammalian) state, which may help to prevent ROS production, occurring *via* reverse electron transport in specific conditions, such as ischaemia-reperfusion injury (Chouchani *et al*, 2014). When turnover resumes the enzyme reverts back to the main cycle.

The mechanism also explains the reverse electron transport in complex I: high pmf would promote reverse reaction by driving charge transitions in ALS in reverse to those in Figure 3.10. Translocation of protons into the matrix would be coupled to transfer of protons from the Q coordinating residues into the central axis, creating a negative charge near Q_d site.

It would promote quinol binding and oxidation, as well as lower the N2 redox potential, enabling reverse electron transfer from N2 to FMN and NAD^+.

Structurally observed (under turnover) states of *Ec*CI likely represent, for the open states a mixture of Steps 1, 4 and 5 in Figure 3.10; and for the closed state - a mixture of Steps 2-3. The mixtures would be present because apart from protonation state these states do not substantially differ and so cannot be resolved by 3D classification. This probably explains why we do not see a clear change in protonation states of $_M$GluTM5, $_M$GluTM12 and $_L$GluTM5, in contrast to E-channel residues (which are charged in Steps 2 and 3).

Overall, this mechanism is straightforward, robust and explains with minimal assumptions the tight coupling of the redox processes and proton translocation over large distances. The existence of the open state with blocked access to the Q_d site is necessary to facilitate quinone movements and to prevent uncoupling which would happen if quinone were to be reduced in the Q_d site with W site open. Therefore, mid/shallow Q sites are used in the open states. The mechanism thus naturally explains the NuoL-only proton exit, why the Q entry site is so narrow, why W site exists and why $_J$TM3 rotates. The arrangement of key *TM12*, *TM8* and *TM7/TM5* sites appears to be a minimum necessary to allow for "domino effect" mechanism.

Despite NuoL-only exit, all three ALS and the E-channel are essential, being responsible for the eventual transfer of one pumped proton each. Therefore, the varying number of ALS is related to the number of protons pumped per cycle in each of evolutionary-related complexes, such as MRP (Steiner & Sazanov, 2020), MBH and MBS, according to the available redox energy (Yu *et al*, 2021). The mechanism appears to be conserved: the Q-like cavity encloses different substrates, such as sodium ions, plastoquinone, hydrogen or polysulfide, while the principle of the redox charge action via the lateral proton transfer along the central axis remains fully applicable (Steiner & Sazanov, 2020).

For clarity and to help understanding, below we provide a full sequence of events during proton ejection (Step 5).

After the first proton from $_L$*TM12* is released into IMS/periplasm (1 H^+ pumped so far), and a sequence of proton hops from $_L$*TM8* to $_L$*TM12*, $_L$*TM7* to $_L$*TM8* and $_M$*TM12* to $_L$*TM5*, repeating twice more in NuoM/N and ending with $_K$E72 donating proton to $_N$*TM5*, the situation

will be as follows, with signs indicating the protonation state (0 or +) of key TM12, TM8, TM7, TM5 residues in antiporters and $_{K}$E72/E36 in the E-channel, respectively:

L (++0+), M (++0+), N(++0+), E-channel (0+).

At this stage the freshly arrived $_{M/N}$*TM12* protons will push $_{L/M}$*TM5* protons onto $_{L/M}$*TM7*. This will re-create the state of NuoL exactly as it was at step 4 and so will push $_{L}$*TM12* proton out (2 H^+ pumped by now). This again initiates a sequence of proton hops from $_{L}$*TM8* to $_{L}$*TM12*, $_{L}$*TM7* to $_{L}$*TM8* and $_{M}$*TM12* to $_{L}$*TM5*, repeating twice more in NuoM/N, and ending now with $_{K}$E36 donating proton to $_{N}$*TM5*, which will create the following distribution:

L (++0+), M (++0+), N(++0+), E-channel (00).

Another repeat of the steps just above (3 H^+ pumped by now), but now ending with $_{N}$*TM12* donating proton to $_{M}$*TM5*, will result in the following distribution:

L (++0+), M (++0+), N(0+0+), E-channel (00).

Finally, another repeat of the steps above (4 H^+ pumped at this stage, accounting for full stoichiometry), but now ending with $_{M}$*TM12* donating proton to $_{L}$*TM5*, will result in the following distribution, i.e. the end of Step 5 as depicted in Figure 3.10:

L (++0+), M (0+0+), N(0+0+), E-channel (00).

Both $_{K}$E36 and $_{K}$E72 can donate protons to $_{N}$*TM5* due to arrival of positive charge to $_{A}$D79 and $_{H}$E157 in the open state – these residues are close enough to interact electrostatically even when $_{J}$TM3 bridge is blocked (as in the open state).

3.6.1 Discussion of the mechanism and alternative proposals

In some recent publications, an alternative path for substrate protons from the cytosol to the Q site was proposed at or near the W site discussed here. It involved either the transfer via specific residues in this area (Grba & Hirst, 2020) or via the branching cavity (similar to cavity leading towards W site) in a gated fashion to allow access of substrate protons instead of waters (Galemou Yoga *et al*, 2020). It is important to note that such a pathway for substrate protons would render them useless for coupling, as the redox reaction would be essentially equilibrated with the cytosol. Only the existence of a very strong and very specific conformational gate linking the Q cavity to all the antiporters would make such a mechanism feasible. However, neither mammalian nor bacterial enzymes show any conformational changes in any of the three ALS under turnover, rendering coupling to quinone protonation via any pathway to the cytosol near PA unlikely. Furthermore, even though there are plenty of charged residues linking the Q cavity to the E-channel, the quinone headgroup in the Q_d site is otherwise well shielded by hydrophobic residues, except for the coordinating residues. In

general, however, this area is relatively hydrophilic as needed for the closed-to-open transition, when NuoA/H loops become exposed to the cytoplasm.

In a recent report, our proposal for ND5/NuoL-only outlet into periplasm was supported based on the structures and MD studies on *Yl*CI (Parey *et al*, 2021). However, the coupling mechanism proposed by the authors involved substrate proton access near the PA interface, which would not allow gating, as noted above. Furthermore, the role for shuttling of the charged quinone intermediates was proposed, which is unlikely due to the extremely short lifetime of such intermediates (Efremov & Sazanov, 2012; Wright *et al*, 2020). Moreover, only one conformational state was observed under turnover conditions, while clearly at least two are necessary for any coupling mechanism to work. The one observed *Yl*CI turnover state resembles the open state of *Oa*CI as similar areas (such as ȷTM3-4 and ND3/NuoA loops) are disordered (but ordered in closed *Oa*CI). Therefore, the reported conformational changes in *Yl*CI may reflect the deactive to open state transition. The failure to observe, so far, a closed state in *Yl*CI could be because it is a high-energy state in enzymes which do not show the apo closed-like state, such as *Ec*CI and *Yl*CI (in contrast to mammalian). More extensive 3D classification (e.g. by focus-revert-classify approach as used here, see Methods) or different data collection strategies may be required to resolve this class in *Yl*CI.

In a recent MD study hydration profiles consistent with ND5-only proton exit were observed in several species, although the interpretation was different (Mühlbauer *et al*, 2020). In another recent report (Kolata & Efremov, 2021) the proposed mechanism involved a key role of the enclosed Q cavity, similar to our arguments, but the rest of mechanism was very vague (in part because only the resting state of *Ec*CI was resolved) and also involved hypothetical conformational changes in the antiporters (which do not happen as we have shown).

A recent publication claimed to overturn our earlier mechanism of complex I (Kampjut & Sazanov, 2020) and suggested yet another alternative (Gu *et al*, 2022), based on a permanently bound ubiquinone shuttling electrons from the deep to the shallow binding site, where they get transferred to a hypothetical loosely bound external molecule of ubiquinone. This proposal is inconsistent with the current knowledge on complex I and the authors do not present any experimental evidence for the binding of an external quinone. On the contrary, there are no clear potential binding sites on the protein surface near Q entry point (even though such sites were computationally predicted recently, they appear to be too far from the Q_s site for efficient electron transfer (Djurabekova *et al*, 2022)). The authors (Gu *et al*, 2022) invoke comparison to Photosystem II (PSII) where primary acceptor Q_A donates electrons to Q_B.

However, in PSII, a non-heme Fe (within iron-quinone complex) promotes electron tunnelling between the quinones (Shevela *et al*, 2012), a feature clearly absent in complex I. There is also no proposal on how protons released from internal quinone would lead to proton translocation. In fact, the main new data in the report was the mode of Q10 binding, while the other structural findings are similar but cover less ground (e.g. there are no turnover conditions) than already published (Kampjut & Sazanov, 2020). Further, the authors reiterated the assignment of the open and closed conformations to the deactive and active states, respectively, without providing any new data for this claim and using the same reasoning which we discussed and dismissed previously (Kampjut & Sazanov, 2020), above and in the recent review (Kampjut & Sazanov, 2022). Another argument was that in the open state quinone cannot bind in the deep site and that the connection between the Q cavity and ALS is interrupted. However, these are exactly the features which allow the open state to act as part of the catalytic cycle in our mechanism. In summary, there is no experimental basis for the mechanistic proposals in this report (Gu *et al*, 2022).

3.7 Conclusions

Complex I is the first and the largest enzyme of the mitochondrial and bacterial respiratory chain. It is essential for energy production and mutations in the enzyme lead to a number of human diseases. The coupling mechanism of the complex I is one of the biggest enigmas in bioenergetics and structural biology and this work was performed to elaborate it. We solved the structures of the minimal complex I version from *E. Coli*, including structures in different redox states, inhibited and under catalytic turnover. It was shown that the bacterial structure has very similar conformational dynamics of the key catalytic elements as the mammalian complex. Thus, based on the compendium of decades of structural, functional and mutagenesis data we suggest the universal coupling mechanism, which can be applied to all species. Further work in this area is likely to involve mutagenesis of key residues in order to verify the mechanism, as well as more direct biophysical experiments to try to visualise open-to-closed transition during turnover by e.g. high-speed AFM (Ando, 2018) or single-molecule FRET (Schuler, 2013).

4 NDH – structure and conformational dynamics

4.1 Structural states and conformational dynamics of NDH

Several NDH structures were published recently, including structures with Fd bound (Schuller *et al*, 2019; Pan *et al*, 2020; Zhang *et al*, 2020). However, none of the works revealed any conformational variability of the enzyme, neither provided insights into the coupling mechanism.

Here we purified NDH from *Thermosynechococcus elongatus* and using extensive cryo-EM data analysis (using several rounds of 3D classifications, see Methods) found that NDH has at least two conformational states which we term open and resting based on similarities to *Ec*CI (Figure 4.1a-c). One of the main features of the open state is the presence of the W site. Although ndhC (NuoA) (NDH/ *E. coli* nomenclature) loop is ordered in the open state, the W site, through which waters can access the Q site, although narrower than in *EcCI/OaCI*, was still identified in NDH (Figure 4.2). Also, both resting and open states contain π-bulge in ndhG (NuoJ) TM3. Other similarities include extended conformation of NuoCD loop, which was discussed above (Figure 3.7c). According to this notation, all the previously published NDH structures showed only the open state of the enzyme (some of them, e.g. PDB 6NBY, show complete disorder of ndhC (NuoA) loop).

Our preparation contained 35 % of complex in the resting state and 65 % in the open state, with the final resolution 4.1 Å and 3.6 Å respectively. These states closely resemble corresponding states of *Ec*CI, where resting state has a wide angle between PA and MA with the active (Q) site mostly disordered. Similarly to *Ec*CI, in NDH resting state the Q cavity is completely disrupted and widely open to solvent on all sides due to protein disorder around the entire cavity. In such a state quinone binding site is not formed and redox coupling is not possible, according to our mechanism discussed above. It represents therefore fully inactive enzyme. The resting state was not previously shown for NDH, probably because of the use of less extensive processing protocols.

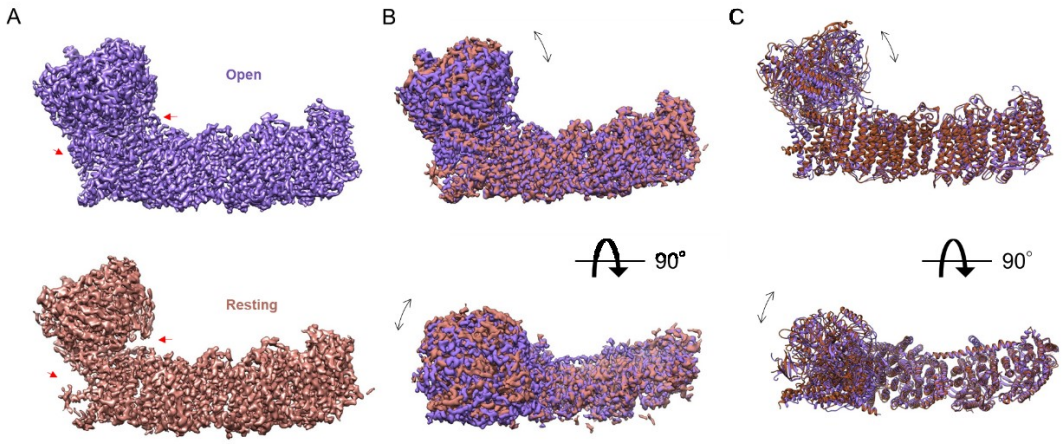

Figure 4.1 Different conformational states of NDH

EM density for the open state is coloured in purple and for the resting state in bronze. A) Overview of the EM density, red arrows point into order/disorder of the active site and ndhL conformational change. B-C) Different states of NDH aligned on the membrane arm and compared.

One of the key functions of the NDH complex is to regulate the balance between ATP and NADPH (Hualing, 2022). This is important because these two molecules have different roles in the cell and need to be kept in balance to support optimal cellular function. For example, ATP is the primary source of energy for cellular processes, while NADPH provides the reducing power that drives numerous anabolic reactions. The NDH-1 complex is also involved in the switch between cyclic and linear electron flows in cyanobacteria. In cyclic electron flow, electrons are recycled within the photosynthetic electron transport chain, whereas in linear electron flow, electrons are transferred from water to NADP+ to generate NADPH. This switch is important because it allows cyanobacteria to adjust their energy production according to their needs. We suggest, that the resting state is an essential regulation factor between those processes as it is necessary to turn off the complex. NDH has exposed FeS cluster due to the absence of the N-module (Figure 4.2), it is possible that it helps the enzyme to sense the redox state of the cell and switch from the open to the resting state.

