Molecular mechanisms of mitochondrial redox-coupled proton pumping enzymes

by

Domen Kampjut

September, 2020

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

The dissertation of Domen Kampjut, titled *"Molecular mechanisms of mitochondrial redox-coupled proton pumping enzymes"*, is approved by:

Supervisor: Leonid Sazanov, IST Austria, Klosterneuburg, Austria

Signature: _____

Committee Member: Martin Loose, IST Austria, Klosterneuburg, Austria

Signature: _____

Committee Member: David Haselbach, Institute of Molecular Pathology, Vienna, Austria

Signature: _____

Exam Chair: Florian Schur, IST Austria, Klosterneuburg, Austria

Signature: _____

signed page is on file

IST Austria Thesis, ISSN: 2663-337X
ISBN: 978-3-99078-008-4

I hereby declare that this dissertation 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: _____

Domen Kampjut

September 8, 2020

signed page is on file

Abstract

Mitochondria are sites of oxidative phosphorylation in eukaryotic cells. Oxidative phosphorylation operates by a chemiosmotic mechanism made possible by redox-driven proton pumping machines which establish a proton motive force across the inner mitochondrial membrane. This electrochemical proton gradient is used to drive ATP synthesis, which powers the majority of cellular processes such as protein synthesis, locomotion and signalling. In this thesis I investigate the structures and molecular mechanisms of two inner mitochondrial proton pumping enzymes, respiratory complex I and transhydrogenase. I present the first high-resolution structure of the full transhydrogenase from any species, and a significantly improved structure of complex I. Improving the resolution from 3.3 Å available previously to up to 2.3 Å in this thesis allowed us to model bound water molecules, crucial in the proton pumping mechanism. For both enzymes, up to five cryo-EM datasets with different substrates and inhibitors bound were solved to delineate the catalytic cycle and understand the proton pumping mechanism. In transhydrogenase, the proton channel is gated by reversible detachment of the NADP(H)-binding domain which opens the proton channel to the opposite sites of the membrane. In complex I, the proton channels are gated by reversible protonation of key glutamate and lysine residues and breaking of the water wire connecting the proton pumps with the quinone reduction site. The tight coupling between the redox and the proton pumping reactions in transhydrogenase is achieved by controlling the NADP(H) exchange which can only happen when the NADP(H)-binding domain interacts with the membrane domain. In complex I, coupling is achieved by cycling of the whole complex between the closed state, in which quinone can get reduced, and the open state, in which NADH can induce quinol ejection from the binding pocket. On the basis of these results I propose detailed mechanisms for catalytic cycles of transhydrogenase and complex I that are consistent with a large amount of previous work. In both enzymes, conformational and electrostatic mechanisms contribute to the overall catalytic process. Results presented here could be used for better understanding of the human pathologies arising from deficiencies of complex I or transhydrogenase and could be used to develop novel therapies.

Acknowledgments

My deepest thanks go to my supervisor Prof Leonid Sazanov FRS for his support during my PhD. I have learned and done more than I could have hoped for in the last four years and none of it would have been possible without Leo's guidance and insights. I am happy and honoured to have been able to see and participate in the rapid advances that the structural bioenergetics field has made since I joined Leo's group in Cambridge as a summer student in 2014 and especially since I started my PhD in 2016.

I would also like to thank my colleagues over the years. Dr James Letts taught me many laboratory techniques and principles of single particle analysis. Karol Fiedorczuk helped me with the initial purification of ovine complex I, preparation of cryo-EM grids and operation of the T12 microscope. I am also indebted to Dr Irene Vercellino for critical discussions of numerous versions of enzyme mechanisms and many ambiguous results. Finally, I am grateful to other members of the Sazanov group and friends in Vienna for their continued support during the highs and lows of the last four years, especially Igor Asanović, Kristina Lukić, Gergely Pinke, Julia Steiner and Vladyslav Kravchuk.

I acknowledge the support of IST facilities, especially the Electron Miscroscopy facility for providing training and resources. Special thanks also go to cryo-EM specialists who helped me to collect the data present here: Dr Valentin Hodirnau (IST Austria), Dr Tom Heuser (IMBA, Vienna), Dr Rebecca Thompson (Uni. of Leeds) and Dr Jirka Nováček (CEITEC). This work has been supported by iNEXT, project number 653706, funded by the Horizon 2020 programme of the European Union. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 665385.

Finally, I am grateful for my family who have always supported me in my curiosity and ambition to become a scientist. Mum and Dad, I would not be the person that I am without you. Lori, my love, thank you for everything and especially for choosing to make Vienna your home with me during a large part of my PhD. I am lucky to have you in my life.

This thesis is dedicated to Barbara, Gregor and Lori.

About the Author

Domen Kampjut completed his B.A. and M.Sci. in Natural Sciences at the University of Cambridge. During his studies he worked with Dr Markus Ralser on the evolution of the Krebs cycle and with Prof John Walker on the mechanism of mitochondrial F_1F_o-ATPase. Domen joined IST Austria in September 2016 to work with Prof Leonid Sazanov on the structure and function of mitochondrial proton pumping machines and to learn electron cryo-microscopy. His results formed the basis of two first-author publications and were presented at several Austrian and international symposia and conferences. His main interests are the structure and function of membrane protein complexes and mitochondria.

List of Publications Appearing in Thesis

Some of the work presented in this thesis has been published elsewhere.

Parts of Chapter 3, as well as subchapters 1.4, 2.3, 2.4, 2.5, 2.7 and 2.9, have been published here:

> **Kampjut, D**. and Sazanov, L.A., 2019. Structure and mechanism of mitochondrial proton-translocating transhydrogenase. *Nature*, *573*(7773), pp.291-295.

Parts of Chapter 4, as well as subchapters 1.6, 2.3, 2.4, 2.6 and 2.8, have been submitted for publication elsewhere:

> **Kampjut, D**. and Sazanov, L.A. 2020. The coupling mechanism of mammalian respiratory complex I. *(submitted)*

"There is grandeur in this view of life, with its several powers, having been originally breathed into a few forms or into one; and that, whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved."

- Charles Darwin, On the Origin of Species

Table of Contents

List of Figures

List of Tables

List of Symbols/Abbreviations

Å	Ångstrom
ATP	Adenosine triphosphate
AU	Absorption units
BCA	Bichinconic acid
BN	Blue native
BSA	Bovine serum albumin
CHAPS	3-[(3-cholamidopropyl)-dimethylammonio]-1-propane sulphonate
CL	Cardiolipin
CMC	Critical micellar concentration
Cryo-EM	electron cryo-microscopy
CTF	Contrast transfer function
DDM	*n*-dodecyl-β-maltoside
DMSO	Dimethylsulfoxide
DOPC	1,2-dioleoyl-*sn*-glycero-3-phosphocholine
DQ	Decylubiquinone
DQE	Detector quantum efficiency
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ETC	Electron transport chain
FeCN	Potassium hexacyanoferrate (III) $K_3[Fe(CN)_6]$
FMN	Flavin mononucleotide
FOM	Fluorinated octyl maltoside
FSC	Fourier shell correlation
GDN	Glyco-diosgenin
GF	Gel filtration
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IMS	Intermembrane space
kDa	Kilodalton
LDH	Lactate dehydrogenase
LMNG	Lauryl maltoside neopentylglycol
MD	Membrane domain

MES	2-(N-Morpholino)ethanesulfonic acid
MS/MS	Tandem mass spectrometry
MTF	Modulation transfer function
NAD^+	Nicotinamide adenine dinucleotide, oxidised
NADH	Nicotinamide adenine dinucleotide, reduced
$NADP^+$	Nicotinamide adenine dinucleotide phosphate, oxidised
NADPH	Nicotinamide adenine dinucleotide phosphate, reduced
NEM	N-ethylmaleimide
NNT	Nicotinamide nucleotide transhydrogenase
OG	Octyl glucoside
PA	Peripheral arm
PAGE	Polyacrylamide gel electrophoresis
PC	Phosphatidylcholine
PDB	Protein data bank
PE	Phosphatidylethanolamine
PI	Phosphatidylinositol
P_i	Inorganic phosphate (PO_4^{3-})
PK	Pyruvate kinase
PMF	Proton motive force
PMSF	Phenylmethylsulphonyl fluoride
POPC	1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine
PS	Phosphatidylserine
Q	Quinone
QH_2	Quinol
RMSD	Root mean square deviation
ROS	Reactive oxygen species
SDS	Sodium dodecyl sulphate
SMP	Submitochondrial particles
SOD	Superoxide dismutase
TM	Transmembrane
TMH	Transmembrane helix
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol
UQ	Ubiquinone

1. Introduction

1.1. The chemiosmotic theory

All cellular life requires a source of energy to drive energetically unfavourable reactions, such as synthesis of peptides, maintenance of ion gradients and trafficking of subcellular components. The majority of these processes are coupled to the hydrolysis of ATP, which is highly favourable under the cellular conditions. Standard free energy change of the reaction ATP -> ADP + Pi is -30.5 kJ/mol at 1 M concentrations of substrates and pH = 7. In cells, concentrations of ATP, ADP and P_i are tightly regulated and because of the disequilibrium between them, the free energy of the above reaction is closer to -52 kJ/mol. ATP was first isolated in 1929 from muscle cells. In 1940 it was proposed to be the universal energy currency of life and in 1948 mitochondria were established as the sites of oxidative phosphorylation in the cell. It was not until 1961 that Peter Mitchell proposed how oxidative phosphorylation replenishes ATP in cells, and only by about 1975 did his hypothesis become widely appreciated. Unravelling the mechanism of oxidative phosphorylation can be placed alongside genetic code determination as one of the grand challenges of biology in the second half of the 20^{th} century.

Initially, a substrate-level phosphorylation was proposed to drive ATP formation, whereby a high-energy phosphorylated intermediate is produced in catabolism that can donate a phosphate group to ADP. This process does exist in glycolysis and provides a small amount of cellular ATP. The reaction is catalysed by phosphoglycerate kinase, 1,3-bisphosphoglycerate + ADP ⇌ glycerate 3-phosphate + ATP in glycolysis. However, no similar intermediate could be found in mitochondria. Secondly, ATP synthesis from ADP to Pi could be driven by a specialized enzyme via a conformational change mechanism. However, such an enzyme could not be isolated and it was not clear where the free energy to drive the ATP synthesis could come from. Finally, in 1961, Peter Mitchell proposed that ADP phosphorylation in mitochondria occurs by coupling this process to a proton gradient which exists across the mitochondrial inner membrane, as well as across thylakoid and bacterial cell membranes (Mitchell, 1961). This hypothesis requires the existence of an ion-impermeable membrane with two types of proton pumps embedded in it. It was widely debated in the 1960s and 1970s and eventually proven. In 1978, Peter Mitchell received a Nobel Prize for developing the theory.

One of the first conclusive proofs came from the Jagendorf experiments. Thylakoids were equilibrated at pH = 4 and then transferred to pH = 8, with ADP and P_i. This procedure

resulted in fast ATP production, even in the absence of light as normally required by thylakoid membranes (Uribe and Jagendorf, 1968).