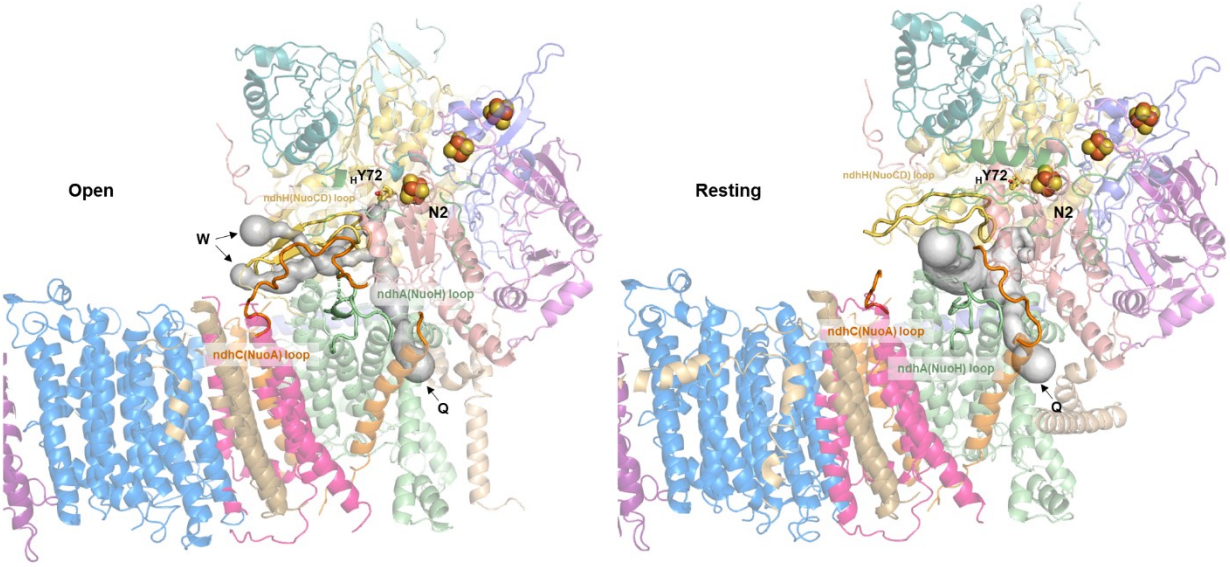

Figure 4.2 Q cavity of the open state NDH

Subunits are coloured in different colours. Important catalytic loops, residues and clusters are highlighted. Q-cavity and W site are shown as grey surface.

Given the analogy to complex I, NDH most likely has an additional closed conformation. Our data analysis did not reveal such a state; however, it is quite possible that it will appear during turnover as in *Ec*CI. We attempted to construct turnover system for NDH but it was unsuccessful (see methods). The main bottleneck is that the electron source for NDH is Fd, which can't be used in large quantities as NADH (in case of complex I) and tends to react with oxygen. Also, we tried to resolve Fd bound state of NDH by adding Fd to the preparation and collecting cryo-EM dataset. However, we did not observe in our reconstructions any bound Fd, likely due to lack of ndhV subunit in our preparation (as was found out later (Zhang *et al*, 2020)) (Figure 1.9). Alternatively, solving apo structures at high pH (current structure was done at pH 6.0, so pH 8-9 may be used) may reveal closed state, if its proportion increases similarly as with *Oa*CI. Solving of the closed state in NDH would prove broad universality of the proposed here coupling mechanism as all of the important features are conserved.

4.2 Analysis of ndhL subunit

One of the differences between the open and the resting states, unique to NDH and reported for the first time in this work, is NdhL conformation. The subunit slides down and twists 90° from vertical to horizontal position upon transition to the resting state, which is a quite remarkable conformational change (Figure 4.3).

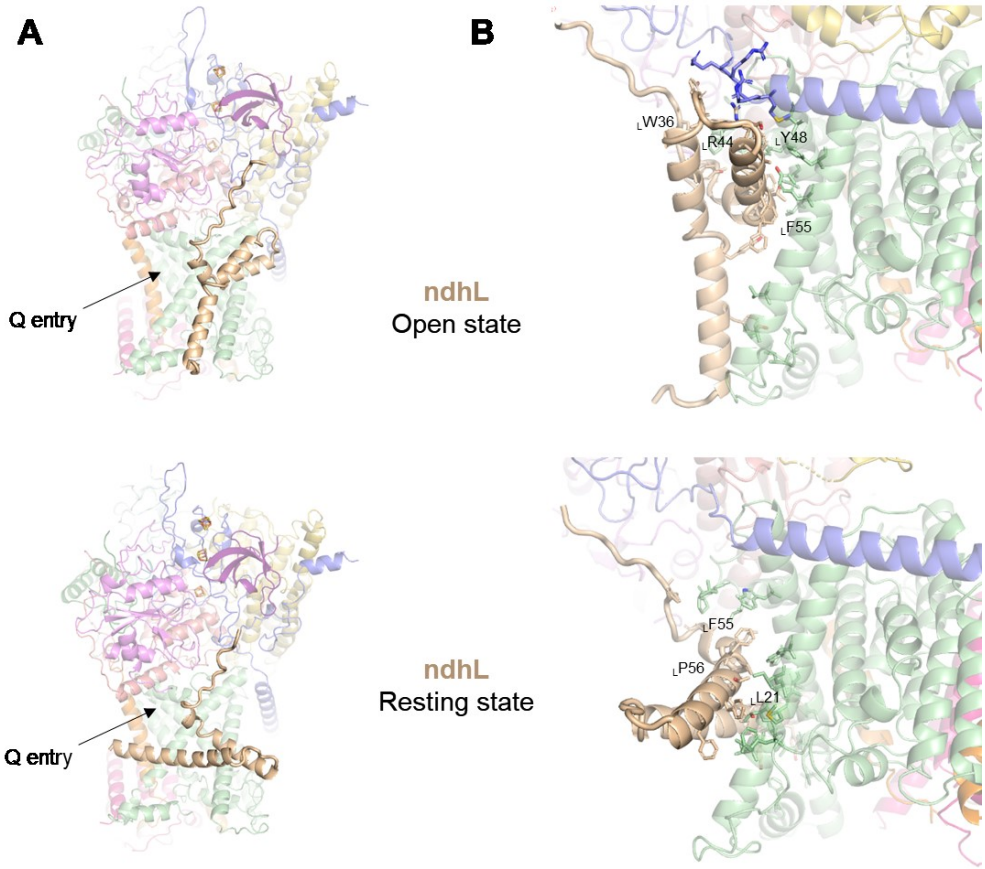

Figure 4.3 Conformational states of ndhL

A) Back view of ndhL subunit in different conformations. B) detailed view on intersubunit interactions. Key conserved residues are indicated.

NdhL is a native NDH subunit which does not have homology to any complex I subunits. It was found that NdhL is important for the docking of hydrophilic subcomplex to hydrophobic subcomplex in both cyanobacteria and plants (Battchikova *et al*, 2005; Shimizu *et al*, 2008). Moreover, deletion of NdhL leads to almost complete absence of NDH-CET activity (Zhang *et al*, 2020). The subunit has relatively small size and consists of a loop, which binds along the peripheral arm, and two intramembrane helices (which we term LH1 and LH2), which bind near the Q-entrance site (Figure 4.3). The subunit is quite hydrophobic (Figure 4.4a-c) and both helices are predicted as transmembrane (Figure 4.4b). However, only LH1 fully traverses the membrane and only in the open state (Figure 4.4a). In the resting state both LH1 and LH2 dip deep into the membrane.

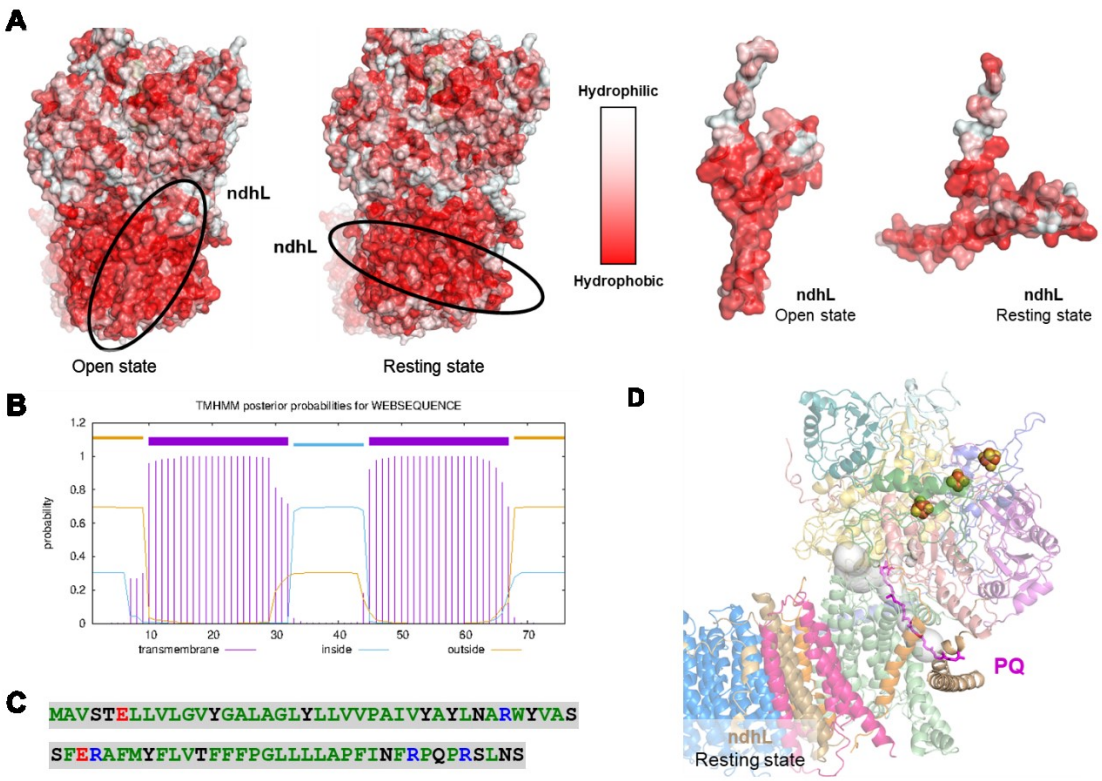

Figure 4.4 Hydrophobicity and structural analysis of ndhL

A) Structure is coloured by the level of hydrophobicity. NdhL subunit is circled. B) Hydrophobicity analysis of sequence using TMHMM web-server (Möller *et al*, 2001)*. Two transmembrane helices are predicted. C) Hydrophobicity analysis. Green: hydrophobic uncharged residues, Red: acidic residues, Blue: basic residues, Black: other residues. D) Plastoquinone (PQ) is modelled (head position based on PDB id: 6khj) into resting state. The conformation of ndhL subunit may hinder PQ movements. Q-cavity is shown in grey transparent surface.*

Several polar amino acids are found near the surface of the membrane and some of those charged residues are highly conserved and interact with the rest of the complex, e.g. $_{L}$R44 (Figure 4.5ab). Other conserved residues are also located mostly at the intersubunit interface, e.g. $_{L}$F55 (Figure 4.5ab).

So what is the role and importance of ndhL? Since ndhL knock-out deactivates the complex, it is clear that ndhL is somehow important for activity. At first we thought that observed ndhL twist would block Q site for PQ entering. However, after careful analysis it was revealed that the Q-site is still accessible, although the entry/exit of Q may be hindered since ndhL might be clashing with the preferred Q tail orientation deep into membrane (Figure 4.4d). Additional and most likely explanation is that the subunit serves as a support for PA-MA interface. In the open state, ndhL interacts closely with PA and ndhJ and ndhA (Nuo I and H in *Ec*CI) while upon shift from open to resting state, ndhL twists and this connection is lost

(Figure 4.3). It is also important to note, that the interaction between ndhJ and ndhL involves the most conserved part of ndhL.

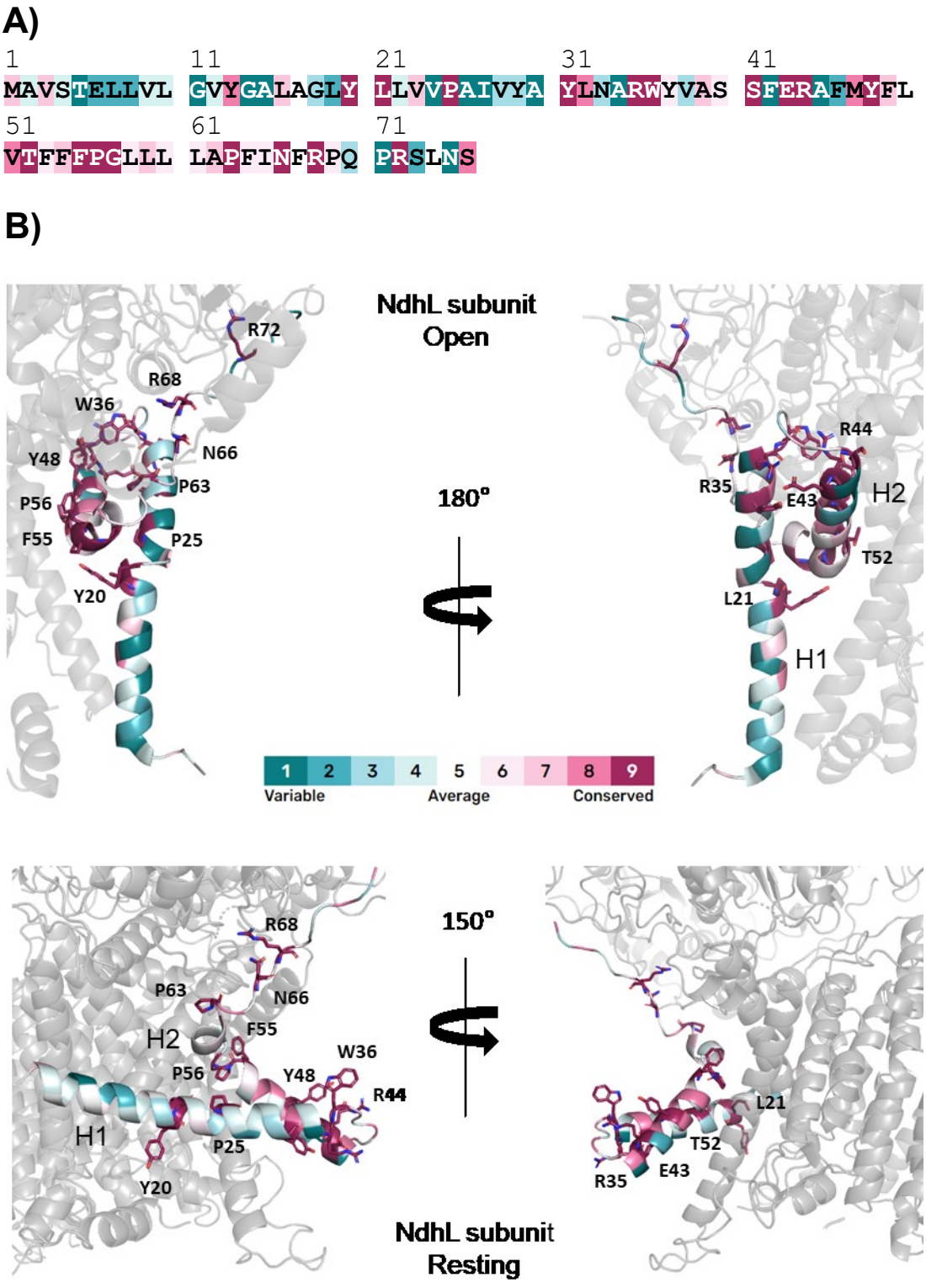

Figure 4.5 ndhL conservation sequence analysis

A-B) Conservation profile. Residues a coloured by conservation score using web-server ConSurf (Ashkenazy *et al*, 2016). *The most conserved residues are shown in sticks.*

Thus, we suggest that this conformational change plays an essential role in NDH regulation. In the open state and potentially also in closed state ndhL interacts with ndhJ and provides support for the PA-MA interface. In the resting state, the support is no longer available, so the PA-MA angle loosens, Q cavity is disassembled and the complex deactivates (Figure 4.2, 4.3 and 4.6). The robust deactivation may thus be achieved via double action of ndhL: destabilisation of PA-MA interface and hindrance for Q access. The deactivation and switch to the linear electron flow is important to balance NADPH/ATP ratio as during the cyclic electron flow only ATP is produced (Huang *et al*, 2015).