The theory envisaged the existence of two types of membrane-embedded proton pumps. Redox-reaction driven proton pumps (NADH oxidase in Figure 1.1) use the free energy from oxidation reactions to pump protons across the membrane. In photosynthetic organisms, light is used to drive proton pumping. In both autotrophic and heterotrophic organisms, however, the ATP synthase couples the energetically favourable flow of protons down their electrochemical gradient to ATP synthesis by a unique rotary catalytic mechanism (Walker, 2012). ATP synthase is one of the most conserved and ancient proteins and it has been suggested that it was one of the first enzymes to have evolved, potentially even before the evolution of cellularity. In primordial life, an ancestral ATP synthase could have been involved in harvesting environmental proton gradients, such as those that exist in deep-sea hydrothermal vents (Martin, Sousa and Lane, 2014). This would evolutionary explain why all life uses proton (or sometimes ion) gradients in the energy metabolism and the conservation of the ATP synthase.

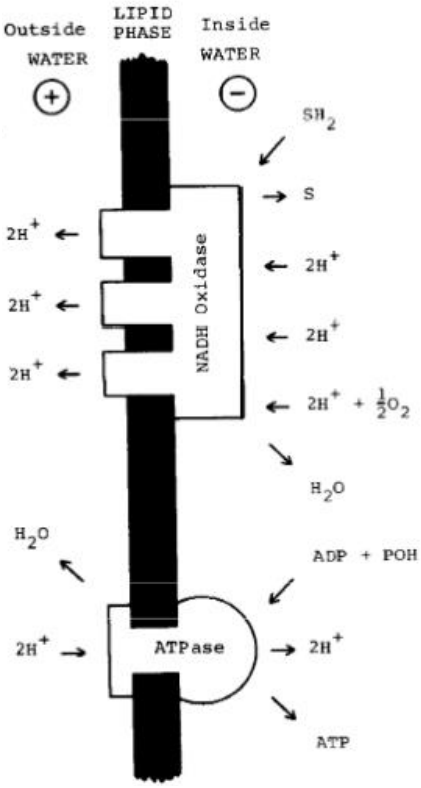


Figure 1.1 Mitchell's chemiosmotic hypothesis.

Energy-transducing membranes are impermeable to protons and contain two types of proton pumps. Redox-coupled proton pumps (NADH oxidase in the scheme) pump protons to the P-side (+) using the free energy from substrate oxidation reactions. ATP synthase (ATPase in the scheme) uses the resulting proton motive force for ATP synthesis.
Figure adapted from Mitchell, 1978

Proton motive force in bioenergetics refers to the difference in the proton electrochemical potential on each side of the membrane. Pmf has two thermodynamically interchangeable components, one due to the concentration difference of protons on either side of the membrane (ΔpH) and one due to the difference in the electric potential (Δψ): Pmf = Δψ - 59ΔpH. In most cases of bioenergetic membranes (with the exception of thylakoid), the majority of pmf comes from electric components.

Evolutionary, mitochondria and chloroplasts are direct successors of bacteria, as they were acquired ~ 2 billion years ago by the eukaryotic cell in the process of endosymbiosis. During the evolution, the symbiotic relationship tightened and modern-day mitochondria have relinquished most of their genes to the nucleus, but retained a handful of genes coding for components of the electron transport chain and gene expression machinery. The reasons for this are not entirely clear, however. Mitochondria are also important in cells for signalling, production of certain metabolites and have a crucial role in apoptosis.

1.2. The mitochondrial electron transport chain

If the primordial life may have depended on the external proton gradient to sustain ATP production, modern life generates these gradients using its own specialised machinery. In the case of mitochondria, the redox reactions result in proton pumping by respiratory complexes I, III and IV. Complex I catalyses the transfer of electrons from NADH to ubiquinone, complex III from ubiquinone to cytochrome c and complex IV from cytochrome c to oxygen. This creates a stepping gradient from the most reducing substrate with midpoint-electric potentials and a sequential release of energy and its storing in proton gradient. The main consumer of the proton gradient is the ATP synthase. Other consumers of the proton gradient include transhydrogenase, which maintains a high NADPH/$NADP^+$ ratio, metabolite carriers (phosphate carrier, Na^+ and Ca^{2+} cycling) and the uncoupling proteins. In mitochondria, some other metabolic reactions contribute to ubiquinone reduction as well, but they are not coupled to proton translocation across the membrane. They include complex II (succinate dehydrogenase), electron transfer flavoprotein-ubiquinone oxidoreductase (ETF-QO), GPDH and dihydroorotate dehydrogenase. In other organisms, especially in bacteria, this basic electron transfer chain framework can be expanded with several other possible entry points and branching pathways, depending on the substrates and environmental conditions.

The ETC complexes are large multi-subunit assemblies with many redox cofactors (Figure 1.2). Bacterial and eukaryotic versions of these proteins are very conserved, with

eukaryotic enzymes typically containing several supernumerary subunits which are not catalytically active but are important for additional layers of regulation or stability. Mitochondrial DNA has retained genes for several of the core subunits of ETC complexes, while all of the other genes, apart from a handful of genes connected with the mitochondrial protein synthesis have been relocated to the nucleus. The reason why genes for ETC complexes remained in mitochondria is not clear but could include high hydrophobicity of the proteins they encode or the need for faster *in situ* gene expression regulation in response to redox signals that cannot be achieved by retrograde signalling to the nucleus (Allen, 2015). Recently it was also discovered that complexes also exist in the form of supercomplexes (Schagger and Pfeiffer, 2000). Supercomplexes exist in varying compositions, depending on the organism and the tissue from which they come. The physiological function of them is currently the subject of intense investigation. The main hypothesis being debated is that they help to organise respiratory complexes in a tightly packed membrane, whereas the substrate channelling hypothesis has largely been disproven (Lobo-Jarne and Ugalde, 2018).

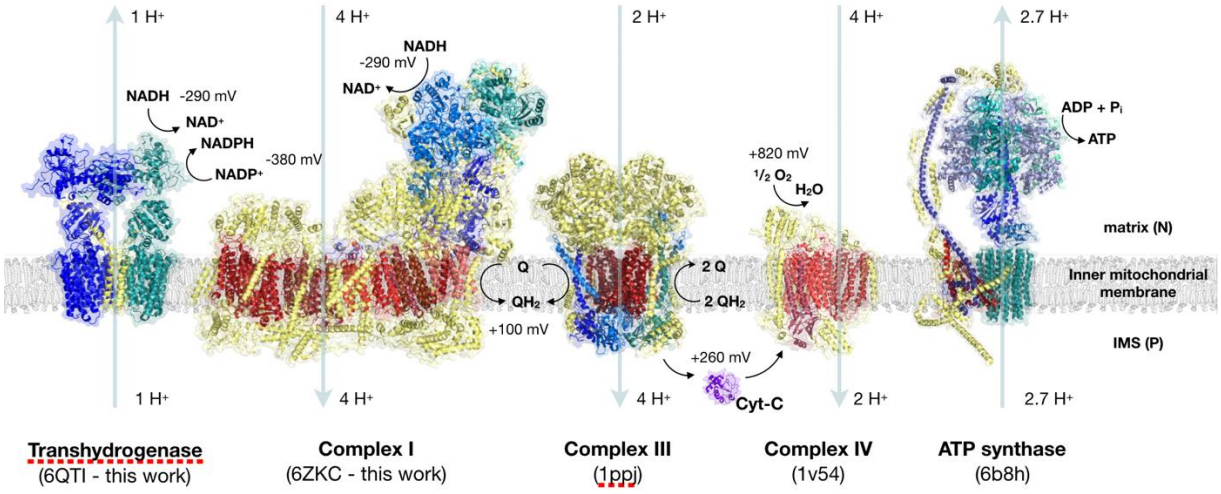


Figure 1.2. Mammalian mitochondrial proton pumping enzymes.

The structures of the mitochondrial proton pumping enzymes have recently become available. Proteins are coloured by subunits: red colours represent subunits encoded in the mitochondrial genome, blue colours are the core subunits conserved across all organisms but not encoded by the mtDNA and yellow colours are supernumerary subunits. Electric potentials of substrate-product pairs are indicated in the scheme for the typical ratios present in mitochondria (hence NADH and NADPH have different potentials).

Biochemically, all of the above complexes I, III and IV (as well as transhydrogenase) catalyse redox-coupled proton pumping but achieve it by radically different means, which reflect their distinct evolutionary ancestry. ATPase is not considered a redox-coupled proton

pump, because proton pumping in ATPase is linked to a phosphorylation reaction by a unique rotary coupling mechanism and will not be discussed further here. Complex I and transhydrogenase will be described in more detail in sections 1.5 and 1.3 respectively. Their structures became available only recently and the mechanism for their reactions is not firmly established yet. Complexes III and IV, however, have been structurally characterised over 20 years ago and their mechanisms are well understood. They are briefly introduced below to enable comparison with complex I and transhydrogenase in the Discussion.

Complex III catalyses electron transfer from ubiquinol to cytochrome c (Figure 1.3). It is the only one of the respiratory complexes which operates by a redox loop, originally envisaged by Mitchell. Mitchell's proposal stated that protons are taken up by a mobile carrier from the N-side of the membrane and then released to the P-side during oxidation of the mobile carrier. The stoichiometry of a redox loop is always 1 proton per 1 electron. Complex III is an obligate dimer with 11 subunits per monomer in mammals (Iwata *et al.*, 1998). The catalytic core subunits are conserved from bacteria and are necessary for the reaction (Smith, Fox and Winge, 2012). They include cytochrome b, which contains haemes b_H and b_L, cytochrome c, which contains haem c_1 and binds cytochrome c, and Rieske protein which contains an Fe_2S_2 cluster that transfers electrons between the two cytochromes. Complex III oxidises two molecules of quinol and reduces one molecule of quinone per catalytic cycle. In the first half of the cycle, one molecule of quinol is oxidised at the Q_o site near the IMS. The resulting two protons are released into the IMS, while the two electrons bifurcate: one goes towards the Q_i site via the two haems, while the other one goes via the FeS cluster of a mobile Rieske protein towards the cytochrome c. In the second half of the cycle, another molecule of quinol is oxidised at the Q_o site and the resulting protons and electrons follow the same pathways as in the first half. At the Q_i site, the two electrons are then used to reduce a molecule of quinone, with the two required protons originating from the matrix. Overall, this results in consumption of two protons from the matrix side, release of four protons to the IMS and reduction of two molecules of cytochrome c. The majority of the charge transfer comes from the movement of an electron from P-side to N-side, hence complex III pumps net one proton per each electron transferred, consistent also with the energy available from the midpoint electric potential difference between the quinone and cytochrome redox pairs.

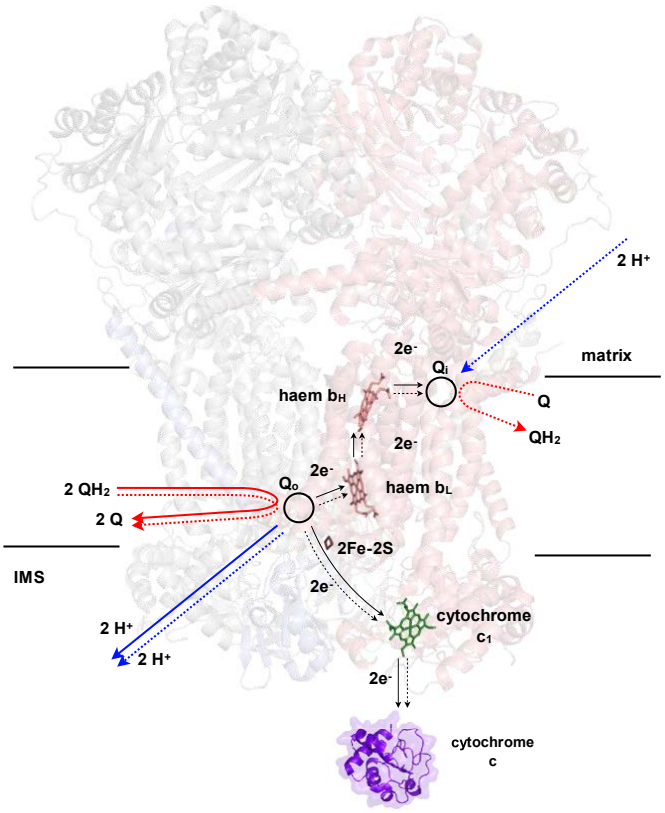


Figure 1.3. Complex III - bc1 complex.

Complex III transfers electrons from quinol to cytochrome c and pumps protons by a redox loop mechanism. One of the monomers is shown in grey and the other one in red (with the exception of the Rieske protein in blue). Redox centres are shown as sticks and the movements of electrons and substrates are shown in full arrows for the first half of the catalytic cycle and in dotted arrows for the second half of the catalytic cycle.
Figure based on PDB 1PPJ

Complex IV was the first redox-coupled proton pump for which the crystal structure was solved (Iwata *et al.*, 1995). Mammalian complex IV contains 13-14 subunits, while its bacterial counterparts contain three core subunits. All three are encoded by the mtDNA in mammals. COX1 contains a haem *a* group and a binuclear haem a_3-Cu_B centre (binuclear cluster, BNC), COX2 contains the Cu_A centre, while COX3 does not have any prosthetic groups. Complex IV catalyses the final electron transfer reaction between cytochrome c and dioxygen via the haems and the metal centres (Figure 1.4). Oxygen is bound by the BNC and its reduction occurs in four steps, during which tyrosine donates an electron and a proton for O-O cleavage, forming a tyrosyl radical. During each step, one “chemical” proton is taken up from the N-side for water formation, and one “pumped” proton is taken up to be pumped to the P-side of the membrane (Blomberg, 2016). Complex IV is thus both a redox loop and a direct proton pump. For the chemical protons, there is charge transfer against the PMF and their release as a part of a neutral water molecule to the same side of the membrane. Translocation

of the pumped protons occurs by vectorial proton transfer from the N-side to the P-side. The process of coupling between electron transfer and proton pumping is linked by almost purely Coulombic principles and does not depend on the protein conformational changes. Preventing slipping (re-release of the proton into the N-side of the membrane) is prevented by kinetic restriction (Wikström and Sharma, 2018).

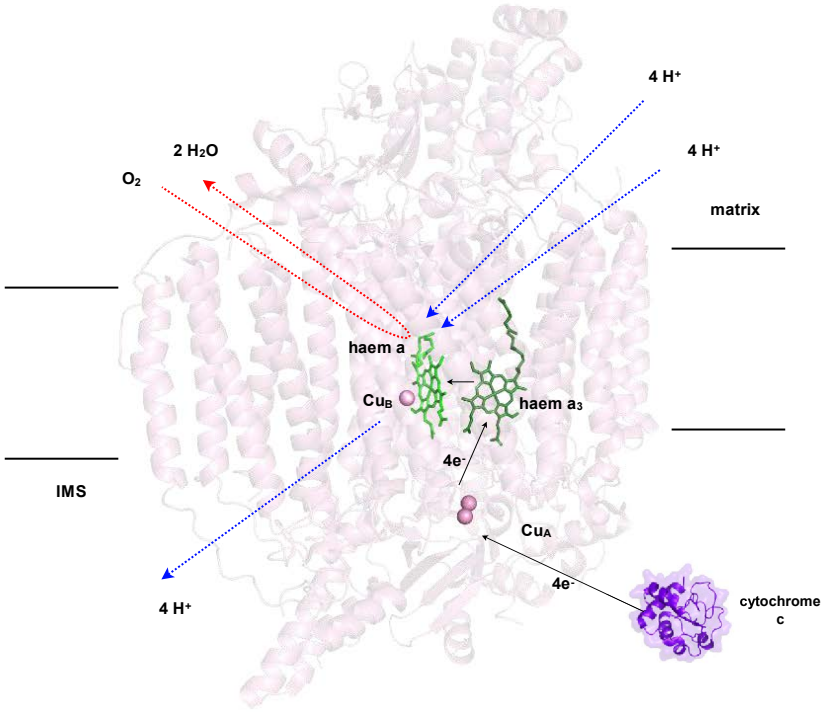


Figure 1.4. Complex IV - cytochrome c oxidase.

Complex IV is the final oxidase which transfers electrons from cytochrome c to oxygen. Details of the mechanism are described above. Complex IV pumps four protons by a direct mechanism and consumes further four protons from the matrix side for the formation of water.
Figure based on PDB 1V54

1.3. Transhydrogenase

Proton-translocating nicotinamide nucleotide transhydrogenase (NNT) is another redox-driven proton pump, which is found in many bacterial species and mitochondria. It catalyses a fully reversible transfer of hydride between NAD(H) and NADP(H) and pumps one proton for each hydride transferred:

$$NADH + NADP^+ + H^+_P <-> NAD^+ + NADPH + H^+_N$$

It has a unique feature of operating on the substrates with the same mid-point electric potential, although due to the different relative oxidized/reduced ratios of the two substrate pools, the electric potentials of NADPH and NADH differ *in vivo* (Figure 1.2) (Jackson *et al.*, 2015). The direction of the reaction depends on the relative ratios between the substrate and product concentrations and the proton motive force (PMF) across the membrane. The reaction from the left-hand side to the right-hand side is referred to as the "forward transhydrogenation" and the opposite reaction is referred to as "reverse transhydrogenation". The fact that forward transhydrogenation is coupled to proton pumping was established by experiments in which transhydrogenase was reconstituted into mixed liposomes with ATP synthase or bacteriorhodopsin. When either of these secondary proton pumps established proton motive force, forward transhydrogenation was increased up to 30 fold (Eytan, Eytan and Rydström, 1987; Eytans *et al.*, 1987). Transhydrogenase can enrich products over substrate by ~500 fold under the PMF typically present across the inner mitochondrial membrane (Jackson, 2012). In vitro, transhydrogenase can also catalyse "cyclic transhydrogenation", in which a reducing equivalent is transferred from NADH (or its analogue APADH) to a tightly bound $NADP^+$ and then to another molecule of NAD^+ (or $APAD^+$). In this process, no protons are pumped across the membrane. The reaction has a low pH optimum and is very unlikely to be of any physiological importance but it was used to investigate the mechanism and proton pumping of transhydrogenase (Jackson, 2012).

All of the known transhydrogenases have a similar obligate homodimeric structure. Each monomer is composed of three domains, NAD(H)-binding dI, transmembrane dII and NADP(H)-binding dIII, split between one, two or three polypeptides. The spatial arrangement of these domains has been revealed for the first time by the architecture of *T. thermophilus* NNT at 7 Å (Figure 1.5) (Leung *et al.*, 2015).

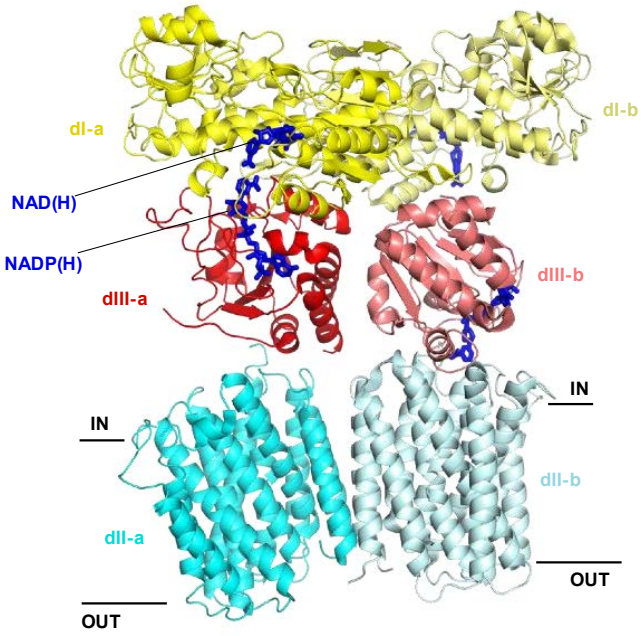


Figure 1.5. Architecture of T. thermophilus transhydrogenase.

Isolated domains of T. thermophilus NNT were crystallised and docked into a 7Å crystal structure of full transhydrogenase. This revealed the asymmetry of the dimer where dIII-a is attached to dI (face-up), while dIII-b is attached to dII (face-down). Hydride transfer is only possible in the face-up orientation.
Figure based on PDB 4O9U.

1.3.1. Evolutionary origin of transhydrogenase

Transhydrogenase can be found in all three domains of life. In Bacteria and Archaea it is localised in the plasma membrane and in Eukarya it is found in the inner mitochondrial membrane. It is present in 57% and 46% of eukaryotic and bacterial phyla, respectively, but only in 3% of archaeal phyla (Marreiros *et al.*, 2016). Within the eukaryotes there is a sharp distinction between animals, fungi and protists, which almost always contain transhydrogenase and higher plants, from which transhydrogenase is entirely absent. This is because in plants other NADPH-producing enzymes are more dominant and because light reactions of photosynthesis produce a lot of NADPH (Corpas and Barroso, 2014). Bacterial species also encode another transhydrogenase enzyme, UdhA, which is soluble and not linked to PMF. These two transhydrogenases do not share structural features or their origin and appear to have opposing metabolic roles (Heinemann *et al.*, 2016). UdhA and NNT are differentially regulated by ArcA, which activates NNT transcription and represses UdhA transcription in a redox-dependent fashion, thus reducing NADH production, while increasing NADPH production (Federowicz et al., 2014). UdhA is important for growth under conditions when too much

NADPH is produced, indicating that UdhA is typically a NADPH consumer, while NNT is a NADPH producer (Sauer *et al.*, 2004).

The exact evolutionary origin of transhydrogenase is unclear. The strong sequence similarity of the known transhydrogenases suggests a single origin. Because it is nearly absent from Archaea, it probably evolved in bacteria after the Archaea split from the bacterial lineage. It was then transferred to Eukarya via endosymbiosis of α-proteobacteria which became modern mitochondria. Evolutionary analysis suggests that transhydrogenase evolved in tandem with an ion-translocating NADH dehydrogenase (complex I), their functions probably being complementary (Studley *et al.*, 1999; Albracht, Meijer and Rydström, 2011). Only the NAD(H)-binding domain, dI, shows any sequence similarity to other proteins, specifically to bacterial alanine dehydrogenase and eukaryotic saccharopine dehydrogenase (Studley *et al.*, 1999). dIII and dII do not share significant sequence similarity to other protein families. Sequence of the dIII domain is, however, the most conserved between species followed by dI and dII respectively.