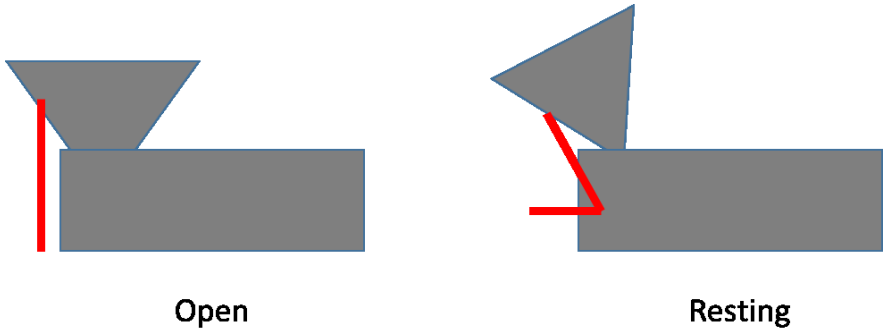

Figure 4.6 Schematic representation of NDH deactivation mechanism

Peripheral arm of the complex is depicted as grey triangle and the membrane part as grey rectangle. NdhL subunit is depicted as red line. NdhL acts a support between two domains in the open state while in the resting state, the conformation change of ndhL leads to loosening of interdomain interface and disorder of the active site.

4.3 Conclusion

In photosynthesis, NDH plays a crucial role in the transfer of electrons from ferredoxin, a small iron-sulfur protein, to plastoquinone. The transfer of electrons from ferredoxin to plastoquinone in cyclic electron transport generates proton gradient and results in the production of ATP, a molecule that is used as an energy currency in cells. NDH is thus important in the regulation of the photosynthetic electron transport chain. It can act as a "bottleneck" in the electron transport chain, limiting the rate at which electrons can be transferred and, as a result, the rate at which ATP is produced. This allows the cell to fine-tune its energy production to match its energy demands. Here we resolved two states of the enzyme: open and resting and suggested the mechanism of NDH activity regulation through the conformational change of ndhL subunit.

References

Adams PD, Afonine P V, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, *et al* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66: 213–221

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: 548–556

Amarneh B & Vik SB (2003) Mutagenesis of subunit N of the Escherichia coli complex I. Identification of the initiation codon and the sensitivity of mutants to decylubiquinone. *Biochemistry* 42: 4800–4808

Ando T (2018) High-speed atomic force microscopy and its future prospects. *Biophys Rev* 10: 285–292

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 Res* 44: W344-50

Barad BA, Echols N, Wang RY-R, Cheng Y, DiMaio F, Adams PD & Fraser JS (2015) EMRinger: side chain-directed model and map validation for 3D cryo-electron microscopy. *Nat Methods* 12: 943–946

Baradaran R, Berrisford JM, Minhas GS & Sazanov LA (2013) Crystal structure of the entire respiratory complex I. *Nature* 494: 443

Battchikova N, Zhang P, Rudd S, Ogawa T & Aro E-M (2005) Identification of NdhL and Ssl1690 (NdhO) in NDH-1L and NDH-1M Complexes of Synechocystis sp. PCC 6803*. *J Biol Chem* 280: 2587–2595

Belevich G, Euro L, Wikström M & Verkhovskaya M (2007) Role of the conserved arginine 274 and histidine 224 and 228 residues in the NuoCD subunit of complex I from Escherichia coli. *Biochemistry* 46: 526–533

Belevich G, Knuuti J, Verkhovsky MI, Wikström M & Verkhovskaya M (2011) Probing the mechanistic role of the long α-helix in subunit L of respiratory Complex I from Escherichia coli by site-directed mutagenesis. *Mol Microbiol* 82: 1086–1095

Belevich N, von Ballmoos C & Verkhovskaya M (2017) Activation of Proton Translocation by Respiratory Complex I. *Biochemistry* 56: 5691–5697

Bernhofer M, Dallago C, Karl T, Satagopam V, Heinzinger M, Littmann M, Olenyi T, Qiu J,

Schütze K, Yachdav G, *et al* (2021) PredictProtein - Predicting Protein Structure and Function for 29 Years. *Nucleic Acids Res* 49: W535–W540

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. *Biochem J* 456: 139–146

Brandt U (2011) A two-state stabilization-change mechanism for proton-pumping complex I. *Biochim Biophys Acta* 1807: 1364–1369

Castro-Guerrero N, Sinha PK, Torres-Bacete J, Matsuno-Yagi A & Yagi T (2010) Pivotal roles of three conserved carboxyl residues of the NuoC (30k) segment in the structural integrity of proton-translocating NADH-quinone oxidoreductase from Escherichia coli. *Biochemistry* 49: 10072–10080

Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS & Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66: 12–21

Chouchani ET, Pell VR, Gaude E, Aksentijević D, Sundier SY, Robb EL, Logan A, Nadtochiy SM, Ord ENJ, Smith AC, *et al* (2014) Ischaemic accumulation of succinate controls reperfusion injury through mitochondrial ROS. *Nature* 515: 431–435

Djurabekova A, Galemou Yoga E, Nyman A, Pirttikoski A, Zickermann V, Haapanen O & Sharma V (2022) Docking and molecular simulations reveal a quinone-binding site on the surface of respiratory complex I. *FEBS Lett* 596: 1133–1146

Dorner K, Vranas M, Schimpf J, Straub IR, Hoeser J & Friedrich T (2017) Significance of [2Fe-2S] Cluster N1a for Electron Transfer and Assembly of Escherichia coli Respiratory Complex I. *Biochemistry* 56: 2770–2778

Efremov RG & Sazanov LA (2011) Structure of the membrane domain of respiratory complex I. *Nature* 476: 414–420

Efremov RG & Sazanov LA (2012) The coupling mechanism of respiratory complex I - a structural and evolutionary perspective. *Biochim Biophys Acta* 1817: 1785–1795

Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66: 486–501

Esterhazy D, King MS, Yakovlev G & Hirst J (2008) Production of reactive oxygen species by complex I (NADH:ubiquinone oxidoreductase) from Escherichia coli and comparison to the enzyme from mitochondria. *Biochemistry* 47: 3964–3971

Esterházy D, King MS, Yakovlev G & Hirst J (2008) Production of reactive oxygen species by complex I (NADH:ubiquinone oxidoreductase) from Escherichia coli and comparison to

the enzyme from mitochondria. *Biochemistry* 47: 3964–3971

Euro L, Belevich G, Bloch DA, Verkhovsky MI, Wikström M & Verkhovskaya M (2009) The role of the invariant glutamate 95 in the catalytic site of Complex I from Escherichia coli. *Biochim Biophys Acta* 1787: 68–73

Euro L, Belevich G, Verkhovsky MI, Wikström M & Verkhovskaya M (2008a) Conserved lysine residues of the membrane subunit NuoM are involved in energy conversion by the proton-pumping NADH:ubiquinone oxidoreductase (Complex I). *Biochim Biophys Acta* 1777: 1166–1172

Euro L, Bloch DA, Wikström M, Verkhovsky MI & Verkhovskaya M (2008b) Electrostatic interactions between FeS clusters in NADH:ubiquinone oxidoreductase (Complex I) from Escherichia coli. *Biochemistry* 47: 3185–3193

Fiedorczuk K, Letts JA, Degliesposti G, Kaszuba K, Skehel M & Sazanov LA (2016) Atomic structure of the entire mammalian mitochondrial complex I. *Nature* 538: 406

Fiedorczuk K & Sazanov LA (2018) Mammalian Mitochondrial Complex I Structure and Disease-Causing Mutations. *Trends Cell Biol* 28: 835–867

Flemming D, Hellwig P & Friedrich T (2003) Involvement of tyrosines 114 and 139 of subunit NuoB in the proton pathway around cluster N2 in Escherichia coli NADH:ubiquinone oxidoreductase. *J Biol Chem* 278: 3055–3062

Flemming D, Hellwig P, Lepper S, Kloer DP & Friedrich T (2006) Catalytic importance of acidic amino acids on subunit NuoB of the Escherichia coli NADH:ubiquinone oxidoreductase (complex I). *J Biol Chem* 281: 24781–24789

Friedrich T, Dekovic DK & Burschel S (2016) Assembly of the Escherichia coli NADH:ubiquinone oxidoreductase (respiratory complex I). *Biochim Biophys Acta - Bioenerg* 1857: 214–223

Friedrich T, van Heek P, Leif H, Ohnishi T, Forche E, Kunze B, Jansen R, Trowitzsch-Kienast W, Höfle G & Reichenbach H (1994) Two binding sites of inhibitors in NADH: ubiquinone oxidoreductase (complex I). Relationship of one site with the ubiquinone-binding site of bacterial glucose:ubiquinone oxidoreductase. *Eur J Biochem* 219: 691–698

Friedrich T & Scheide D (2000) The respiratory complex I of bacteria, archaea and eukarya and its module common with membrane-bound multisubunit hydrogenases. *FEBS Lett* 479: 1–5

Friedrich T, Steinmüller K & Weiss H (1995) The proton-pumping respiratory complex I of bacteria and mitochondria and its homologue in chloroplasts. *FEBS Lett* 367: 107–111

Galemou Yoga E, Parey K, Djurabekova A, Haapanen O, Siegmund K, Zwicker K, Sharma V,

Zickermann V & Angerer H (2020) Essential role of accessory subunit LYRM6 in the mechanism of mitochondrial complex I. *Nat Commun* 11: 6008

Gnandt E, Schimpf J, Harter C, Hoeser J & Friedrich T (2017) Reduction of the off-pathway iron-sulphur cluster N1a of Escherichia coli respiratory complex I restrains NAD(+) dissociation. *Sci Rep* 7: 8754

Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH & Ferrin TE (2018) UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* 27: 14–25

Gostimskaya IS, Grivennikova VG, Cecchini G & Vinogradov AD (2007) Reversible dissociation of flavin mononucleotide from the mammalian membrane-bound NADH: ubiquinone oxidoreductase (complex I). *FEBS Lett* 581: 5803–5806

Grba DN & Hirst J (2020) Mitochondrial complex I structure reveals ordered water molecules for catalysis and proton translocation. *Nat Struct Mol Biol*

Grivennikova VG & Vinogradov AD (2013) Partitioning of superoxide and hydrogen peroxide production by mitochondrial respiratory complex I. *Biochim Biophys Acta - Bioenerg* 1827: 446–454

Gu J, Liu T, Guo R, Zhang L & Yang M (2022) The coupling mechanism of mammalian mitochondrial complex I. *Nat Struct Mol Biol* 29: 172–182

Guerrero-Castillo S, Baertling F, Kownatzki D, Wessels HJ, Arnold S, Brandt U & Nijtmans L (2017) The Assembly Pathway of Mitochondrial Respiratory Chain Complex I. *Cell Metab* 25: 128–139

Gutiérrez-Fernández J, Kaszuba K, Minhas GS, Baradaran R, Tambalo M, Gallagher DT & Sazanov LA (2020) Key role of quinone in the mechanism of respiratory complex I. *Nat Commun* 11: 4135

He Z, Zheng F, Wu Y, Li Q, Lv J, Fu P & Mi H (2015) NDH-1L interacts with ferredoxin via the subunit NdhS in Thermosynechococcus elongatus. *Photosynth Res* 126: 341–349

Hedderich R (2004) Energy-converting [NiFe] hydrogenases from archaea and extremophiles: ancestors of complex I. *J Bioenerg Biomembr* 36: 65–75

Hinttala R, Kervinen M, Uusimaa J, Maliniemi P, Finnilä S, Rantala H, Remes AM, Hassinen IE & Majamaa K (2010) Analysis of functional consequences of haplogroup J polymorphisms m.4216T>C and m.3866T>C in human MT-ND1: mutagenesis of homologous positions in Escherichia coli. *Mitochondrion* 10: 358–361

Hirst J (2013) Mitochondrial complex I. *Annu Rev Biochem* 82: 551–575

Holt PJ, Efremov RG, Nakamaru-Ogiso E & Sazanov LA (2016) Reversible FMN dissociation

from Escherichia coli respiratory complex I. *Biochim Biophys Acta* 1857: 1777–1785

Hualing M (2022) Cyanobacterial NDH-1 Complexes. *Front Microbiol* 13: 933160

Huang W, Yang Y-J, Hu H & Zhang S-B (2015) Different roles of cyclic electron flow around photosystem I under sub-saturating and saturating light intensities in tobacco leaves. *Front Plant Sci* 6: 923

Ingledew WJ & Ohnishi T (1980) An analysis of some thermodynamic properties of iron-sulphur centres in site I of mitochondria. *Biochem J* 186: 111–117

Kaila VRI (2018) Long-range proton-coupled electron transfer in biological energy conversion: towards mechanistic understanding of respiratory complex I. *J R Soc Interface* 15

Kaila VRI & Wikström M (2021) Architecture of bacterial respiratory chains. *Nat Rev Microbiol* 19: 319–330

Kampjut D & Sazanov LA (2020) The coupling mechanism of mammalian respiratory complex I. *Science* 370

Kampjut D & Sazanov LA (2022) Structure of respiratory complex I - An emerging blueprint for the mechanism. *Curr Opin Struct Biol* 74: 102350

Kao M-C, Di Bernardo S, Nakamaru-Ogiso E, Miyoshi H, Matsuno-Yagi A & Yagi T (2005a) Characterization of the membrane domain subunit NuoJ (ND6) of the NADH-quinone oxidoreductase from Escherichia coli by chromosomal DNA manipulation. *Biochemistry* 44: 3562–3571

Kao M-C, Di Bernardo S, Perego M, Nakamaru-Ogiso E, Matsuno-Yagi A & Yagi T (2004) Functional roles of four conserved charged residues in the membrane domain subunit NuoA of the proton-translocating NADH-quinone oxidoreductase from Escherichia coli. *J Biol Chem* 279: 32360–32366