Interdomain fusion and splicing occurred several times during the evolution and some species have all the three domains encoded on a single polypeptide, while others have them distributed over two or three polypeptides or encoded in a different sequence (Figure 1.6). The number of transmembrane helices also varies between 12 and 14 in different species. High-resolution structure of transhydrogenase would provide an explanation of how these modular domains with distinct evolutionary origins work together.

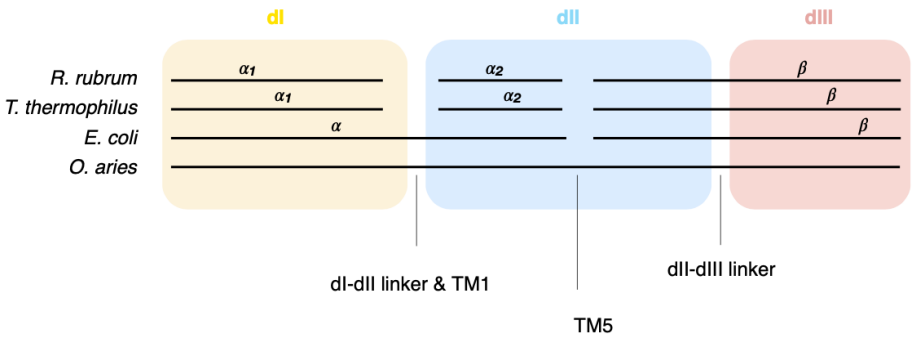


Figure 1.6. Domain split of model transhydrogenases.

NNT is split into one, two or three polypeptide chains. dII-dIII linker is conserved in all transhydrogenases, while the dI-dII linker and TM1 are limited to NNTs with - polypeptide split (e.g. E. coli) and single polypeptide NNTs (metazoans, including mammals). TM5 is only present in metazoans.
Adapted from Kampjut & Sazanov, 2019.

1.3.2. Physiological role of transhydrogenase

Transhydrogenase operates at the interface between catabolism and anabolism by interconverting NADP(H) and NAD(H) pools. Furthermore, it interacts also with the proton motive force, which makes it unique among the metabolic enzymes. Physiologically, mitochondrial transhydrogenase runs in the forward direction and provides NADPH at the expense of NADH and the pmf (Figure 1.7). NADPH is a substrate for anabolic reactions and the reactive oxygen species detoxification system. Finely tuned NADPH/$NADP^+$ ratio is also an important signifier of the oxidative status within the cell (Rydström, 2006; Jackson, 2012). NADPH regulates the redox state of the mitochondrial matrix by maintaining a highly reduced glutathione pool via NADPH-linked glutathione reductase. NADPH level is also important for hydrogen peroxide detoxification in mitochondria. Peroxidases depend on electrons from thioredoxins that are in turn reduced by electrons ultimately originating from NADPH. This system is particularly crucial for mitochondria, as they generate a lot of hydrogen peroxide as a by-product of mitochondrial metabolism in particular by complexes I and III as well as other NADH-linked dehydrogenase enzymes. Indeed, when NNT is knocked out from a superoxide dismutase-deficient mouse strain, the effect is lethal, despite neither of the individual knock-outs being lethal by themselves (Kim et al., 2010). The role of NNT in ROS detoxification is further made clear by a finding that the loss of NNT increases the expression of mitochondrial catalase, which is an alternative hydrogen peroxide detoxifier (Dogar *et al.*, 2020). Overall, respiratory state of the mitochondria and presence of other sources of NADPH in liver mitochondria determines the contribution of NNT to peroxide detoxification (Ronchi *et al.*, 2016).

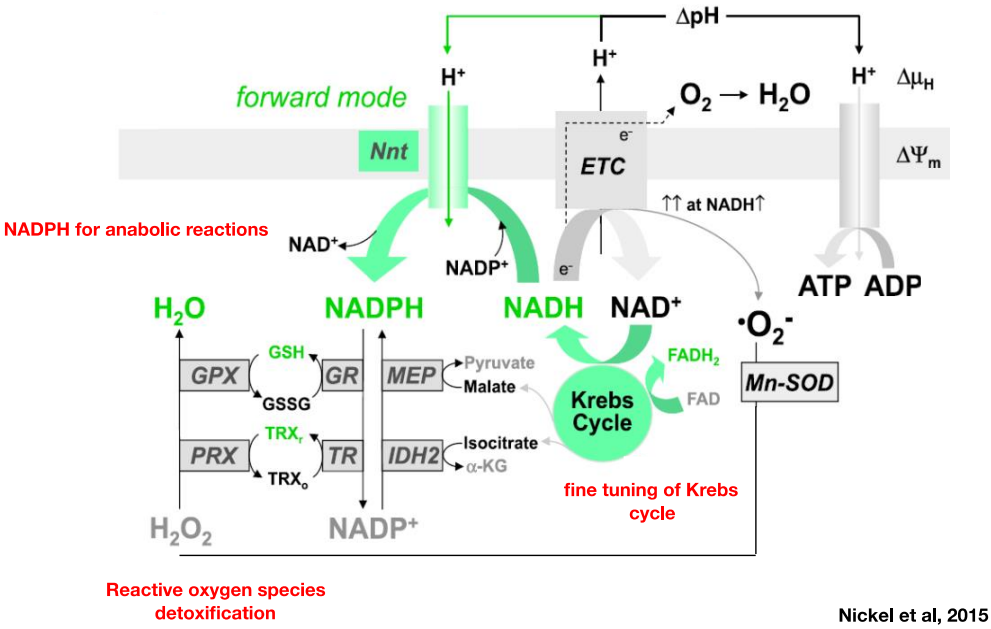


Figure 1.7. Physiological roles of transhydrogenase.

Transhydrogenase operates in the forward mode physiologically. NADPH which it produces is used for anabolic reactions directly. NADPH is also a substrate for glutathione reductase (GR), which is an essential part of the ROS detoxification system. It could also be involved in regulation of the Krebs cycle.
Adapted from Nickel et al., 2015.

Another possible role of NNT was suggested to be fine-tuning of the Krebs cycle and linking its regulation to NAD(P)H/NAD(P)$^+$ levels and pmf. If NADP-linked isocitrate dehydrogenase (IDH2) operates in the reverse direction and NAD-linked isocitrate dehydrogenase in the forward direction they establish a futile cycle of α-ketoglutarate and isocitrate. This makes the metabolic flow at this stage of the Krebs cycle much more sensitive to allosteric regulation (Sazanov and Jackson, 1994). Since transhydrogenase is pmf-linked, this would also allow coupling of Krebs cycle regulation to the pmf. At lower pmf NADPH levels would diminish and reverse flow via IDH2 would dimish as well. This would stimulate the Krebs cycle to regenerate NADH and consequently pmf. This hypothesis also elegantly explains why transhydrogenase is linked to the pmf and the seemingly superfluous presence of two mitochondrial isocitrate dehydrogenases. So far, the experimental evidence in favour of the proposal has demonstration the reverse operation of IDH2 (Rosiersl *et al.*, 1994) but not the full futile cycle yet.

In bacteria NNT is also believed to provide NADPH for anabolic reactions. This has been exploited extensively for the biotechnological purposes because overexpression of either of the bacterial transhydrogenases often significantly increases yields of valuable secondary

metabolites (Jan *et al.*, 2013; Shi *et al.*, 2013). In *E. coli* without either of the transhydrogenase enzymes, concentrations of the Krebs cycle intermediates drastically increase, citrate up to 60 fold, while glutamate levels decrease (Timischl *et al.*, 2008). This can be explained by NADH-mediated inhibition of Krebs cycle and increased glutamate catabolism to supply necessary NADPH. The only exception to NADPH-producing role of NNT are specialized anaerobic parasites, such as a cestode *H. diminuta.* They operate NNT in the reverse direction which generates proton motive force and thus supports anaerobic ATP production (Mercer-Haines and Fioravanti, 2008).

Despite the fact that NNT is a core bioenergetic enzyme, its loss is not lethal in humans or mice because mitochondria have other sources of NADPH that are able to partly compensate for NNT loss (Francisco *et al.*, 2018). These sources include the NADP-linked isocitrate dehydrogenase (IDH2), malic enzyme and glutamate dehydrogenase. NNT expression also varies widely between tissues. In mice it is the highest in heart, kidney, liver and testicles. The lowest expression (more than 10 times lower than in heart) was found in skin, skeletal muscle and brain (Arkblad *et al.*, 2002). Nevertheless, recent proteomic data suggest that differences between tissues are lower and that brain tissue in fact highly expresses NNT (Uhlén *et al.*, 2015). The previously observed differences could also reflect intra-tissue heterogeneity, different densities of mitochondria in different tissues or different rates of proteolytic turnover of NNT in different tissues.

1.3.3. Relevance of transhydrogenase in disease

Despite not being lethal, loss of transhydrogenase nevertheless results in pathological conditions in humans. In 2012, a link between NNT deficiency and a rare form familial glucocorticoid deficiency (FGD) was described (Meimaridou *et al.*, 2012; Flück, 2017). NNT deficiency accounts for ~10% of all FGD cases. Patients of FGD suffer primarily of isolated glucocorticoid deficiency often accompanied by hypoglycaemia, seizures, failure to thrive, muscle weakness and frequent infections. Disease is currently treatable well with hormone supplementation but needs to be diagnosed promptly before a fatal hypoglycaemic attack could occur. The pathophysiological cascade between transhydrogenase deficiency and glucocorticoid deficiency has been proposed to involve either diminished NADPH levels, ROS dysregulation or both. Adrenal tissue might be particularly sensitive to ROS stress which would increase in the absence of NNT (R Prasad *et al.*, 2014). The ROS hypothesis is supported by the fact that mutations in thioredoxin reductase 2 (TXNRD2), another ROS homeostasis system, also cause FGD (Rathi Prasad *et al.*, 2014). Diminished NADPH levels could on the

other hand directly affect steroid hormone synthesis as the mitochondrial cytochrome enzymes responsible for corticoid synthesis depend on NADPH. This hypothesis is partly supported by the fact that some FGD patients also exhibit gonadal symptoms and hypothyroidism (Roucher-Boulez *et al.*, 2016). It is not clear why relatively few FGD patients also have decreased mineralocorticoid levels. Overall, the phenotypic heterogeneity of FGD is large and even patients from the same families can be affected very differently, which suggests that there is a large effect of the environment and genetic background on the outcome of the disease. Phenotypes of heterozygotes are also not clear in humans nor in mice and more research also needs to be done to uncover potential non-symptomatic patients with NNT deficiency.

In mice, a widely used strain C57BL/6J (6J) has been shown to have dysfunctional transhydrogenase due to a missense mutation and several exon deletions in its NNT gene which predisposes it to metabolic abnormalities similar to the type II diabetes (Toye *et al.*, 2005). The glucose intolerance is caused by insulin response to glucose stress in 6J mice compared to 6N (sister strain with intact NNT gene) mice (Fergusson *et al.*, 2014). Absence of NNT interferes with release of insulin from pancreatic β-islet cells because lack of NNT prevents glucose from closing the K^+ATP channels and opening the Ca^{2+} channels in pancreatic beta cells (Attané *et al.*, 2016). It has been rather controversially proposed that reverse NNT operation could be responsible for the release of insulin in β-cells (Santos *et al.*, 2017). Pancreatic β-cells lacking NNT also show a profound increase in ROS (Ronchi *et al.*, 2013). In humans, there also appears to be a link between NNT deficiency and diabetes as described recently for a patient with adrenal insufficiency and diabetes as a cause of NNT loss (Scott, Van Vliet and Deladoëy, 2017). NNT expression was also found to be higher in visceral fat of obese patients compared to normal weight people (Heiker *et al.*, 2013).

Under non-physiological circumstances, i.e. heart failure, transhydrogenase can strikingly operate in the reverse direction to compensate for the decreased proton motive force (Nickel *et al.*, 2015). Oxidation of NADPH to re-establish pmf, however, comes at a cost of increased oxidative stress and ROS production (Nickel *et al.*, 2015). This is consistent with a finding that brain damage resulting from postnatal hypoxic ischemia brain injury was reduced in 6J mice compared to 6N (Wolf *et al.*, 2016). Inhibition of transhydrogenase could thus have a potentially therapeutic effect in ischemia-reperfusion injury and in situations involving increased ROS in general.

NNT could also be a target for inhibitors in several cancers. The occurrence of spontaneous as well as some types of chemically induced cancers is much lower in C57BL/6J

mice than in other mouse strains, possibly due to increased susceptibility to ROS (Albracht, Meijer and Rydström, 2011). NNT knockdowns were also found to lead to reduced tumorigenicity and proliferation of SK-Hep1 cells (Ho *et al.*, 2017). Furthermore, NNT was found to be over-expressed in several tumours: in non-small lung cancer (He, Zhang and Xia, 2020), bladder cancer (Wu *et al.*, 2019), gastric cancer (Li *et al.*, 2018) and cholangiocarcinoma (Gu *et al.*, 2019; Huang *et al.*, 2019). NNT is downregulated in gastric tumour tissue but upregulated in invasive tumours (Esfandi *et al.*, 2019). In breast cancer, NNT expression does not differ but NNT-AS1 (antisense long non-coding RNA against NNT) was found to be downregulated (Gargari *et al.*, 2019). Inhibition of transhydrogenase thus deserves more research in the coming years and new NNT inhibitors need to be developed as no specific nano-molar affinity inhibitors are currently known.

Finally, NNT inhibition could be used to treat infections. 6J mice were more resistant to pulmonary infection with *Streptococcus pneumoniae* due to the reduced macrophage inflammatory response (Ripoll *et al.*, 2012). The malaria parasite has also been recently shown to be dependent on NNT to provide NADPH to apicoplast and crystalloid organelles (Saeed *et al.*, 2020). Overall, transhydrogenase is an important emerging drug target in cancer and metabolic diseases. NNT inhibitors could be particularly useful because NNT inhibition is not lethal in humans.

1.4. Recent structural and mechanistic understanding of transhydrogenase

Kaplan and co-workers first described transhydrogenation activity in *Pseudomonas fluorescens* and in animal tissues (Colowick *et al.*, 1952; Kaplan, Colowick and Neufeld, 1952, 1953). Later transhydrogenases were described in many organisms but most biochemical and structural work has been done on the four model transhydrogenases: from *Escherischia coli*, *Rhodospirillum rubrum, Thermus thermophilus* and *Bos taurus.*

Since 1999 structures of isolated domains and dI_2dIII complexes from several of these species were solved by x-ray crystallography in native, reduced or oxidised states (Prasad *et al.*, 1999; Cotton *et al.*, 2001; Singh *et al.*, 2003; Sundaresan *et al.*, 2003; Mather *et al.*, 2004; Johansson *et al.*, 2005). While these structures showed the atomic details of protein folds and interaction with substrates, they failed to explain the unique mechanism of NNT which depends on the arrangement and coordination of all the domains in the dimer. Below I briefly describe the structures of isolated domains of NNT and then discuss the architecture of intact transhydrogenase and its relevance for the mechanism.

dI always purifies and crystallises as a dimer. Dimer of two ~40 kDa monomers is stabilised by a large buried surface extended by an elongated hairpin (ovine residues 172-184) reaching out from each monomer into the neighbour. NAD(H) occupies a classical nucleotide binding site in one of the two Rossmann folds that comprise the dI monomer (Prasad *et al.*, 2002; Johansson *et al.*, 2005; Sundaresan *et al.*, 2005).

dIII domain on the other hand crystallises as a monomer. The ~20 kDa domain comprises a single classical Rossmann fold with six-stranded β-sheet but NADP(H) binds in a non-classical manner with a reverse nucleotide orientation which is important for bringing the NADP nucleotide into close contact with NAD nucleotide on the dI domain for direct hydride transfer (Prasad *et al.*, 1999; Jeeves *et al.*, 2000; White *et al.*, 2000). All the solved structures of dIII domain to date have essentially the same conformation. Loops D and E are transhydrogenase-specific and highly conserved structural elements that have been proposed to be involved in the mechanism. Loop E closes over the nucleotide binding pocket, whereas loop D could provide contact surface for interactions with other domains. Loop D has been observed in two different conformations in *R. rubrum* dIII in the two molecules in one asymmetric unit and since loop D is also not close to the nucleotide binding pocket this transition probably is not catalytically important (Sundaresan *et al.*, 2003). The alternative, non-dI binding conformation of dIII is therefore still not resolved and only a structure of an intact enzyme can reveal different redox-dependent conformations of dIII.

Structures of dI_2dIII heterotrimers from *T. thermophilus* and *R. rubrum* showed the nicotinamide rings closely poised for hydride transfer between NAD(H) and NADP(H) across the dI/dIII interface and established that hydride transfer between the two nucleotide must be direct as they approach each other to within 3-4 Å. Surprisingly, they also demonstrated that only one copy of dIII can bind to dI_2 at a time (Cotton *et al.*, 2001; Singh *et al.*, 2003; Mather *et al.*, 2004; Sundaresan *et al.*, 2005). Binding of two molecules of dIII to the dI dimer in the same orientation is sterically prohibited and binding of the second dIII monomer would either require reorganization of its structure or the first dIII to detach.

Structure of the transmembrane dII was elucidated only recently. Isolated *T. thermophilus* dII was solved *in meso* and revealed a unique fold and a putative proton translocation pathway (Leung *et al.*, 2015). An improved 2.2 Å resolution structure of *T. thermophilus* dII structure delineated the proton transfer pathway in more detail and observed the presence of relatively large solvent-accessible cavities on either side of the membrane (Padayatti *et al.*, 2017). These cavities are separated by a couple of highly conserved residues,

α_2H42 and βT214, which are involved in proton transport. *E. coli* transhydrogenase which is the most biochemically characterized bacterial enzyme also contains a histidine residue near the centre of the membrane but it appears to be βH91 based on mutagenesis studies (Hu *et al.*, 1999).

To date, the only intact architecture of NNT was determined at 6.9 Å resolution in the presence of $NADP^+$ from *T. thermophilus*$^+$ (Leung *et al.*, 2015). It showed that NNT is an asymmetric dimer, with dIII adopting a face-up position (suitable for hydride transfer) in one monomer and a face-down conformation in the other monomer, compatible with the finding that only one dIII can bind to dI_2 at a time (Leung *et al.*, 2015).

The existence of the asymmetric conformation of NNT led to a proposal of a "division of labour" mechanism (Figure 1.8) in which dIII rotates 180º and alternatively functions in proton translocation (face-down) and hydride transfer (face-up) (Jackson *et al.*, 2015). The key feature of the mechanism which prevents slipping (i.e. decoupling proton transfer from hydride transfer) is the occlusion (kinetic restriction) of the nucleotide bound to dIII. Kinetic studies of the cyclic NAD(H) reaction (where NADP(H) remains bound to dIII) revealed that the occluded NADP(H) state is elevated at a low pH, when the key protonatable residue in dII is protonated (Hutton *et al.*, 1994; Sazanov and Jackson, 1995). It was also shown that the binding affinities of NAD(H) to dI are not altered during the catalytic cycle; this domain usually remains open (Hutton *et al.*, 1994; Hu *et al.*, 1995; Venning *et al.*, 1998, 2000). The proposed mechanism involved six tightly regulated consecutive steps which required strict out-of-phase functioning of the two monomers and a long-lived face-up intermediate.

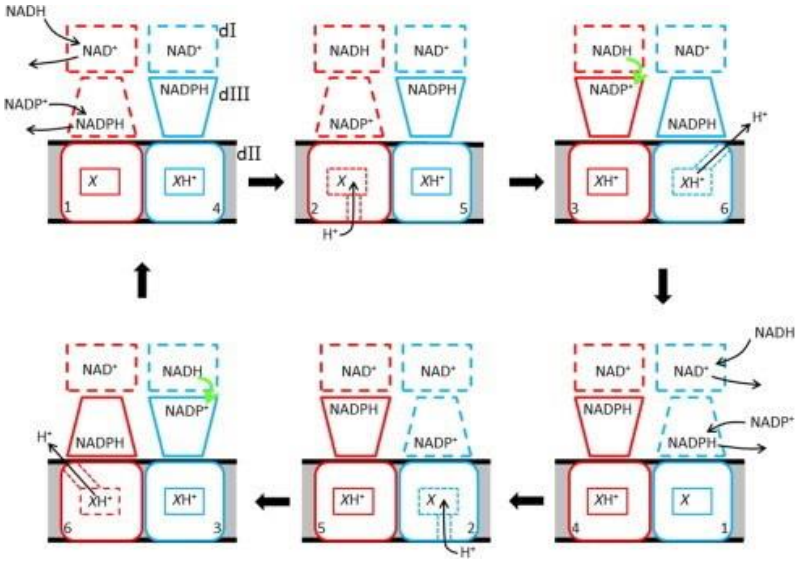


Figure 1.8. Division of labour mechanism of NNT.

Reactions on the two monomers (blue and red) proceed in an anti-phase fashion. dIII can exist in occluded (full outline) and open (dashed outline) conformation, while dI is always open. dIII cycles between face-up and face-down conformations. dII can exist in occluded, open-to-N-side or open-to-P-side conformations. Coupling is achieved by interaction between dII and dIII.
Reproduced from Jackson et al., 2015

However, the details of this mechanism were largely speculative in the absence of high-resolution structure of NNT in several of the proposed intermediate steps. While being consistent with the *T. thermophilus* architecture and most of the previous biochemical work they also failed to account for the key finding that $NADP^+$/NADPH have different effects on the conformation of NNT (Yamaguchi and Hatefi, 1989; Yamaguchi, Wakabayashi and Hatefi, 1990). NADPH has been shown to increase the susceptibility of several bonds outside of the dIII domain (in particular within the dI-dII linker) to tryptic hydrolysis which suggests that NADPH can bring about long-range conformational changes (Yamaguchi, Wakabayashi and Hatefi, 1990). Similarly, NADPH and $NADP^+$ lead to different susceptibility of Cys893 to chemical modification in mitochondrial and bacterial enzyme (Yamaguchi and Hatefi, 1989; Bragg and Hou, 1999). Clearly, NADPH and NADP+ promote radically different conformations of transhydrogenase, whereas according to the proposed mechanism, their effects on the NNT should not differ. This indicates a clear need to understand the mechanism of NNT better and provide a structural explanation for these effects.