Kao M-C, Nakamaru-Ogiso E, Matsuno-Yagi A & Yagi T (2005b) Characterization of the membrane domain subunit NuoK (ND4L) of the NADH-quinone oxidoreductase from Escherichia coli. *Biochemistry* 44: 9545–9554

Kelley LA, Mezulis S, Yates CM, Wass MN & Sternberg MJE (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 10: 845–858

Kervinen M, Hinttala R, Helander HM, Kurki S, Uusimaa J, Finel M, Majamaa K & Hassinen IE (2006) The MELAS mutations 3946 and 3949 perturb the critical structure in a conserved loop of the ND1 subunit of mitochondrial complex I. *Hum Mol Genet* 15: 2543–2552

Kervinen M, Pätsi J, Finel M & Hassinen IE (2004) A pair of membrane-embedded acidic

residues in the NuoK subunit of Escherichia coli NDH-1, a counterpart of the ND4L subunit of the mitochondrial complex I, are required for high ubiquinone reductase activity. *Biochemistry* 43: 773–781

Klusch N, Senkler J, Yildiz Ö, Kühlbrandt W & Braun H-P (2021) A ferredoxin bridge connects the two arms of plant mitochondrial complex I. *Plant Cell* 33: 2072–2091

Kolata P & Efremov RG (2021) Structure of Escherichia coli respiratory complex I reconstituted into lipid nanodiscs reveals an uncoupled conformation. *Elife* 10

Kotlyar AB & Vinogradov AD (1990) Slow active/inactive transition of the mitochondrial NADH-ubiquinone reductase. *Biochim Biophys Acta* 1019: 151–158

Kramer DM & Evans JR (2011) The importance of energy balance in improving photosynthetic productivity. *Plant Physiol* 155: 70–78

Kravchuk V, Petrova O, Kampjut D, Wojciechowska-Bason A, Breese Z & Sazanov L (2022) A universal coupling mechanism of respiratory complex I. *Nature* 609: 808–814

Kühlbrandt W (2014) The Resolution Revolution. *Science (80-)* 343: 1443–1444

Langen R, Jensen GM, Jacob U, Stephens PJ & Warshel A (1992) Protein control of iron-sulfur cluster redox potentials. *J Biol Chem* 267: 25625–25627

Letts JA, Fiedorczuk K, Degliesposti G, Skehel M & Sazanov LA (2019) Structures of Respiratory Supercomplex I+III2 Reveal Functional and Conformational Crosstalk. *Mol Cell* 75: 1131-1146.e6

Letts JA & Sazanov LA (2015) Gaining mass: the structure of respiratory complex I-from bacterial towards mitochondrial versions. *Curr Opin Struct Biol* 33: 135–145

Maliniemi P, Kervinen M & Hassinen IE (2009) Modeling of human pathogenic mutations in Escherichia coli complex I reveals a sensitive region in the fourth inside loop of NuoH. *Mitochondrion* 9: 394–401

Masuya T, Uno S, Murai M & Miyoshi H (2021) Pinpoint Dual Chemical Cross-Linking Explores the Structural Dynamics of the Ubiquinone Reaction Site in Mitochondrial Complex I. *Biochemistry* 60: 813–824

Michel J, DeLeon-Rangel J, Zhu S, Van Ree K & Vik SB (2011) Mutagenesis of the L, M, and N subunits of Complex I from Escherichia coli indicates a common role in function. *PLoS One* 6: e17420

Millar AH, Whelan J, Soole KL & Day DA (2011) Organization and Regulation of Mitochondrial Respiration in Plants. *Annu Rev Plant Biol* 62: 79–104

MITCHELL P (1961) Coupling of Phosphorylation to Electron and Hydrogen Transfer by a Chemi-Osmotic type of Mechanism. *Nature* 191: 144–148

Möller S, Croning MD & Apweiler R (2001) Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* 17: 646–653

Moparthi VK & Hägerhäll C (2011) The evolution of respiratory chain complex I from a smaller last common ancestor consisting of 11 protein subunits. *J Mol Evol* 72: 484–497

Mühlbauer ME, Saura P, Nuber F, Di Luca A, Friedrich T & Kaila VRI (2020) Water-Gated Proton Transfer Dynamics in Respiratory Complex I. *J Am Chem Soc* 142: 13718–13728

Nakamaru-Ogiso E, Kao M-C, Chen H, Sinha SC, Yagi T & Ohnishi T (2010) The membrane subunit NuoL(ND5) is involved in the indirect proton pumping mechanism of Escherichia coli complex I. *J Biol Chem* 285: 39070–39078

Nicholls DG & Ferguson SJ (2002) Bioenergetics Elsevier

Nosek J & Fukuhara H (1994) NADH dehydrogenase subunit genes in the mitochondrial DNA of yeasts. *J Bacteriol* 176: 5622–5630

Nowaczyk MM, Wulfhorst H, Ryan CM, Souda P, Zhang H, Cramer WA & Whitelegge JP (2011) NdhP and NdhQ: two novel small subunits of the cyanobacterial NDH-1 complex. *Biochemistry* 50: 1121–1124

Ohnishi T (1998) Iron-sulfur clusters/semiquinones in complex I. *Biochim Biophys Acta* 1364: 186–206

Ohnishi T, Ohnishi ST & Salerno JC (2018) Five decades of research on mitochondrial NADH-quinone oxidoreductase (complex I). *Biol Chem* 399: 1249–1264

Pan X, Cao D, Xie F, Xu F, Su X, Mi H, Zhang X & Li M (2020) Structural basis for electron transport mechanism of complex I-like photosynthetic NAD(P)H dehydrogenase. *Nat Commun* 11: 610

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

Parey K, Lasham J, Mills DJ, Djurabekova A, Haapanen O, Yoga EG, Xie H, Kühlbrandt W, Sharma V, Vonck J, *et al* (2021) High-resolution structure and dynamics of mitochondrial complex I-Insights into the proton pumping mechanism. *Sci Adv* 7: eabj3221

Parey K, Wirth C, Vonck J & Zickermann V (2020) Respiratory complex I — structure, mechanism and evolution. *Curr Opin Struct Biol* 63: 1–9

Pätsi J, Kervinen M, Finel M & Hassinen IE (2008) Leber hereditary optic neuropathy mutations in the ND6 subunit of mitochondrial complex I affect ubiquinone reduction kinetics in a bacterial model of the enzyme. *Biochem J* 409: 129–137

Pätsi J, Maliniemi P, Pakanen S, Hinttala R, Uusimaa J, Majamaa K, Nyström T, Kervinen M

& Hassinen IE (2012) LHON/MELAS overlap mutation in ND1 subunit of mitochondrial complex I affects ubiquinone binding as revealed by modeling in Escherichia coli NDH-1. *Biochim Biophys Acta* 1817: 312–318

Peng G, Meyer B, Sokolova L, Liu W, Bornemann S, Juli J, Zwicker K, Karas M, Brutschy B & Michel H (2018) Identification and characterization two isoforms of NADH:ubiquinone oxidoreductase from the hyperthermophilic eubacterium Aquifex aeolicus. *Biochim Biophys acta Bioenerg* 1859: 366–373

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC & Ferrin TE (2004) UCSF Chimera - A visualization system for exploratory research and analysis. *J Comput Chem* 25: 1605–1612

Pintilie G, Zhang K, Su Z, Li S, Schmid MF & Chiu W (2020) Measurement of atom resolvability in cryo-EM maps with Q-scores. *Nat Methods* 17: 328–334

Rochaix J-D (2011) Reprint of: Regulation of photosynthetic electron transport. *Biochim Biophys Acta - Bioenerg* 1807: 878–886

Rohou A & Grigorieff N (2015) CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* 192: 216–221

Sato M, Sinha PK, Torres-Bacete J, Matsuno-Yagi A & Yagi T (2013) Energy transducing roles of antiporter-like subunits in Escherichia coli NDH-1 with main focus on subunit NuoN (ND2). *J Biol Chem* 288: 24705–24716

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

Sazanov LA, Carroll J, Holt P, Toime L & Fearnley IM (2003) A role for native lipids in the stabilization and two-dimensional crystallization of the Escherichia coli NADH-ubiquinone oxidoreductase (complex I). *J Biol Chem* 278: 19483–19491

Sazanov LA & Hinchliffe P (2006) Structure of the hydrophilic domain of respiratory complex I from Thermus thermophilus. *Science* 311: 1430–1436

Schertl P & Braun H-P (2014) Respiratory electron transfer pathways in plant mitochondria. *Front Plant Sci* 5

Schuler B (2013) Single-molecule FRET of protein structure and dynamics - a primer. *J Nanobiotechnology* 11: S2

Schuller JM, Birrell JA, Tanaka H, Konuma T, Wulfhorst H, Cox N, Schuller SK, Thiemann J, Lubitz W, Setif P, *et al* (2019) Structural adaptations of photosynthetic complex I enable ferredoxin-dependent electron transfer. *Science* 363: 257–260

Schulte M, Frick K, Gnandt E, Jurkovic S, Burschel S, Labatzke R, Aierstock K, Fiegen D,

Wohlwend D, Gerhardt S, *et al* (2019) A mechanism to prevent production of reactive oxygen species by Escherichia coli respiratory complex I. *Nat Commun* 10: 2551

Shevela D, Eaton-Rye JJ, Shen J-R & Govindjee (2012) Photosystem II and the unique role of bicarbonate: a historical perspective. *Biochim Biophys Acta* 1817: 1134–1151

Shimizu H, Peng L, Myouga F, Motohashi R, Shinozaki K & Shikanai T (2008) CRR23/NdhL is a Subunit of the Chloroplast NAD(P)H Dehydrogenase Complex in Arabidopsis. *Plant Cell Physiol* 49: 835–842

Sinha PK, Castro-Guerrero N, Patki G, Sato M, Torres-Bacete J, Sinha S, Miyoshi H, Matsuno-Yagi A & Yagi T (2015) Conserved amino acid residues of the NuoD segment important for structure and function of Escherichia coli NDH-1 (complex I). *Biochemistry* 54: 753–764

Sinha PK, Nakamaru-Ogiso E, Torres-Bacete J, Sato M, Castro-Guerrero N, Ohnishi T, Matsuno-Yagi A & Yagi T (2012) Electron transfer in subunit NuoI (TYKY) of Escherichia coli NADH:quinone oxidoreductase (NDH-1). *J Biol Chem* 287: 17363–17373

Sinha PK, Torres-Bacete J, Nakamaru-Ogiso E, Castro-Guerrero N, Matsuno-Yagi A & Yagi T (2009) Critical roles of subunit NuoH (ND1) in the assembly of peripheral subunits with the membrane domain of Escherichia coli NDH-1. *J Biol Chem* 284: 9814–9823

Steimle S, Schnick C, Burger E-M, Nuber F, Krämer D, Dawitz H, Brander S, Matlosz B, Schäfer J, Maurer K, *et al* (2015) Cysteine scanning reveals minor local rearrangements of the horizontal helix of respiratory complex I. *Mol Microbiol* 98: 151–161

Steimle S, Willistein M, Hegger P, Janoschke M, Erhardt H & Friedrich T (2012) Asp563 of the horizontal helix of subunit NuoL is involved in proton translocation by the respiratory complex I. *FEBS Lett* 586: 699–704

Steiner J & Sazanov L (2020) Structure and mechanism of the Mrp complex, an ancient cation/proton antiporter. *Elife* 9: e59407

Stroud DA, Surgenor EE, Formosa LE, Reljic B, Frazier AE, Dibley MG, Osellame LD, Stait T, Beilharz TH, Thorburn DR, *et al* (2016) Accessory subunits are integral for assembly and function of human mitochondrial complex I. *Nature* 538: 123–126

Torres-Bacete J, Nakamaru-Ogiso E, Matsuno-Yagi A & Yagi T (2007) Characterization of the NuoM (ND4) subunit in Escherichia coli NDH-1: conserved charged residues essential for energy-coupled activities. *J Biol Chem* 282: 36914–36922

Torres-Bacete J, Sinha PK, Castro-Guerrero N, Matsuno-Yagi A & Yagi T (2009) Features of subunit NuoM (ND4) in Escherichia coli NDH-1: TOPOLOGY AND IMPLICATION

OF CONSERVED GLU144 FOR COUPLING SITE 1. *J Biol Chem* 284: 33062–33069

Torres-Bacete J, Sinha PK, Sato M, Patki G, Kao M-C, Matsuno-Yagi A & Yagi T (2012) Roles of subunit NuoK (ND4L) in the energy-transducing mechanism of Escherichia coli NDH-1 (NADH:quinone oxidoreductase). *J Biol Chem* 287: 42763–42772

Vercellino I & Sazanov LA (2022) The assembly, regulation and function of the mitochondrial respiratory chain. *Nat Rev Mol Cell Biol* 23: 141–161

Verkhovskaya M & Bloch DA (2013) Energy-converting respiratory Complex I: on the way to the molecular mechanism of the proton pump. *Int J Biochem Cell Biol* 45: 491–511

Verkhovskaya ML, Belevich N, Euro L, Wikstrom M & Verkhovsky MI (2008) Real-time electron transfer in respiratory complex I. *Proc Natl Acad Sci U S A* 105: 3763–3767

de Vries S, Dörner K, Strampraad MJF & Friedrich T (2015) Electron tunneling rates in respiratory complex I are tuned for efficient energy conversion. *Angew Chem Int Ed Engl* 54: 2844–2848

Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, Verma V, Keedy DA, Hintze BJ, Chen VB, *et al* (2018) MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci* 27: 293–315

Wright JJ, Fedor JG, Hirst J & Roessler MM (2020) Using a chimeric respiratory chain and EPR spectroscopy to determine the origin of semiquinone species previously assigned to mitochondrial complex I. *BMC Biol* 18: 54

Yip C, Harbour ME, Jayawardena K, Fearnley IM & Sazanov LA (2011) Evolution of respiratory complex I: 'supernumerary' subunits are present in the alpha-proteobacterial enzyme. *J Biol Chem* 286: 5023–5033

Yu H, Schut GJ, Haja DK, Adams MWW & Li H (2021) Evolution of complex I-like respiratory complexes. *J Biol Chem* 296: 100740

Zhang C, Shuai J, Ran Z, Zhao J, Wu Z, Liao R, Wu J, Ma W & Lei M (2020) Structural insights into NDH-1 mediated cyclic electron transfer. *Nat Commun* 11: 888

Zheng SQ, Palovcak E, Armache J-P, Verba KA, Cheng Y & Agard DA (2017) MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* 14: 331–332

Zickermann V, Wirth C, Nasiri H, Siegmund K, Schwalbe H, Hunte C & Brandt U (2015) Structural biology. Mechanistic insight from the crystal structure of mitochondrial complex I. *Science* 347: 44–49

Zilberstein D, Agmon V, Schuldiner S & Padan E (1984) Escherichia coli intracellular pH, membrane potential, and cell growth. *J Bacteriol* 158: 246–252

Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E & Scheres SH (2018) New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* 7

Zivanov J, Nakane T & Scheres SHW (2019) A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCrJ* 6: 5–17

Zu Y, Di Bernardo S, Yagi T & Hirst J (2002) Redox properties of the [2Fe-2S] center in the 24 kDa (NQO2) subunit of NADH:ubiquinone oxidoreductase (complex I). *Biochemistry* 41: 10056–10069

A. Appendix 1. Model and data collection statistics for DDM datasets

	DDM_APO	DDM_NADH
	Resting	Resting
Data collection and processing		
Microscope	Titan Krios	Titan Krios
Camera	Falcon 3 linear	Falcon 3 linear
Magnification	75000	75000
Voltage (kV)	300	300
Electron exposure (e/Å2)	89	89
Automation software	EPU	EPU
Number of frames	40	40
Defocus range (μm)	~ -1 to -2	~ -1 to -2
Pixel size (Å)	1.061	1.061
Symmetry imposed	C1	C1
Number of micrographs	3658	3492
Initial particle images	755k	441k
Final particle images	121k	145k
Map resolution (Å) at 0.143 FSC threshold	3.58 Entire 3.18 PA, 3.40 MD	3.22 Entire 2.99 PA, 3.30 MD
Refinement		
Initial model used (PDB code)	3RKO + 4HEA homology models	3RKO + 4HEA homology models
Refinement package	Phenix, real space	Phenix, real space
Model resolution (Å) 0.5 FSC threshold	3.34	3.21
Local resolution range (Å)	3.2-7.0	3.0-6.2
Cross-correlation		
Mask	0.88	0.87
Volume	0.87	0.87
Map sharpening B factor (Å2)	-99 Entire, -81 PA, -100 MD	-82 Entire -76 PA, -85 MD
Model composition		
Protein residues	4675	4681
Ligands	13	15
B factors (Å)		
Protein	87	75
Ligand	104	88
Q-scores		
Protein + ligands	0.60	0.62
R.m.s. deviations		
Bond length (Å)	0.004	0.004
Bond angles (°)	0.821	0.813
Validation		
MolProbity score	1.67	1.66
EMRinger score	2.07	2.36
Clashscore	7.03	7.18
Poor rotamers (%)	0.45	0.29
C-beta deviations (%)	0	0
CaBLAM outliers (%)	1.94	1.99
Ramachandran plot		
Favored (%)	95.94	96.07
Allowed (%)	4.06	3.93
Disallowed (%)	0	0

B. Appendix 2. Model and data collection statistics for DDM/LMNG datasets

	DDM/LMNG APO		DDM/LMNG Turnover pH6		
	Resting	Open	Resting	Open	Closed
Data collection and processing					
Microscope	Titan Krios		Titan Krios		
Camera	K3		K3		
Magnification	81000		81000		
Voltage (kV)	300		300		
Electron exposure (e/Å2)	78		78		
Automation software	Serial-EM		Serial-EM		
Number of frames	60		60		
Defocus range (μm)	~-1 to -2		~-1 to -2		
Pixel size (Å)	1.06		1.06		
Symmetry imposed	C1		C1		
Number of micrographs	3815		2757		
Initial particle images	918k		773k		
Final particle images	199.3k	167k	92.8k	97.9k	8.2k
Map resolution (Å) at 0.143 FSC threshold	2.7 Entire 2.4 PA, 2.6 MD	2.4 Entire 2.2 PA, 2.4 MD	3.1 Entire 2.6 PA, 2.8 MD	2.5 Entire 2.3 PA, 2.6 MD	3.4 Entire 3.2 NuoFEG, 3.2 NuoN-CD, 3.3 NuoNML
Refinement					
Initial model used (PDB code)	3RKO + 4HEA homology models		3RKO + 4HEA homology models		
Refinement package	Phenix, real space refine		Phenix, real space refine		
Model resolution (Å) 0.5 FSC threshold	2.52	2.5	2.72	2.48	3.34
Local resolution range(Å)	2.3-6.2	2.2-7.0	2.7-6.7	2.2-5.5	3.0-20
Cross-correlation					
Mask	0.9	0.89	0.92	0.92	0.85
Volume	0.9	0.89	0.92	0.91	0.84
Map sharpening B factor (Å2)	-33 Entire, -13 PA, -28 MD	-15 Entire, -9 PA, -17 MD	-50 Entire -18 PA, -20 MD	-19 Entire -8 PA, -20 MD	-29 Entire -10 NuoFEG, 0 NuoN-CD, -10 NuoNML
Model composition					
Protein residues	4612	4717	4645	4732	4772
Ligands	23	25	30	36	25
Waters	Total: 837 PA: 665 MD: 172	Total: 948 PA: 683 MD: 265	Total: 581 PA: 421 MD: 160	Total: 987 PA: 639 MD: 348	-
B factors (Å)					
Protein	57	54	61	54	61
Ligand	85	84	100	93	83
Waters	43	43	42	46	-
Q-scores					
Protein + ligands	0.75	0.74	0.74	0.77	0.62
Water	0.89	0.89	0.90	0.90	-
R.m.s. deviations					
Bond length (Å)	0.004	0.005	0.006	0.005	0.004
Bond angles (°)	0.769	0.803	0.709	0.736	0.754
Validation					
MolProbity score	1.61	1.51	1.57	1.41	1.55
EMRinger score	4.17	4.62	4.49	4.61	3.07
Clashscore	4.94	4.99	4.23	4.47	6.47
Poor rotamers (%)	1.94	1.61	2.11	1.47	0.49
C-beta deviations	0	0.02	0	0	0
CaBLAM outliers (%)	1.18	1.40	1.18	1.18	1.51
Ramachandran plot					
Favored (%)	97.28	97.56	97.41	97.76	96.81
Allowed	2.68	2.4	2.56	2.2	3.17
Disallowed (%)	0.04	0.04	0.02	0.04	0.02

	DDM/LMNG DQ		DDM/LMNG NADH+FMN	DDM/LMNG PiericidinA + FMN
	Resting	Open	Open	Open
Data collection and processing				
Microscope	Titan Krios		Titan Krios	Glacios
Camera	K3		K3	Falcon3
Magnification	81000		81000	120000
Voltage (kV)	300		300	200
Electron exposure (e/Å2)	78		78	90
Automation software	Serial-EM		Serial-EM	EPU
Number of frames	60		60	65
Defocus range (μm)	~ -1 to -2		~ -1 to -2	~ -1.2 to -2.4
Pixel size (Å)	1.06		1.06	1.21
Symmetry imposed	C1		C1	C1
Number of micrographs	3150		3033	3582
Initial particle images	852k		563k	463k
Final particle images	66.8k	56.2k	36.8k	99.9k
Map resolution (Å) at 0.143 FSC threshold	3.1 Entire 2.6 PA, 2.8 MD	2.7 Entire 2.5 PA, 2.8 MD	3.0 Entire 2.8 PA, 3.0 MD	3.2 Entire 3.0 PA, 3.2 MD
Refinement				
Initial model used (PDB code)	3RKO + 4HEA homology models		3RKO + 4HEA homology models	
Refinement package	Phenix, real space refine		Phenix, real space refine	
Model resolution (Å) 0.5 FSC threshold	2.88	2.81	3.0	3.21
Local resolution range(Å)	2.6-8.3	2.3-7.4	2.7-8.6	2.8-8.8
Cross-correlation				
Mask	0.88	0.9	0.88	0.88
Volume	0.89	0.9	0.89	0.88
Map sharpening B factor (Å2)	-33 Entire, -17 PA, -24 MD	-20 Entire, -11 PA, -21 MD	-29 Entire -21 PA, -24 MD	-65 Entire -62 PA, -64 MD
Model composition				
Protein residues	4630	4719	4710	4720
Ligands	23	28	32	26
B factors (Å)				
Protein	70	62	73	86
Ligand	102	92	109	110
Q-scores				
Protein + Ligand	0.68	0.71	0.65	0.61
R.m.s. deviations				
Bond length (Å)	0.006	0.005	0.005	0.004
Bond angles (°)	0.7	0.756	0.769	0.739
Validation				
MolProbity score	1.63	1.43	1.52	1.36
EMRinger score	3.63	4.09	3.47	2.89
Clashscore	4.97	4.92	6.11	5.31
Poor rotamers (%)	2.21	1.25	1.3	0.39
C-beta deviations	0	0	0	0
CaBLAM outliers (%)	1.19	1.38	1.41	1.51
Ramachandran plot				
Favored (%)	97.47	97.54	97.51	97.63
Allowed (%)	2.51	2.44	2.46	2.35
Disallowed (%)	0.02	0.02	0.02	0.02

	DDM/LMNG Turnover pH8			
	Resting	Open	Open-ready	Closed
Data collection and processing				
Microscope	Titan Krios			
Camera	K3			
Magnification	81000			
Voltage (kV)	300			
Electron exposure (e/Å2)	80			
Automation software	Serial-EM			
Number of frames	80			
Defocus range (μm)	~ -1 to -2			
Pixel size (Å)	1.06			
Symmetry imposed	C1			
Number of micrographs	8659			
Initial particle images	1.65M			
Final particle images	82.6k	67.2k	40.8k	32.6k
Map resolution (Å) at 0.143 FSC threshold	3.1 Entire 2.6 PA, 3.1 MD	3.0 Entire 2.6 PA, 3.1 MD	3.0 Entire 2.7 PA, 3.0 MD	3.0 Entire 2.7 PA, 3.1 MD
Refinement				
Initial model used (PDB code)	3RKO + 4HEA homology models			
Refinement package	Phenix, real space refine			
Model resolution (Å) 0.5 FSC threshold	3.0	3.0	3.0	3.0
Local resolution range(Å)	2.6-8.3	2.3-7.4	2.7-8.6	2.8-8.8
Cross-correlation				
Mask	0.87	0.88	0.88	0.87
Volume	0.86	0.87	0.88	0.87
Map sharpening B factor (Å2)	-34 Entire, -25 PA, -38 MD	-30 Entire, -19 PA, -34 MD	-28 Entire -22 PA, -33 MD	-23 Entire -20 PA, -27 MD
Model composition				
Protein residues	4645	4732	4751	4764
Ligands	23	25	26	28
B factors (Å)				
Protein	66	62	52	61
Ligand	93	88	83	82
Q-scores				
Protein + Ligand	0.67	0.68	0.70	0.69
R.m.s. deviations				
Bond length (Å)	0.007	0.004	0.004	0.004
Bond angles (°)	0.848	0.77	0.747	0.776
Validation				
MolProbity score	1.68	1.7	1.47	1.65
EMRinger score	3.49	2.92	3.18	2.63
Clashscore	5.41	5.58	5.5	5.73
Poor rotamers (%)	1.96	2.59	1.65	2.08
C-beta deviations	0	0	0	0
CaBLAM outliers (%)	1.32	1.29	1.35	1.65
Ramachandran plot				
Favored (%)	97.02	97.59	97.98	97.53
Allowed (%)	2.96	2.39	2.0	2.45
Disallowed (%)	0.02	0.02	0.02	0.02

C. Appendix 3. Model and data collection statistics for LMNG datasets

	LMNG APO	LMNG Turnover pH6			
	Open-ready	Resting	Open	Open-ready	Closed
Data collection and processing					
Microscope	Glacios	Titan Krios			
Camera	Falcon3	K3			
Magnification	120000	81000			
Voltage (kV)	200	300			
Electron exposure (e/Å2)	74	80			
Automation software	EPU	Serial-EM			
Number of frames	53	80			
Defocus range (μm)	~ -1.2 to -2.4	~ -1 to -2			
Pixel size (Å)	1.21	1.06			
Symmetry imposed	C1	C1			
Number of micrographs	4295	11316			
Initial particle images	1M	4.5M			
Final particle images	109k	189.7k	31.8k	317.1k	170k
Map resolution (Å) at 0.143 FSC threshold	3.4 Entire 3.2 NuoFEG, 3.3 NuoN-CD, 3.4 NuoNML	2.7 Entire 2.4 PA, 2.95 MD	3.1 Entire 2.88 PA, 3.14 MD	2.35 Entire 2.15 PA, 2.38 MD	2.5 Entire 2.3 PA, 2.5 MD
Refinement					
Initial model used (PDB code)	3RKO + 4HEA homology models				
Refinement package	Phenix, real space refine				
Model resolution (Å) 0.5 FSC threshold	3.5	2.61	3.5	2.12	2.27
Local resolution range(Å)	3.0-11.2	2.2-10.8	2.7-13.7	2.2-5.5	2.2-5.8
Cross-correlation					
Mask	0.85	0.89	0.87	0.93	0.92
Volume	0.84	0.89	0.86	0.93	0.92
Map sharpening B factor (Å2)	-70 Entire -50 NuoFEG, -50 NuoN-CD, -60 NuoNML	-26 Entire, -18 PA, -39 MD	-32 Entire -24 PA, -35 MD	-16 Entire -13 PA, -22 MD	-15 Entire -12 PA, -25 MD
Model composition					
Protein residues	4750	4650	4721	4748	4774
Ligands	14	28	23	28	25
Waters				Total: 1298 PA: 865 MD: 433	Total: 1184 PA: 850 MD: 334
B factors (Å)					
Protein	92	57	61	43	48
Ligand	95	94	100	67	71
Waters	-	-	-	40	41
Q-scores					
Protein + ligands	0.57	0.74	0.67	0.82	0.80
Water	-	-	-	0.93	0.92
R.m.s. deviations					
Bond length (Å)	0.004	0.006	0.005	0.005	0.006
Bond angles (°)	0.849	0.815	0.746	0.735	0.794
Validation					
MolProbity score	1.51	1.56	1.37	1.44	1.45
EMRinger score	2.15	3.87	2.91	5.2	4.92
Clashscore	6.72	4.84	5.86	4.62	4.21
Poor rotamers (%)	0.41	1.56	0.23	1.68	1.67
C-beta deviations	0	0	0	0	0
CaBLAM outliers (%)	1.46	1.18	1.27	1.35	1.27
Ramachandran plot					
Favored (%)	97.26	97.39	97.78	97.86	97.66
Allowed	2.72	2.56	2.16	2.1	2.26
Disallowed (%)	0.02	0.04	0.02	0.04	0.08