1.5. Complex I

Complex I is the largest of the mitochondrial respiratory complexes with a total mass of ~ 1 MDa. It can be divided into the membrane domain (MD) and the peripheral arm (PA), in total composed of 45 polypeptide subunits. Fourteen of them are the "minimal catalytically active" core subunits conserved throughout bacterial and eukaryotic complex I (Figure 1.9) (Baradaran *et al.*, 2013). The rest of the subunits are species-specific and are believed to be important for the regulation and stability of the complex (Stroud *et al.*, 2016). Mammalian complex I is not catalytically active until fully assembled with all of its supernumerary subunits. Supernumerary subunits are mostly small (10-20 kDa with the exception of the 39 kDa and 42 kDa subunits) and form a "cage" around the evolutionary conserved core subunits (Figure 1.10) (Fiedorczuk *et al.*, 2016). It has been described that supernumerary subunits are necessary for the stability and activity of complex I but the exact function of many of the supernumerary subunits have not yet been established. Particularly interesting are the supernumerary subunits with bound cofactors. 13 kDa binds a Zn^{2+} ion, which could mean that it has a redox signaling role. Because of the vicinity of the Zn^{2+} ion to the FeS chain (13 Å away from N6a and 15 Å away from N5 clusters), it could directly influence the electron transfer reactions and electric potentials of FeS clusters. Similarly, 39 kDa has a NADPH molecule bound and could have a role in sensing the $NADP^+$/NADPH ratio. It could also have a role in anchoring the PA to the membrane via the amphipathic helix. Complex I also contains two SDAP subunits which bind an acyl-phosphopantetheine molecule. This could link complex I assembly to the fatty acid metabolism. Finally, subunit B14.7 could be important for supercomplex I-III formation and B12 and AGGG subunits could be important for complex I-IV contacts in the respirasome formation (Letts, Fiedorczuk and Sazanov, 2016).

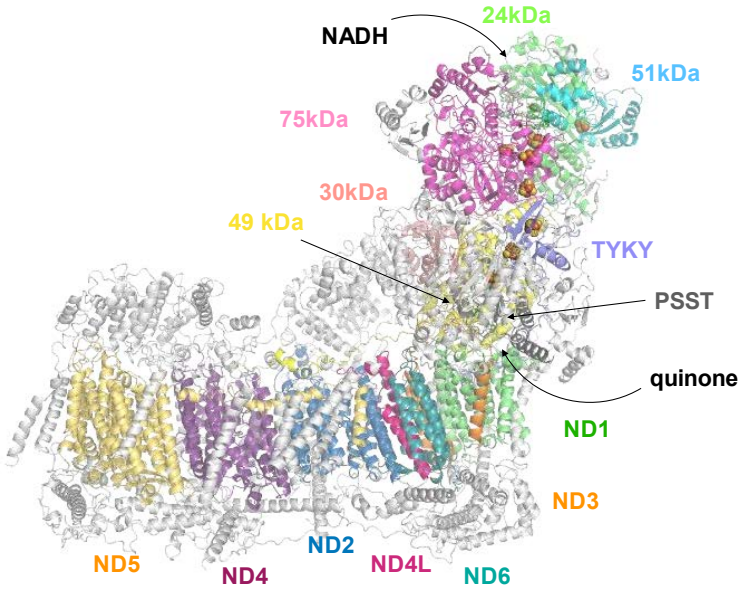


Figure 1.9. Core subunits of complex I.

Subunits conserved in all organisms and necessary for the reaction of complex I are labelled with corresponding colours and mammalian supernumerary subunits are shown in grey.
Adapted from Kampjut & Sazanov, 2020.

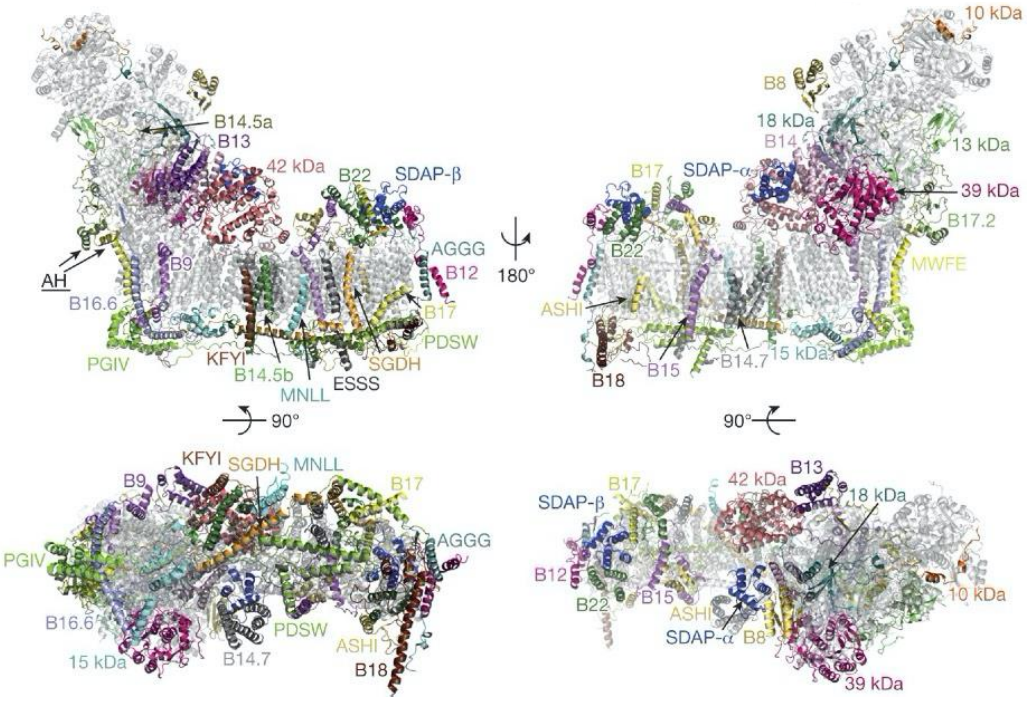


Figure 1.10. Supernumerary subunits of complex I.

Supernumerary subunits in ovine complex I are coloured and labelled and core subunits are shown in grey.
Reproduced from Fiedorczuk et al., 2016.

1.5.1. Evolution of complex I

Complex I has a modular structure and each of the several modules has a unique function and evolutionary history (Figure 1.11). At some point during the evolution of complex I, several pre-existing modules combined together to form a modern complex I. This event probably preceded the evolution of mitochondria because complex I is present in modern bacteria and archaea (Gabaldón, Rainey and Huynen, 2005).

N-module of complex I is the dehydrogenase module consisting of 75 kDa, 51 kDa and 24 kDa subunits, containing the FMN cofactor and several FeS clusters. It catalyses NADH dehydrogenation. It is the latest evolutionary addition to complex I and it evolved as an electron-receiver from the Krebs cycle in aerobic respiratory chains. Some organisms lack 51 kDa and 24 kDa subunits and probably get electrons from flavodoxins (Finel, 1998).

Q-module is the connecting module between the peripheral and membrane arms and is homologous to soluble [NiFe] hydrogenases. It consists of the 49 kDa, PSST, TYKY and 30 kDa subunits. 49 kDa and PSST are ancient electron transferring enzymes and the quinone binding site in complex I directly corresponds to the NiFe active site in hydrogenase (Brandt, 2006). TYKY is a ferredoxin-like protein that was added later to the assembly and the 30 kDa was added last and has no catalytic role but provides stability to the assembly.

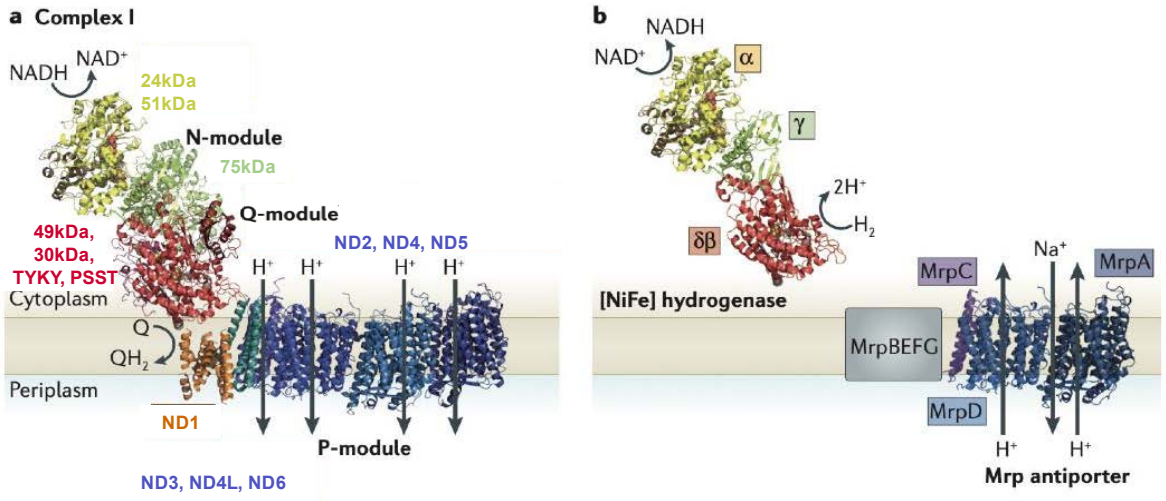


Figure 1.11. Modular architecture of complex I.

a. *Complex I is composed of N, Q and P modules. Nqo8 (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 N/Q-module and P-modules respectively.*
Adapted from Sazanov, 2015.

P-module is the proton pumping module which can be further divided into the distal P_d and proximal P_p modules. It evolved from the H^+/Na^+ antiporter (Mrp). Subunits ND5 (MrpA), ND4 and ND2 (MrpD) and ND4L (MrpC) have direct homologues with the Mrp subunits (Mathiesen and Hägerhäll, 2002, 2003). Subunit ND6 is also homologous to the C-terminal part of MrpA (Dzioba-Winogrodzki *et al.*, 2009).

Acquisition of the rest of membrane arm, which is critical for the connection between the redox and proton pumping arms of complex is not clear evolutionary as subunits ND3 and ND1 do not have any known homologues. ND1 is the most conserved of the membrane subunits and is crucial for the coupling mechanism in complex I, hence it is probably that it emerged only once during the evolution (Efremov and Sazanov, 2012). Unexpectedly, the core fold of ND1 is similar to a half-channel of the antiporter-like subunit, the similarity that only became apparent once the structure of complete complex I was solved as the sequence identity is below 20% (Baradaran *et al.*, 2013). Nevertheless, helices in ND1 are much more tilted relative to the membrane than in antiporters and some of them are very short (e.g. TMH5) and others are very long (e.g. TMH1).

Supernumerary subunits were incrementally acquired during the evolution (Elurbe and Huynen, 2016). It is likely that the complex I in the α-proteobacterial endosymbiont which gave rise to the modern mitochondria already contained some supernumerary subunits. *Paracoccus denitrificans* for example contains three supernumerary subunits (B17.2, 18kDa, 13kDa homologues) which were likely present already in the mitochondrial ancestor (Yip *et al.*, 2011). In total six supernumerary subunits were acquired from the proto-mitochondrial ancestor's genome: SDAP, 39kDa (homologous to short chain dehydrogenase), 13kDa, B17.2, 18kDa, 42kDa. The rest of the supernumerary subunits originated in the host genome. Fourteen of them are shared between all eukaryotes: B14.7, ESSS, B8, PDSW, B13, PGIV, 15kDa, B22, B16.6, B18, B14, B12, MWFE, NUXM. After the plants split from the fungal and animal lineage, three additional subunits were acquired which are not present in plants: B15, ASHI, B9. Finally, after the animal lineage has diverged, eight more subunits were added (AGGG, SGDH, B14.5a, B14.5b, B17, 10kDa, MNLL, KFYI) and NUXM subunit was lost. This leads to the modern 45-subunit mammalian complex I.

Apart from these different versions of complex I in different organisms, there exists a larger superfamily of complex I-like enzymes. The closest to the canonical complex I is its chloroplast analogue, NDH. It is also found in cyanobacteria and some archaea and bacteria. It

contains only 11 subunits and lacks the entire N-module, hence the electron donor to NDH is likely ferredoxin. Several structures of NDH recently became available which show a high degree of conservation with complex I (Laughlin *et al.*, 2019; Schuller *et al.*, 2019; Zhang *et al.*, 2020). The folds of the conserved subunits are the same, and the quinone binding site is largely superimposable, although it is disordered to varying degrees in all of the NDH structures. Because sequences of NDH have been shown to display variability greater than the sequences of the full 14-subunit complex I, it was concluded that the 11-subunit NDH-like complex is likely an ancestor to all present-day complex I enzymes (Moparthi and Hägerhäll, 2011).

Structure of the hydrogen gas evolving membrane bound hydrogenase (Mbh) also recently became available (Yu *et al.*, 2018). This protein contains 14 subunits, 10 of which are homologous to the minimal complex I without the N-module and one of the antiporter-like subunits and 4 of which are homologous to the sodium-translocating module of Mrp. Strikingly, the P-module-like part of the Mbh is rotated 180º around the membrane plane normal. How exactly such an arrangement still results in a coupled reaction that drives proton and sodium pumping is not yet understood but will be crucial for any generalization of complex I mechanism to other similar redox-coupled pumps.

Other complex I-related proteins which have not been structurally characterised yet include the 11-subunit Fpo complex from *M. mazei* (also without the N-module), formate hydrogenlyases FHL-2 (10-subunit, lacking the flavoprotein part of the N-module and ND3 and ND6) and FHL-1 (7-subunit, lacking additional two antiporter-like subunits ND4 and ND2 compared to FHL-2), as well as the energy-converting oxidoreductase (Ech), the simplest of the complex I homologues (6-subunit, containing only the Q-module and ND1 and ND5 homologues).

The structures of these proteins will be crucial for our understanding of the adaptations of complex I-like proteins for different substrates and could provide important insights into the shared elements of the mechanism. The high modularity of the complex I not only in with respect to the electron input module but also in terms of the proton-pumping modules is striking and will need to be understood in terms of the coupling mechanism. The structure of the complex I ancestor Mrp will also be crucial for the understanding of the proton pumping mechanism.

1.5.2. Assembly of complex I

Assembly of complex I is also modular as it involves formation of independent assemblies which are later combined together (Vartak *et al.*, 2015; Stroud *et al.*, 2016; Guerrero-Castillo *et al.*, 2017). Interestingly, many supernumerary subunits assemble with the separate modules before the assembly of the full core of complex I. Assembly also involves several assembly factors which later dissociate.

The main five modules of complex I which are assembled first include:

- N module consisting of 51kDa,24kDa,75kDa, B8 subunits,
- Q/Pp-a module consisting of Q module (49 kDa, PSST, TYKY and 30 kDa subunits) with added ND1, PGIV, B9, B16.6 subunits and TIMMDC1 assembly factor,
- Pp-b module consisting of ND2, ND3, ND4L, ND6, KFYI and B14.5b subunits and assembly factors ACAD9, ECSIT, NDUFAF1, TMEM126B and putatively also COA1 TMEM186
- Pd-a module consisting of the Pd' module (SGDH, PDSW, ESSS, B17 subunits) with added MNLL and ND4 subunits and assembly factors TMEM70 and FOXRED1 and putatively also ATP5SL assembly factor
- Pd-b module consisting of the Pd' module with added ND5, B18, B12, ASHI, AGGG, SDAP and B22 subunits.

During the maturation process, Pd-a and Pd-b subunits join together with the 42 kDa and B15 subunits to form a Pd-680kDa subcomplex. Similarly, Q/Pp-a and Q/Pp-b modules combine together with the B9 and MWFE subunits as well as the NDUFAF2 assembly factor to form the Q/Pp-736kDa subcomplex. In the final maturation step, the main assemblies are combined into the Q/P subassembly, and in the final step, N-module as well as B14.7 and SDAP subunits are added and all the assembly factors dissociate.

1.5.3. Diseases of complex I

Complex I deficiency is also implicated in a number of inborn human diseases and is the most frequently diagnosed single enzyme deficiency of a mitochondrial enzyme. The two conditions with which complex I deficiency is most often correlated are LHON (Leber's Hereditary Optic Neuropathy) and MELAS (Mitochondrial Encephalomyopathy, Lactic Acidosis, and Stroke-like episodes) (Rodenburg, 2016). However, clinical manifestations of complex I deficiency are varied and can sometimes be hard to diagnose, firstly because of the

heteroplasmy effects and secondly because of secondary effects of complex I mutations which can include increased ROS production, altered mitochondrial morphology and membrane potential. At the moment, the only available treatment for complex I deficiency is symptomatic. Detailed review of human patient mutations in relation to complex I structure and mechanism has identified several mechanistically important hot-spots of mutations that will be discussed in Chapter 4 (Fiedorczuk and Sazanov, 2018). Furthermore, several known drugs, such as metformin, bind to it and inhibit it but the mechanisms of inhibition and physiological consequences are not well understood (Fontaine, 2018).

1.6. Recent structural and mechanistic understanding of complex I

1.6.1. History of structure determination

Structural understanding of complex I begun with X-ray crystal structures of bacterial homologues. In 2006, the structure of the peripheral arm of *T. thermophilus* was solved, which showed the organization of the redox cofactors (Sazanov and Hinchliffe, 2006). In 2011 this was followed by *E. coli* complex I membrane domain, which revealed a unique antiporter fold of antiporter-like subunits and their connections (Efremov and Sazanov, 2011). The crystal did not contain NuoH subunit (ND1 homologue), which is necessary for connecting the PA and MD. This was finally resolved in 2013 in a crystal structure of the entire *T. thermophilus* complex I (Baradaran *et al.*, 2013). This structure revealed the connection between the PA and MD and the quinone binding cavity, which is a narrow and long tunnel extending up to ~20 Å above the lipid membrane surface.

The crystal structure of *Yarrowia lipolytica* complex I was the first crystal structure (at 3.6-3.9 Å) of eukaryotic complex I and revealed the architecture of several eukaryotic supernumerary subunits (Zickermann *et al.*, 2015). The first complete structure of mammalian (ovine) complex I was solved in 2016 by the Sazanov group using cryo-EM (Fiedorczuk *et al.*, 2016). Since then cryo-EM structures of several mammalian complexes at resolutions up to 3.3 Å, including bovine, mouse, human and porcine complex I have allowed ever more complete and accurate models to be built (Wu *et al.*, 2016; Zhu, Jiapeng, Vinothkumar and Hirst, 2016; Guo *et al.*, 2017; Agip *et al.*, 2018). Structures of human, porcine and ovine complex I were solved also as parts of various supercomplexes, placing the complex I within the physiological context (Letts, Fiedorczuk and Sazanov, 2016; Letts *et al.*, 2019).

1.6.2. Catalytic reaction of complex I

Complex I catalyzes NADH oxidation and ubiquinone reduction, the first reaction step in the electron transfer chain. In the peripheral arm (PA) of the enzyme, electrons from NADH are initially accepted by FMN and then passed along a chain of iron-sulfur clusters to the final acceptor ubiquinone (Figure 1.12). There are eight or nine FeS clusters, depending on the species. Two clusters are binuclear (N1a and N1b) while the rest are tetranuclear. Electron transfer between the iron-sulfur clusters occurs by electron tunneling, a process during which electron occupies either the donor or the acceptor but avoids the protein bridge between them. This ensures transfer to a specific location along a specific direction. The rate of tunneling falls sharply with the increasing distance between the donor and the acceptor and for electron transfer to occur at the enzymatically-useful timescales of milliseconds they need to be < 14 Å apart. Seven clusters lie directly on the electron path the connects the NADH and quinone binding sites, while N1a lies off-path "upstream" of FMN. Because it is close enough to FMN to be able to accept electrons its role could be to temporarily store an electron to prevent flavosemiquinone generation and reduce ROS generation. FeS cluster N7 occurs only in bacteria and is located too far off the main path (~20 Å) to be reducible and is probably an evolutionary vestige. In steady state with an excess of NADH, not all of the FeS clusters are reduced and one electron resides approximately on every other FeS cluster (Bridges, Bill and Hirst, 2012).

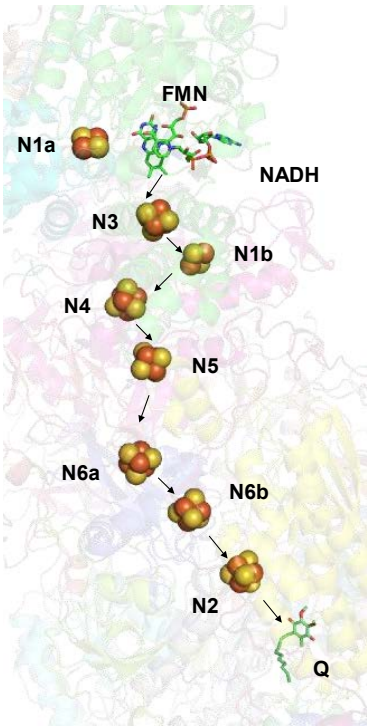


Figure 1.12. Electron transport chain of complex I.

Transport of electrons from NADH to quinone proceeds via the FMN and eight iron-sulphur clusters.

FMN, NADH and quinone are two-electron carriers (NADH donates 2 e^- and 1 H^+, while FMN quinone donates 2 e^- and 2 H^+), while FeS clusters are one electron carriers. Partly reduced FMN or quinone are not stable and can react with oxygen to produce superoxide radicals. Because the quinone binding site is deep within the protein environment, while FMN is on the surface complex I has been suggested to mainly generate ROS at the FMN site, thus being a major contributor to ROS generation in the mitochondrial matrix. The quinone binding site is a narrow, ~25 Å long tunnel found at the interface between the peripheral and membrane domains that connects the lipid bilayer to the last FeS cluster N2 at the end of the redox chain.

NADH-quinone oxidoreduction is coupled over a distance of more than 200 Å to pumping of four protons by four separate proton pumps from the matrix to the intermembrane space side. The closest to the quinone binding cavity is the E channel composed of ND1, ND6 and ND4L subunits, followed by three antiporter-like subunits ND2, ND4 and ND5 (Figure 1.13). These pumps are connected by a string of conserved charged residues along the center of the entire length of the membrane domain, forming the central hydrophilic axis, a distinct feature of complex I. The antiporter-like subunits contain an N-terminal proton-entry half-channel and a C-terminal proton-exit half-channel that are related to each other by pseudosymmetry. Both entry and exit paths contain a conserved lysine residue (or glutamate in case of ND4 C-terminal half-channel), that is likely the key to proton pumping. In each antiporter, the N-terminal lysine forms a pair with conserved TMH5 glutamate, which is thought to modulate pK_a of lysine (Efremov and Sazanov, 2011). The C- and N-terminal lysines are connected by a central lysine (or histidine in the case of ND5), which likely acts to transfer protons between the N- and C-terminal half-channels. All of these key residues sit on breaks in TM helices (N-terminal TMH7, connecting/central TMH8 and C-terminal TMH12), which likely renders the central hydrophilic axis flexible. This flexibility may have a role in the conformational coupling mechanism and/or could help with pK_a modulations of key residues.