D. Appendix 4. Model and data collection statistics for NDH dataset

	NDH_APO	
	Open	Resting
Data collection and processing		
Microscope	Titan Krios	
Camera	Falcon 3 linear	
Magnification	75000	
Voltage (kV)	300	
Electron exposure (e/Å2)	89	
Automation software	EPU	
Number of frames	40	
Defocus range (μm)	~ -1 to -2	
Pixel size (Å)	1.061	
Symmetry imposed	C1	
Number of micrographs	2980	
Initial particle images	438k	
Final particle images	57k	34k
Map resolution (Å) at 0.143 FSC threshold	3.57	4.06
Refinement		
Initial model used (PDB code)	6HUM	
Refinement package	Phenix, real space	
Model resolution (Å) 0.5 FSC threshold	3.61	4.16
Local resolution range (Å)	3.3-7.0	3.9-9.3
Cross-correlation		
Mask	0.86	0.79
Volume	0.83	0.76
Map sharpening B factor (Å2)	-89.7	-130
Model composition		
Protein residues	4675	3754
Ligands	13	3
B factors (Å)		
Protein	87	75
Ligand	104	88
Q-scores		
Protein + ligands	0.58	0.46
R.m.s. deviations		
Bond length (Å)	0.008	0.007
Bond angles (°)	1.153	1.255
Validation		
MolProbity score	2.14	2.10
EMRinger score	1.75	0.58
Clashscore	10.05	11.03
Poor rotamers (%)	1.26	0.82
C-beta deviations (%)	0.03	0.06
CaBLAM outliers (%)	5.67	5.94
Ramachandran plot		
Favored (%)	90.31	90.45
Allowed (%)	9.56	9.36
Disallowed (%)	0.13	0.19

E. Appendix 5. *E. coli* complex I mutation data

The list of mutations in *E.coli* complex I. Oxidoreductase activities were measured with either O_2, DQ or UQ_1 as final acceptor and the range observed is shown. FMN-site activities were measured with either ferricyanide or hexaammineruthenium III as final acceptor. Proton pumping rates are very approximate. "Reduced" corresponds to the value ~50-80% and "low" to below ~50%. Abbreviations: SB - salt bridge, HB – hydrogen bond, HL – helix HL.

a) Mutations in the peripheral arm

Mutation	Amino acid location	Expression/ assembly	Effect			Reference and comments
			FMN-site activity	Oxido-reductase activity	Proton pumping activity	
NuoF						
E95Q	NADH binding site	normal	20-40%	40%	NA	(Euro *et al*, 2009)
NuoCD						
S104A	surface	NA	103%	106-114%	NA	(Castro-Guerrero *et al*, 2010)
A134S	surface	NA	82%	NA	NA	(Castro-Guerrero *et al*, 2010)
E138A	Intersubunit interface, before $_{CD}$LHL	low	49%	2-3%	0%	(Castro-Guerrero *et al*, 2010)
E138Q		normal	44%	20-29%	40%	(Castro-Guerrero *et al*, 2010)
E138D		reduced	51%	4-6%	10%	(Castro-Guerrero *et al*, 2010)
R139A		NA	87%	102-118%	NA	(Castro-Guerrero *et al*, 2010)
E140A		reduced	42%	5-7%	0%	(Castro-Guerrero *et al*, 2010)
E140Q		reduced	33%	7-11%	0%	(Castro-Guerrero *et al*, 2010)

E140D		normal	94%	85-104%	100%	(Castro-Guerrero *et al*, 2010)
D143A		reduced	50%	2-3%	0%	(Castro-Guerrero *et al*, 2010)
D143N		reduced	69%	5-9%	10%	(Castro-Guerrero *et al*, 2010)
D143E		normal	103%	128-132%	90%	(Castro-Guerrero *et al*, 2010)
G146A	surface	NA	102%	114-117%	NA	(Castro-Guerrero *et al*, 2010)
F149A		NA	109%	NA	NA	(Castro-Guerrero *et al*, 2010)
R156A		NA	124%	123-131%	NA	(Castro-Guerrero *et al*, 2010)
G166A		NA	113%	117%	NA	(Castro-Guerrero *et al*, 2010)
H167A		NA	93%	74-97%	NA	(Castro-Guerrero *et al*, 2010)
P168A	$_{CD}$LHL	NA	91%	75-84%	NA	(Castro-Guerrero *et al*, 2010)
K171A		NA	76%	71-80%	NA	(Castro-Guerrero *et al*, 2010)
K171R		NA	80%	78-88%	NA	(Castro-Guerrero *et al*, 2010)
P182A		NA	100%	71-88%	NA	(Castro-Guerrero *et al*, 2010)
G221V	NuoCD loop, Q-site	NA	50%	1-6%	NA	(Sinha *et al*, 2015)
P222A		reduced	51%	7-9%	NA	(Sinha *et al*, 2015)
H224A		NA	88%	64%	65%	(Belevich *et al*, 2007)
H228A		NA	100%	48%	50%	(Belevich *et al*, 2007)
H228R		normal	85%	67-98%	NA	(Sinha *et al*, 2015)

G229A		normal	67%	24-38%	NA	(Sinha *et al*, 2015)
G229V		reduced	42%	1-4%	NA	(Sinha *et al*, 2015)
R232A		NA	76%	50-74%	NA	(Sinha *et al*, 2015)
R232K		NA	80%	61-93%	NA	(Sinha *et al*, 2015)
G239A	NuoA loop interface	normal	96%	88-118%	NA	(Sinha *et al*, 2015)
G239V		reduced	62%	35-52%	NA	(Sinha *et al*, 2015)
E240A	NuoH interface	normal	103%	14-26%	NA	(Sinha *et al*, 2015)

a) Mutations in the peripheral arm, continuation

Mutation	Amino acid location	Expression/ assembly	Effect			Reference and comments
			FMN-site activity	Oxido-reductase activity	Proton pumping activity	
NuoCD						
E240D	NuoH interface	normal	71%	46-75%	NA	(Sinha *et al*, 2015)
H253A	N2 environment	low	24%	13-18%	NA	(Sinha *et al*, 2015)
H253K		low	26%	4-7%	NA	(Belevich *et al*, 2007)
G255A		normal	81%	82-105%	NA	(Sinha *et al*, 2015)
G255V		reduced	67%	35-51%	NA	(Sinha *et al*, 2015)
R274A		normal	47%	3-4%	NA	(Sinha *et al*, 2015)
R274K		normal	77%	45-60%	NA	(Sinha *et al*, 2015)
R274A		NA	77%	22%	25%	(Belevich *et al*, 2007)
Y277A	N2 environment, Q-site	normal	66%	3-7%	NA	(Sinha *et al*, 2015)
Y277W		normal	60%	2-8%	NA	(Sinha *et al*, 2015)
Y277F		NA	76%	17-27%	NA	(Sinha *et al*, 2015), capsaicin-40 insensitive

E292Q	Intersubunit interface, SB with $_{c}$R307	low	29%	4-6%	NA	(Sinha *et al*, 2015)
E292D		low	14%	2-8%	NA	(Sinha *et al*, 2015)
R302A	surface	normal	78%	78-97%	NA	(Sinha *et al*, 2015)
R302K		normal	94%	92-105%	NA	(Sinha *et al*, 2015)
E312Q	Intersubunit interface, SB with cR439	low	19%	3-7%	NA	(Sinha *et al*, 2015)
E312D		normal	82%	80-95%	NA	(Sinha *et al*, 2015)
R315A	Intersubunit interface	low	30%	2-3%	NA	(Sinha *et al*, 2015)
R315K		reduced	49%	41-49%	NA	(Sinha *et al*, 2015)
H319A	Intersubunit interface	normal	82%	84-106%	NA	(Sinha *et al*, 2015)
H319R		normal	47%	30-38%	NA	(Sinha *et al*, 2015)
D329A	Q-site, HB with $_{c}$H228	normal	92%	42-56%	NA	(Sinha *et al*, 2015)
D329E		normal	87%	54-63%	NA	(Sinha *et al*, 2015)
H359A	HB with N2	reduced	74%	13-28%	NA	(Sinha *et al*, 2015)
H359K		reduced	68%	36-52%	NA	(Sinha *et al*, 2015)
R560K	Part of internal β-sheet	low	40%	8-14%	NA	(Sinha *et al*, 2015)
R600A	Intersubunit interface	reduced	61%	47-56%	NA	(Sinha *et al*, 2015)
NuoB						
E67Q	Close to N2, HB with $_{c}$R357	NA	89%	10%	NA	(Flemming *et al*, 2006), piericidin A insensitive
E67D		NA	100%	78%	NA	(Flemming *et al*, 2006), piericidin A insensitive
D77N	NuoH interface, part of Q-cavity	NA	78%	12%	NA	(Flemming *et al*, 2006), piericidin A insensitive
D77E		NA	78%	54%	NA	(Flemming *et al*, 2006)
D94N	NuoH interface, part of Q-cavity	NA	78%	12%	NA	(Flemming *et al*, 2006)
D94E		NA	89%	83%	NA	(Flemming *et al*, 2006)

Y114C	Intersubunit interface	normal	80%	100%	NA	(Flemming *et al*, 2003)
D115N		NA	78%	45%	NA	(Flemming *et al*, 2006)
Y114C/ Y139F		normal	50%	20%	NA	(Flemming *et al*, 2003)
E119Q	Intersubunit interface	NA	78%	88%	NA	(Flemming *et al*, 2006)
Y139C	cdLHL interface	normal	80%	100%	NA	(Flemming *et al*, 2003)
D146N	NuoI interface	NA	89%	59%	NA	(Flemming *et al*, 2006)
D152N	surface	NA	78%	59%	NA	(Flemming *et al*, 2006)
Y154H	Part of internal β-sheet	normal	90%	100%	NA	(Flemming *et al*, 2003)
E163Q	HB with $_{B}$R161	NA	78%	76%	NA	(Flemming *et al*, 2006)

a) Mutations in the peripheral arm, continuation

Mutation	Amino acid location	Expression/ assembly	Effect			Reference and comments
			FMN-site activity	Oxido-reductase activity	Proton pumping activity	
NuoI						
C60A	Bond with N6a Fe-S cluster	NA	17%	1-2%	0%	(Sinha *et al*, 2012)
C60S		NA	31%	5-7%	5%	(Sinha *et al*, 2012)
C60H		NA	20%	2%	0%	(Sinha *et al*, 2012)
C63A		NA	23%	2%	0%	(Sinha *et al*, 2012)
C63S		normal	50%	20%	20%	(Sinha *et al*, 2012)
C66S		NA	18%	22%	0%	(Sinha *et al*, 2012)
C66H		NA	19%	1-2%	0%	(Sinha *et al*, 2012)
C70S	Bond with N6b Fe-S cluster	NA	19%	4%	0%	(Sinha *et al*, 2012)
C70H		NA	15%	1-2%	0-5%	(Sinha *et al*, 2012)

C99S		NA	17%	1-2%	0%	(Sinha *et al*, 2012)
C99H		NA	15%	1%	0%	(Sinha *et al*, 2012)
C102S		NA	15%	2-5%	0-5%	(Sinha *et al*, 2012)
C102H		NA	17%	1-2%	0%	(Sinha *et al*, 2012)
C105S		NA	17%	1-2%	0%	(Sinha *et al*, 2012)
C109S		NA	21%	1-4%	0%	(Sinha *et al*, 2012)
T30A	NuoH interface	NA	64%	58-61%	reduced	(Sinha *et al*, 2012)
P34A	H1 connecting loop	NA	65%	57-65%	reduced	(Sinha *et al*, 2012)
P42A	NuoCD interface	NA	108%	61-65%	reduced	(Sinha *et al*, 2012)
R43A		NA	84%	46-53%	reduced	(Sinha *et al*, 2012)
Y44A	NuoB interface	NA	139%	120%	100%	(Sinha *et al*, 2012)
G46A	Close to N6b	NA	79%	78%	100%	(Sinha *et al*, 2012)
R52A	Surface, part of N6a interacting loop	NA	84%	64-75%	100%	(Sinha *et al*, 2012)
P54A		NA	95%	73-79%	100%	(Sinha *et al*, 2012)
G56A		NA	94%	47-54%	reduced	(Sinha *et al*, 2012)
E58A		NA	99%	64-85%	reduced	(Sinha *et al*, 2012)
V61A		normal	67%	31-33%	reduced	(Sinha *et al*, 2012)
L65A	Close to N6a	normal	74%	67-72%	70%	(Sinha *et al*, 2012)
P71A	N6b interface	normal	75%	45-56%	reduced	(Sinha *et al*, 2012)
I75A		normal	66%	28-40%	reduced	(Sinha *et al*, 2012)
G85A	surface	NA	106%	88-116%	100%	(Sinha *et al*, 2012)
F92A	N6a interface	NA	79%	76-81%	100%	(Sinha *et al*, 2012)
R93A	surface	NA	101%	92-101%	100%	(Sinha *et al*, 2012)
I94A	Nba interface	normal	108%	92-99%	100%	(Sinha *et al*, 2012)

I94G		NA	92%	93-119%	NA	(Sinha *et al*, 2012)
R98A	Intersubunit interface, near $_{B}$C99	normal	53%	9-18%	20%	(Sinha *et al*, 2012)
I100A		normal	101%	85-86%	NA	(Sinha *et al*, 2012)
I100G		NA	86%	61-85%	NA	(Sinha *et al*, 2012)
G103A	near N6b	NA	88%	80-100%	100%	(Sinha *et al*, 2012)
E107A	NuoCD interface,	normal	67%	58-71%	reduced	(Sinha *et al*, 2012)
P110A	N6a interface	normal	78%	9-19%	reduced	(Sinha *et al*, 2012)
I114A	near N6a and N6b	NA	94%	107-121%	100%	(Sinha *et al*, 2012)
E121A	surface	normal	99%	74-85%	80%	(Sinha *et al*, 2012)
Y132A	surface	NA	123%	84-105%	100%	(Sinha *et al*, 2012)

b) Mutations in the membrane arm

Mutation	Amino acid location	Expression/ assembly	Effect			Reference and comments
			FMN-site activity	Oxido-reductase activity	Proton pumping activity	
NuoL						
D82A	β-hairpin - TM8 SB	normal	90%	90%	80%	(Nakamaru-Ogiso *et al*, 2010)
D82N		normal	90%	75%	80%	(Nakamaru-Ogiso *et al*, 2010)
D134N	surface, interacts with β-hairpin	normal	110%	110%	70%	(Nakamaru-Ogiso *et al*, 2010)
E144A	TM5, central axe, interface with NuoM	reduced	90%	20%	30%	(Nakamaru-Ogiso *et al*, 2010)
E144Q		normal	105%	15%	10%	(Nakamaru-Ogiso *et al*, 2010)
K169C	TM6 – HL SB	normal	87%	65%	reduced	(Michel *et al*, 2011)

K169E		normal	117%	67%	reduced	(Michel *et al*, 2011)
K169R		normal	118%	94%	reduced	(Michel *et al*, 2011)
D178A	TM6, central axe	normal	125%	95%	80%	(Nakamaru-Ogiso *et al*, 2010)
D178N		normal	125%	70%	50%	(Nakamaru-Ogiso *et al*, 2010), (EIPA insensitive)
R175A	NuoL/M interface	reduced	60%	17%	low	(Sato *et al*, 2013), (EIPA insensitive)
K229A	TM7, central axe	low	60%	10%	NA	(Nakamaru-Ogiso *et al*, 2010)
K229R		normal	125%	30%	NA	(Nakamaru-Ogiso *et al*, 2010)
K229E		low	100%	20%	NA	(Nakamaru-Ogiso *et al*, 2010)
P234A	Intermembrane surface	reduced	85%	67%	reduced	(Sato *et al*, 2013)
Q236H	TM7b – HL HB	normal	108%	86%	NA	(Michel *et al*, 2011)
Q236K		normal	99%	57%	low	(Michel *et al*, 2011)
Q236C		normal	106%	86%	NA	(Michel *et al*, 2011)
Q236E		normal	117%	84%	NA	(Michel *et al*, 2011)
W238A	TM7b - TM7a HB	normal	130%	80%	NA	(Nakamaru-Ogiso *et al*, 2010)
W238Y		low	100%	50%	NA	(Nakamaru-Ogiso *et al*, 2010)
W238C		low	90%	30%	NA	(Nakamaru-Ogiso *et al*, 2010)
D303A	surface, interacts with TM11-12 loop	normal	110%	110%	80%	(Nakamaru-Ogiso *et al*, 2010), (EIPA insensitive)
D303N		normal	115%	100%	80%	(Nakamaru-Ogiso *et al*, 2010)

H334A	central axe, proton channel	low	100%	50%	NA	(Nakamaru-Ogiso *et al*, 2010)
H334Q		normal	150%	120%	NA	(Nakamaru-Ogiso *et al*, 2010)
H338A		normal	110%	100%	NA	(Nakamaru-Ogiso *et al*, 2010)
H338Q		normal	100%	100%	NA	(Nakamaru-Ogiso *et al*, 2010)
K342A	central axe, proton channel	low	63%	11%	low	(Sato *et al*, 2013)
E359A	surface	normal	110%	100%	normal	(Nakamaru-Ogiso *et al*, 2010)
P390A	Intermembrane surface	reduced	90%	68%	NA	(Sato *et al*, 2013)
K399A	central axe, proton channel	low	105%	20%	NA	(Nakamaru-Ogiso *et al*, 2010)
K399E		low	100%	15%	NA	(Nakamaru-Ogiso *et al*, 2010)
D400A	central axe, proton channel, surface	normal	120%	70%	50%	(Nakamaru-Ogiso *et al*, 2010), (EIPA insensitive)
D400N		normal	105%	90%	70%	(Nakamaru-Ogiso *et al*, 2010)
D400E		normal	130%	100%	90%	(Nakamaru-Ogiso *et al*, 2010)
R431A	surface, near HL	low	60%	10%	NA	(Nakamaru-Ogiso *et al*, 2010)
R431H		normal	120%	100%	NA	(Nakamaru-Ogiso *et al*, 2010)

b) Mutations in the membrane arm, continuation

Mutation	Amino acid location		Effect	

		Expression/ assembly	FMN-site activity	Oxido-reductase activity	Proton pumping activity	Reference and comments
NuoL						
R529C	Horizontal amphipathic helix	NA	NA	86%	NA	(Steimle *et al*, 2015)
D542R		NA	61%	64%	102%	(Belevich *et al*, 2011)
D542N		NA	78%	91%	105%	(Belevich *et al*, 2011)
D542N		NA	NA	81%	77%	(Steimle *et al*, 2012)
D546N		NA	NA	93%	90%	(Steimle *et al*, 2012)
K551C		NA	NA	100%	NA	(Steimle *et al*, 2015)
K551Q		NA	89%	104%	96%	(Belevich *et al*, 2011)
K551E		NA	80%	98%	96%	(Belevich *et al*, 2011)
V550C		NA	NA	74%	NA	(Steimle *et al*, 2015)
P552A		NA	100%	121%	109%	(Belevich *et al*, 2011)
P552C		NA	88%	108%	103%	(Belevich *et al*, 2011)
P552Q		NA	96%	120%	98%	(Belevich *et al*, 2011)
F553C		NA	NA	104%	NA	(Steimle *et al*, 2015)
L554C		NA	NA	93%	NA	(Steimle *et al*, 2015)
L560C		NA	NA	50%	NA	(Steimle *et al*, 2015)
K561C		NA	NA	97%	NA	(Steimle *et al*, 2015)
R562C		NA	NA	81%	NA	(Steimle *et al*, 2015)
D563N		NA	NA	75%	49%	(Steimle *et al*, 2012)
D563E		NA	NA	106%	74%	(Steimle *et al*, 2012)
D563Q		NA	NA	81%	56%	(Steimle *et al*, 2012)
D563A		NA	NA	118%	82%	(Steimle *et al*, 2012)

N566C		NA	NA	69%	NA	(Steimle *et al*, 2015)
I571C		NA	NA	80%	NA	(Steimle *et al*, 2015)
P572C		NA	NA	119%	NA	(Steimle *et al*, 2015)
A573C		NA	NA	102%	NA	(Steimle *et al*, 2015)
V574C		NA	NA	86%	NA	(Steimle *et al*, 2015)
Y590C		NA	NA	102%	NA	(Steimle *et al*, 2015)
NuoM						
D84A	β-hairpin - TM8 SB	normal	92%	83%	NA	(Torres-Bacete *et al*, 2007)
D84N		normal	97%	89%	NA	(Torres-Bacete *et al*, 2007)
D135A	surface, interacts with β - hairpin	reduced	80%	44%	normal	(Torres-Bacete *et al*, 2007)
D135N		normal	86%	78%	normal	(Torres-Bacete *et al*, 2007)
D135E		normal	93%	87%	normal	(Torres-Bacete *et al*, 2007)
E144A	TM5, central axe, interface with NuoN	normal	103%	2-10%	0%	(Torres-Bacete *et al*, 2007)
E144Q		normal	98%	2%	0%	(Torres-Bacete *et al*, 2007)
E144D		normal	100%	89-100%	normal	(Torres-Bacete *et al*, 2007)
E144A/M145E		normal	86%	3-10%	0%	(Torres-Bacete *et al*, 2009)
E144A/W143E		normal	91%	13-15%	15%	(Torres-Bacete *et al*, 2009)

b) Mutations in the membrane arm, continuation

Mutation	Amino acid location	Expression/ assembly	Effect			Reference and comments
			FMN-site activity	Oxido-reductase activity	Proton pumping activity	
NuoM						
E144A/V148E	TM5, central axe, interface with NuoN	normal	113%	3-12%	0%	(Torres-Bacete *et al*, 2009)
E144A/F140E		normal	88%	39-60%	60%	(Torres-Bacete *et al*, 2009)
E144A/F152E		normal	85%	3-11%	0%	(Torres-Bacete *et al*, 2009)
E144A/F141E		normal	99%	3-15%	0%	(Torres-Bacete *et al*, 2009)
E144A/L147E		normal	95%	30-45%	50%	(Torres-Bacete *et al*, 2009)
E144A/F139E		normal	104%	2-15%	0%	(Torres-Bacete *et al*, 2009)
E144A/F142E		normal	103%	3-15%	0%	(Torres-Bacete *et al*, 2009)
E144A/M146E		normal	86%	2-11%	0%	(Torres-Bacete *et al*, 2009)
E144A/P149E		normal	88%	4-10%	0%	(Torres-Bacete *et al*, 2009)
E144A/M150E		normal	93%	2-12%	0%	(Torres-Bacete *et al*, 2009)
E144A/Y151E		normal	98%	3-10%	0%	(Torres-Bacete *et al*, 2009)
E144A/L153E		normal	91%	3-13%	0%	(Torres-Bacete *et al*, 2009)
E144A/V127E		normal	95%	3%	0%	(Torres-Bacete *et al*, 2009)
E144A/I128E		normal	109%	3-13%	0%	(Torres-Bacete *et al*, 2009)

E144A/G129E		normal	85%	5-12%	0%	(Torres-Bacete *et al*, 2009)
E144A/I189E		normal	111%	3-13%	0%	(Torres-Bacete *et al*, 2009)
E144A/L190E		normal	98%	2-12%	0%	(Torres-Bacete *et al*, 2009)
E144A/A191E		normal	85%	3-12%	0%	(Torres-Bacete *et al*, 2009)
K173C	TM6 – HL HB	normal	78%	70%	reduced	(Michel *et al*, 2011)
K173E		normal	102%	50%	reduced	(Michel *et al*, 2011)
K173R		normal	94%	91%	reduced	(Michel *et al*, 2011)
H196A	surface	normal	99%	79%	NA	(Torres-Bacete *et al*, 2007)
K234A	TM7, central axe	normal	91%	5-10%	low	(Torres-Bacete *et al*, 2007)
K234R		reduced	65%	5-20%	low	(Euro *et al*, 2008a)
H241A	TM7b - HL interaction	normal	90%	88%	NA	(Torres-Bacete *et al*, 2007)
H241E		normal	79%	71%	NA	(Michel *et al*, 2011)
H241K		normal	94%	40%	low	(Michel *et al*, 2011)
H241R		normal	81%	46%	NA	(Michel *et al*, 2011)
W243A	TM7b – TM7a HB	normal	97%	103%	normal	(Torres-Bacete *et al*, 2007)
W243Y		normal	108%	104%	NA	(Torres-Bacete *et al*, 2007)
P245A	TM7b surface	normal	100%	102%	NA	(Torres-Bacete *et al*, 2007)
K265A	TM8, central axe	normal	96%	35-80%	low	(Torres-Bacete *et al*, 2007)

R273A	TM8 - β-hairpin SB	normal	105%	92%	NA	(Torres-Bacete *et al*, 2007)
H322A	central axe	normal	99%	100%	NA	(Torres-Bacete *et al*, 2007)
H322A		NA	NA	61%	74-92%	(Mühlbauer *et al*, 2020)
H348A	central axe	normal	100%	92%	NA	(Torres-Bacete *et al*, 2007)
H348A		NA	NA	82%	61-79%	(Mühlbauer *et al*, 2020)
H322/H348		NA	NA	48	40-63%	(Mühlbauer *et al*, 2020)
R365A	surface, SB TM11-TM14	normal	85%	87%	NA	(Torres-Bacete *et al*, 2007)
R369H	surface, HB to TM7b	normal	94%	63-68%	normal	(Torres-Bacete *et al*, 2007)
P399A	Broken helix at M/L interface	normal	78%	52-63%	reduced	(Sato *et al*, 2013)
E407A	TM12, NuoL interface, central axe	normal	65%	5-8%	low	(Sato *et al*, 2013)
Y435A	proton channel	normal	99%	94	NA	(Torres-Bacete *et al*, 2007)

b) Mutations in the membrane arm, continuation

Mutation	Amino acid location	Expression/ assembly	Effect			Reference and comments
			FMN-site activity	Oxido-reductase activity	Proton pumping activity	
NuoN						
M74K	surface	normal	NA	90%	100%	(Amarneh & Vik, 2003)
C88S	TM3, interior	normal	NA	100%	100%	(Amarneh & Vik, 2003)
C88V		normal	NA	100%	100%	(Amarneh & Vik, 2003)
E104C	surface, near NuoJ	normal	NA	90%	100%	(Amarneh & Vik, 2003)

E133A	E-channel / interface with NuoK	normal	NA	70%	100%	(Amarneh & Vik, 2003)
E133C		normal	NA	70%	100%	(Amarneh & Vik, 2003)
E133D		normal	NA	80%	100%	(Amarneh & Vik, 2003)
E133A		normal	102%	88%	100%	(Sato *et al*, 2013)
E133A/ $_{K}$E72A		normal	71%	19%	low	(Sato *et al*, 2013)
R151C	surface	normal	NA	90%	100%	(Amarneh & Vik, 2003)
E154C	surface, interacts with NuoK N-terminus	normal	NA	70%	90%	(Amarneh & Vik, 2003)
K158C	TM6 – HL HB	normal	86%	50%	80%	(Michel *et al*, 2011)
K158R		normal	71%	70%	80%	(Michel *et al*, 2011)
K158E		normal	95%	47%	reduced	(Michel *et al*, 2011)
K158A		normal	79	57	90%%	(Sato *et al*, 2013)
K158R		normal	75	41	reduced	(Sato *et al*, 2013)
T160I	interface with NuoK	normal	NA	80%	NA	(Amarneh & Vik, 2003)
K217C	TM7, central axe	No expression	NA	NA	NA	(Amarneh & Vik, 2003)
K217R		normal	NA	40%	80%	(Amarneh & Vik, 2003)
K217A		normal	92	55%	reduced	(Sato *et al*, 2013)
K217C		NA	102	57	100%	(Sato *et al*, 2013)
K217R		normal	87	44	reduced	(Sato *et al*, 2013)
P222A	TM7 broken HL	normal	91	77	90%	(Sato *et al*, 2013)
H224A	TM7b - HL interaction	normal	NA	100%	NA	(Amarneh & Vik, 2003)
H224A		normal	90	73	80%	(Sato *et al*, 2013)
H224Y		normal	NA	90%	NA	(Amarneh & Vik, 2003)
H224K		normal	95	37-40%	reduced	(Amarneh & Vik, 2003;

						Michel *et al*, 2011)
H224E		normal	100%	67%	NA	(Michel *et al*, 2011)
H224R		normal	69%	32%	NA	(Michel *et al*, 2011)
W226C	TM7b – TM7a HB	normal	NA	90%	100%	(Amarneh & Vik, 2003)
D229C	TM7b – LTM16 and HL interaction	normal	NA	70%	100%	(Amarneh & Vik, 2003)
K247C	TM8, central axe	normal	NA	0-7%	50%	(Amarneh & Vik, 2003)
K247R		normal	NA	80%	100%	(Amarneh & Vik, 2003)
K247R		NA	101%	94%	normal	(Sato *et al*, 2013)
K247A		NA	70%	32%	reduced	(Sato *et al*, 2013)
K295C	surface	normal	NA	80%	70%	(Amarneh & Vik, 2003)
K295R		normal	NA	90%	80%	(Amarneh & Vik, 2003)
Y300C	TM10 - HL interaction	normal	NA	70%	80%	(Amarneh & Vik, 2003)
Y300S		normal	NA	50%	80%	(Amarneh & Vik, 2003)

b) Mutations in the membrane arm, continuation

Mutation	Amino acid location	Expression/ assembly	Effect			Reference and comments
			FMN-site activity	Oxido-reductase activity	Proton pumping activity	
NuoN						
P387A		normal	83%	52%	reduced	(Sato *et al*, 2013)
P387G		NA	103%	91%	90%	(Sato *et al*, 2013)
G391S	near central axe	normal	NA	90%	NA	(Amarneh & Vik, 2003)
K395C	TM12, central axe, NuoM interface	normal	NA	5%	NA	(Amarneh & Vik, 2003)
K395R		normal	NA	30%	NA	(Amarneh & Vik, 2003)

K395R		NA	90%	37%	reduced	(Sato *et al*, 2013)
K395A		normal	71%	4%	low	(Sato *et al*, 2013)
Y424C	intermembrane	normal	NA	90%	NA	(Amarneh & Vik, 2003)
V469A	Intermembrane surface	normal	88%	72%	90%	(Sato *et al*, 2013)
M482C	Surface/ N/M interface	normal	NA	100%	NA	(Amarneh & Vik, 2003)
NuoK						
F15A	interface with NuoN	normal	102%	90%	NA	(Kao *et al*, 2005b)
G21V	interface with NuoJ	normal	99%	61%	NA	(Kao *et al*, 2005b)
R25A	HBs to backbone of JTM1 C-terminus	normal	101%	26%	30%	(Kao *et al*, 2005b)
R25A		reduced	85%	54-73%	70%	(Torres-Bacete *et al*, 2012)
R25K		normal	97%	28-31%	NA	(Kao *et al*, 2005b)
R25K		reduced	97%	58-78%	100%	(Torres-Bacete *et al*, 2012)
R25C		reduced	97%	53-81%	90%	(Torres-Bacete *et al*, 2012)
R25S		reduced	915	51-64%	70%	(Torres-Bacete *et al*, 2012)
R26A	surface, near LTM16	normal	91%	39%	40%	(Kao *et al*, 2005b)
R26A		reduced	98%	55-73%	90%	(Torres-Bacete *et al*, 2012)
R26K		normal	96%	100%	NA	(Kao *et al*, 2005b)
R26K		normal	106%	80-95%	100%	(Torres-Bacete *et al*, 2012)
R25A/R26A		normal	90%	14%	30%	(Kao *et al*, 2005b)
N27C	surface, near LTM16	normal	95%	67-82%	70%	(Torres-Bacete *et al*, 2012)

N27S		normal	75%	59-67%	70%	(Torres-Bacete *et al*, 2012)
E36A	E-channel, central axe, interface with NuoJ	normal	95%	1-7%	NA	(Kao *et al*, 2005b)
E36A		normal	86%	3-7%	0%	(Torres-Bacete *et al*, 2012)
E36A/M31E		normal	94%	3-5%	0%	(Torres-Bacete *et al*, 2012)
E36A/L32E		normal	101%	52-56%	70%	(Torres-Bacete *et al*, 2012)
E36A/I33E		normal	85%	2-4%	0%	(Torres-Bacete *et al*, 2012)
E36A/G34E		normal	94%	2-3%	0%	(Torres-Bacete *et al*, 2012)
E36A/L35E		normal	80%	2-7%	0%	(Torres-Bacete *et al*, 2012)
E36A/I37E		normal	98%	3-8%	0%	(Torres-Bacete *et al*, 2012)
E36A/M38E		normal	87%	65-69%	90%	(Torres-Bacete *et al*, 2012)
E36A/I39E		normal	101%	49-75%	70%	(Torres-Bacete *et al*, 2012)
E36A/N40E		normal	78%	47-77%	65%	(Torres-Bacete *et al*, 2012)
E36A/A41E		normal	79%	3-4%	0%	(Torres-Bacete *et al*, 2012)
E36Q		normal	99%	3-8%	0%	(Kao *et al*, 2005b; Kervinen *et al*, 2004)
E36D		normal	118%	120%	normal	(Kervinen *et al*, 2004)
E36Q/E72Q		normal	64%	5%	impaired	(Kervinen *et al*, 2004)
E36Q/I39D		normal	76%	21%	impaired	(Kervinen *et al*, 2004)

E36Q/A69D		normal	52%	91%	normal	(Kervinen *et al*, 2004)

b) Mutations in the membrane arm, continuation

Mutation	Amino acid location	Expression/ assembly	Effect			Reference and comments
			FMN-site activity	Oxido-reductase activity	Proton pumping activity	
NuoK						
I39D	interface with NuoN	normal	84%	140%	normal	(Kervinen *et al*, 2004)
A69D	opposite E36	normal	60%	119%	normal	(Kervinen *et al*, 2004)
E72A	E-channel, central axe, interface with NuoN	normal	103%	43-48%	~50%	(Kao *et al*, 2005b)
E72Q		normal	99%	22-77%	~20%	(Kao *et al*, 2005b; Kervinen *et al*, 2004)
E72Q/I39D		normal	54%	180%	normal	(Kervinen *et al*, 2004)
E72Q/A69D		normal	92%	77%	impaired	(Kervinen *et al*, 2004)
E72Q/G34D		normal	60%	77%	impaired	(Kervinen *et al*, 2004)
E36Q/I39D A69D/E72Q		normal	118%	200%	impaired	(Kervinen *et al*, 2004)
E72A		normal	105%	47-52%	70%	(Torres-Bacete *et al*, 2012)
E72A/S67E		reduced	94%	11-16%	5%	(Torres-Bacete *et al*, 2012)
E72A/L68E		normal	84%	57-73%	70%	(Torres-Bacete *et al*, 2012)
E72A/A69E		normal	114%	63-74%	90%	(Torres-Bacete *et al*, 2012)
E72A/A71E		normal	94%	23-30%	30%	(Torres-Bacete *et al*, 2012)

E72A/A73E		normal	77%	28-39%	60%	(Torres-Bacete *et al*, 2012)
E72A/I75E		reduced	99%	54-62%	70%	(Torres-Bacete *et al*, 2012)
E72A/G75E		normal	106%	68-84%	100%	(Torres-Bacete *et al*, 2012)
E72A/L77E		normal	93%	45-55%	50%	(Torres-Bacete *et al*, 2012)
R85A	surface	normal	104%	100%	NA	(Kao *et al*, 2005b)
R85K		normal	103%	98%	NA	(Kao *et al*, 2005b)
R87A	surface	normal	103%	99%	NA	(Kao *et al*, 2005b)
R87K		normal	106%	100%	NA	(Kao *et al*, 2005b)
NuoA						
K46A	TM1-TM2 loop, not in the structure	normal	100%	94-100%	NA	(Kao *et al*, 2004)
E51A	TM1-TM2 loop, not in the structure	normal	97%	30%	NA	(Kao *et al*, 2004)
D79A	interface between JTM3 and NuoH, central axe	normal	102%	86-95%	NA	(Kao *et al*, 2004)
D79N		normal	99%	37-44%	NA	(Kao *et al*, 2004)
E81A	interface with NuoJ	normal	96%	36-42%	NA	(Kao *et al*, 2004)
E81Q		normal	94%	50-77%	NA	(Kao *et al*, 2004)
D79N/E81Q		normal	98%	2-10%	NA	(Kao *et al*, 2004)
NuoJ						
Y59C	TM3, interface with NuoK	normal	95%	57-93%	normal	(Kao *et al*, 2005a; Pätsi *et al*, 2008)
Y59F		normal	102%	44-50%	normal	(Kao *et al*, 2005a; Pätsi *et al*, 2008)
G61V	TM3 kink, interface with NuoA	normal	99%	48-53%	normal	(Kao *et al*, 2005a)
G61L		normal	98%	69-72%	normal	(Kao *et al*, 2005a)

M64V	TM3	normal	96%	78-89%	60%	(Kao *et al*, 2005a; Pätsi *et al*, 2008)
M64C		normal	105%	47%	NA	(Pätsi *et al*, 2008)
M64I		normal	107%	100%	normal	(Kao *et al*, 2005a)
V65G	near TM3 kink, interface with NuoA	normal	99%	3-13%	0%	(Kao *et al*, 2005a; Pätsi *et al*, 2008)
V65L		normal	98%	21-23%	reduced	(Kao *et al*, 2005a)
F67A	TM3	normal	86%	85%	100%	(Kao *et al*, 2005a)

b) Mutations in the membrane arm, continuation

Mutation	Amino acid location	Expression/ assembly	Effect			Reference and comments
			FMN-site activity	Oxido-reductase activity	Proton pumping activity	
NuoJ						
M72V	interface with NuoA/H	NA	100%	38%	NA	(Pätsi *et al*, 2008)
M72A		NA	170%	126%	NA	(Pätsi *et al*, 2008)
M72C		NA	92%	48%	NA	(Pätsi *et al*, 2008)
M64V/M72A		NA	57%	53%%	NA	(Pätsi *et al*, 2008)
E80Q	surface, interacts with NuoK	normal	101%	100%,	normal	(Kao *et al*, 2005a)
E80A		normal	102%	90%	reduced	(Kao *et al*, 2005a)
Y109F	surface	NA	89%	112%	NA	(Pätsi *et al*, 2008)
NuoH						
E36D	Q-site	normal	74%	52-77%	reduced	(Sinha *et al*, 2009)
E36D		NA	91%	57%	NA	(Pätsi *et al*, 2012)
E36A		normal	86%	20-27	reduced	(Sinha *et al*, 2009)
E36K		normal	13%	7%	BA	(Pätsi *et al*, 2012)
E36Q		reduced	32%	18%	NA	(Pätsi *et al*, 2012)
R37A		no assembly	19%	1-3%	0%	(Sinha *et al*, 2009)

R37K	Q-site, NuoB interface	low	36%	8-13	0%	(Sinha *et al*, 2009)
Q44A	Loop1, Q site	normal	121%	45-54	NA	(Sinha *et al*, 2009)
R46A	Loop1, Q site	no assembly	47%	6-13	0%	(Sinha *et al*, 2009)
R46K		low	62%	23-42	reduced	(Sinha *et al*, 2009)
P49A	Loop1	normal	78%	39-83	reduced	(Sinha *et al*, 2009)
D63E	Q-site (entry)	normal	112%	91-93	normal	(Sinha *et al*, 2009)
D63E		NA	78%	35%	NA	(Pätsi *et al*, 2012)
D63A		no assembly	18%	1-2%	0%	(Sinha *et al*, 2009)
D63N		no assembly	24%	1%	0%	(Sinha *et al*, 2009)
M64T	Q-entry	NA	86%	90%	NA	(Pätsi *et al*, 2012)
K70A	NuoB interface	normal	111%	93-98%	NA	(Sinha *et al*, 2009)
E71A	NuoB, A interface	normal	72%	37-65%	reduced	(Sinha *et al*, 2009)
G134A	near TM5-6 loop	normal	69%	70%	normal	(Sinha *et al*, 2009)
G134L		no assembly	44%	7-18%	0%	(Sinha *et al*, 2009)
G134V		low	54%	3-6%	0%	(Sinha *et al*, 2009)
S137A	near TM5-6 loop	normal	90%	72-76%	NA	(Sinha *et al*, 2009)
G145A	near TM5-6 loop	normal	114%	70-75%	normal	(Sinha *et al*, 2009)
G145V		low	50%	4-5%	0%	(Sinha *et al*, 2009)
R148A	NuoCD interface, SB with $_{c}$E240	low	49%	4-18	0%	(Sinha *et al*, 2009)
S155A	NuoA interface	normal	95%	95-104%	NA	(Sinha *et al*, 2009)
Y156A	Intermembrane surface	normal	80%	51-61%	reduced	(Sinha *et al*, 2009)
E157A	E-channel, Central axe	normal	95%	24-29%	reduced	(Sinha *et al*, 2009)
E157K		normal	160%	80-111%	NA	(Sinha *et al*, 2009)
I201V		NA	105%	67%	NA	(Pätsi *et al*, 2012)

I201T	Intermembrane surface	NA	65%	59-72%	NA	(Hinttala *et al*, 2010)

b) Mutations in the membrane arm, continuation

Mutation	Amino acid location	Expression/ assembly	Effect			Reference and comments
			FMN-site activity	Oxido-reductase activity	Proton pumping activity	
NuoH						
V206G	Deep Intermembrane	normal	110%	95-100%	NA	(Sinha *et al*, 2009)
V206E		NA	67%	63%	NA	(Kervinen *et al*, 2006)
R209F	Intermembrane, points into Q site	NA	59%	43%	NA	(Kervinen *et al*, 2006)
R209A		normal	85%	63-87%	reduced	(Sinha *et al*, 2009)
H210T	Q-site, NuoCD interface	NA	82%	63%	NA	(Kervinen *et al*, 2006)
D213A	Q-site, NuoCD interface	low	53%	12%	0%	(Sinha *et al*, 2009)
D213E		NA	69%	43%	NA	(Kervinen *et al*, 2006)
D213N		NA	95%	71%	NA	(Kervinen *et al*, 2006)
E216A	TM5-6 loop "beginning"	normal	112%	63-80%	NA	(Sinha *et al*, 2009)
E216A		NA	95%	80%	NA	(Kervinen *et al*, 2006)
E218A	TM5-6 loop	normal	124%	35-43%	reduced	(Sinha *et al*, 2009)
E220A	TM5-6 loop	no assembly	49%	2%	0%	(Sinha *et al*, 2009)
E220Q		no assembly	45%	1%	0%	(Sinha *et al*, 2009)
E228A	TM5-6 loop, NuoA, B loop interface	no assembly	40%	8%	0%	(Sinha *et al*, 2009)
E228Q		no assembly	32%	1%	0%	(Sinha *et al*, 2009)
E228D		NA	79%	43%	NA	(Kervinen *et al*, 2006)
Y229H		NA	71%	39%	NA	(Kervinen *et al*, 2006)

E241A	E-channel, Q-site, Central axe	normal	94%	59-64%	NA	(Sinha *et al*, 2009)
E241Q		normal	95%	58-60%	NA	(Sinha *et al*, 2009)
R286A	Q site	normal	79%	67-63%	reduced	(Sinha *et al*, 2009)
R291A	Q-site	normal	126%	85-90%	NA	(Sinha *et al*, 2009)
L289C	NuoI_H1 interface	NA	74%	74-81%	NA	(Maliniemi *et al*, 2009)
R291M	near TM5-6 loop	NA	57%	57-107%	105%	(Maliniemi *et al*, 2009)
R293M	Surface	NA	46%	46-63%	NA	(Maliniemi *et al*, 2009)
Y294L	100_{DQ}/89 NADH-O_2	NA	89%	89-100%	NA	(Maliniemi *et al*, 2009)
D295A	Surface, salt bridge with R291	NA	39%	39-52%	NA	(Maliniemi *et al*, 2009)
D295A		low	58%	16-20%	0%	(Sinha *et al*, 2009)
D295E		normal	98%	41-49%	reduced	(Sinha *et al*, 2009)
Q296T	NuoCD interface, Interacts with lipid near $_{I}$H1	NA	63%	63-100%	NA	(Maliniemi *et al*, 2009)
V297P	Interacts with lipid near $_{I}$H1	NA	62%	62-97%	NA	(Maliniemi *et al*, 2009)
L289C/V297P	near $_{I}$H1	NA	36%	36-73%	NA	(Maliniemi *et al*, 2009)
G301C	Intermembrane surface	NA	72%	72-90%	NA	(Maliniemi *et al*, 2009)
W302L	NuoA interface	NA	62%	62%%	100%	(Maliniemi *et al*, 2009)
K303A	Surface/ NuoA interface	normal	71%	47-75%	reduced	(Sinha *et al*, 2009)
T316H	Surface	NA	73%	72%	NA	(Hinttala *et al*, 2010)