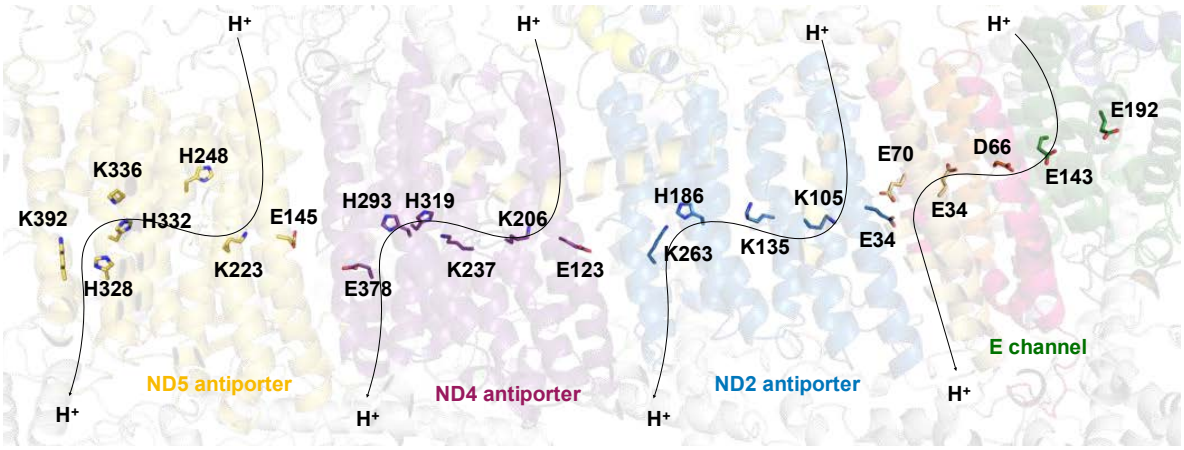


Figure 1.13. Proton pumps and the central hydrophilic axis of complex I.

Membrane arm contains four separate proton pumping channels, three in the antiporter-like subunits and one in the E-channel. They all contain N-terminal and C-terminal entry points with a conserved lysine and one or more connecting lysines and histidines. E-channel consists of several conserved glutamate and aspartate residues.

1.6.3. Open and closed conformations of complex I

Both isolated complex I and complex I within supercomplexes have been observed in two conformations, the closed and the open conformation which differ in the angle between their peripheral and membrane arms (Vinothkumar, Zhu and Hirst, 2014; Fiedorczuk *et al.*, 2016). The opening coincides with unfolding of several loops that form the interface between the peripheral and membrane arms as is described in detail in Chapter 4.

It has been argued that the open conformation of complex I corresponds to the deactive state in mouse and bovine complex I (Agip *et al.*, 2018; Blaza, Vinothkumar and Hirst, 2018). The basis for this was that proportion of open complex I in the sample increases upon incubation at 37ºC without substrates. The authors argued that disordering of the quinone binding site loops (ND1, ND3, 49 kDa), TMH4 of ND6 and the amphipathic loops of 39 kDa in the open complex I is too extensive to be part of the catalytic cycle, hence the open structures must correspond to the deactive state.

Deactive complex I is a catalytically inactive state present in many eukaryotic species that is probably important *in vivo* for the prevention of excessive ROS production and reverse electron transfer by complex I under ischemic conditions (Kotlyar and Vinogradov, 1990; Dröse, Stepanova and Galkin, 2016). Deactivation of mammalian complex I *in vitro* proceeds slowly at elevated temperatures (30-37ºC) in the absence of substrates and can be stimulated by increased pH (9.4), biguanides or divalent metal ions (Matsuzaki and Humphries, 2015). Deactive state can slowly (1-4 min^{-1}) revert back to the active state in the presence of substrates.

In *Y. lipolytica* both the deactivation and reactivation proceed with much faster kinetics and do not require elevated temperatures (Maklashina, Kotlyar and Cecchini, 2003).

A number of recent experimental data question the notion that active complex I must necessarily exist in the closed state and whether opening only occurs during deactivation and not also during the catalytic cycle. Firstly, open and closed conformations complex I were observed within active preparations of ovine respiratory supercomplexes and indeed mixtures of open and closed complex were also observed in the mouse and bovine preparations that led to the notion of open complex I being deactive (Letts *et al.*, 2016, 2019). Secondly, MD simulations predicted that unfolding and refolding of mobile loops can happen on a millisecond timescale and should not limit the reaction (Haapanen, Djurabekova and Sharma, 2019). Thirdly, a sample of *Y. lipolytica* complex I was solved in the deactive state and during turnover and both of these structures exhibited a disordered ND3 loop, indicating that it is in the open state (Parey *et al.*, 2018). Trapping of the ND3 loop by reversible disulfide bridge to the nearby PSST subunit decoupled the complex I reaction (Cabrera-Orefice *et al.*, 2018). This shows that subtle movements of the ND3 loop and potentially other structural elements in the vicinity are crucial for the coupling and activity of complex I and that closed and static complex I cannot catalyze the full reaction. Unlike the alkylation of ND3 loop, locking it in the closed state did not interfere with the A/D transition. Finally, all the observed bacterial complex I structures contain a π-bulge in the ND6 TMH3, which is in mammalian enzyme only observed in the open conformation (Efremov and Sazanov, 2011; Baradaran *et al.*, 2013).

1.6.4. Coupling mechanism in complex I

Despite the accelerated advance in structural understanding of complex I from several different model species, the catalytic mechanism of complex I remains elusive. In particular the coupling between quinone reduction and proton pumping, which happen more than 200 Å apart from each other has so far not been satisfactorily explained. The precise mechanisms of quinone reduction, its binding and movement in the cavity also remain to be resolved. On the basis of the currently available structures of complex I, both conformational coupling mechanisms and “electrostatic wave” mechanisms have been proposed, as described below. However, the current experimental data is insufficient to discriminate between these proposals. None of the structures solved so far display different conformations that could be reliably mapped to different catalytic intermediates. Furthermore, most of the structures have been solved in the absence of substrates, with the exception of *Y. lipolytica* during turnover but only at modest resolution 4.5 Å (Parey *et al.*, 2018).

Nevertheless, several principles of the coupling mechanism can be worked out based on the midpoint potential measurements of complex I and the thermodynamics governing the redox reactions. In the absence of experimental data, our understanding of the coupling mechanism also relies heavily on the molecular dynamics (MD) simulations. There is no drop in mid-point electric potential until the electron reaches the final cluster N2 (Euro *et al.*, 2008; Wikström, Sharma, Ville R I Kaila, *et al.*, 2015). The mid-point electric potential of the mid-chain Fe-S clusters was estimated to around -250 mV, of the terminal FeS cluster N2 -150 mV and of the Q/QH_2 pair ~ +100 mV (for ubiquinone freely dissolved in lipid) (Euro *et al.*, 2008). The crucial energy-releasing step in the reaction is therefore quinone reduction, quinone protonation or even its release out of the cavity as the mid-point electric potential of Q/QH_2 pair bound near N2 is likely close to the N2 potential (Wikström, Sharma, Ville R.I. Kaila, *et al.*, 2015). The main challenge in understanding coupling is therefore to work out how quinone reduction happens and how it drives the proton pumps.

Transmission of energy in the quinone binding cavity has been proposed to involve the key quinone-coordinating residues, H59/D160/Y108 of the 49 kDa subunit. They could act as proton donors for the quinone and their re-protonation would elicit conformational changes in the quinone binding cavity and further in the proton pumps (Wikström, Sharma, Ville R I Kaila, *et al.*, 2015).

The Sazanov group proposed that the negative charge within the quinone cavity, probably held on the 49 kDa subunit residues H59/D160/Y108, which coordinate the quinone headgroup, initiates conformational changes in the E-channel, propagating further into the antiporters via the central hydrophilic axis (Figure 1.14) (Baradaran *et al.*, 2013; Sazanov, 2015). These movements would influence interactions between the key lysine/glutamate residues within and between antiporters, leading to pK_a changes and proton translocation (Efremov and Sazanov, 2011). This has since been observed also in MD simulations, which suggest that negative charge within the quinone cavity somehow couples quinone reduction to proton pumping, either via deprotonated 49kDa/Y108, negatively charged semiquinone or quinol radicals directly or by rearrangements of the ND1 loop, which contains many negatively charged residues (Sharma *et al.*, 2015).

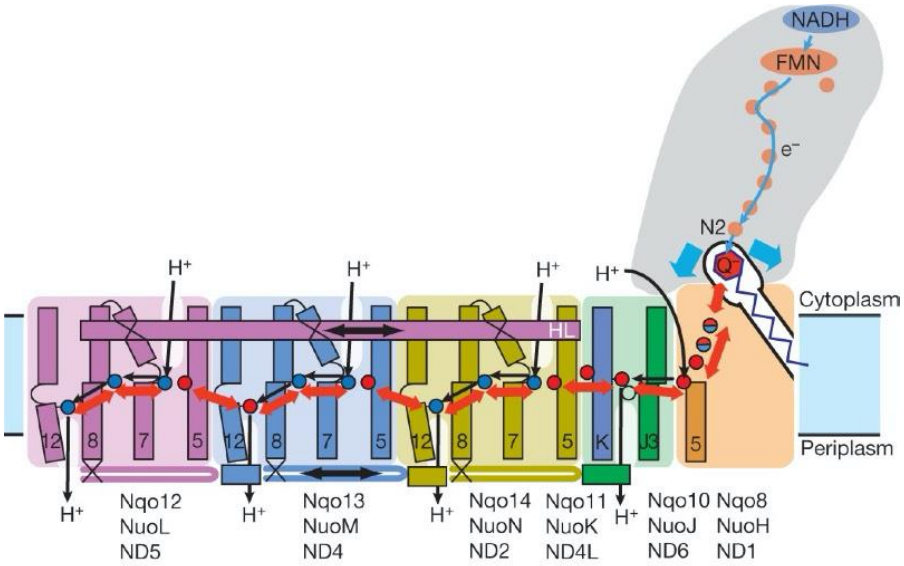


Figure 1.14. "Single stroke" coupling mechanism of complex I based on T. thermophilus structure.

Proton pumping is initiated by the negative charge accumulated in the quinone binding cavity during quinone reduction. Conformational changes propagate via the E channel into the rest of the membrane domain of complex I. Conformational coupling was proposed to be enhanced by concerted movements of the structural elements connecting the antiporters: long horizontal helix HL of ND5 and β-hairpin loops between the antiporters.
Adapted from Baradaran et al., 2013.

Molecular dynamics (MD) simulations instead suggested that rapid diffusion of quinol within the cavity is accompanied by the rearrangement of the quinone-binding 49 kDa, ND1 and PSST loops and that negative charge on the semiquinone/quinol radicals, deprotonated 49kDa/Y108 or negatively charged ND1 loop initiates these changes (Sharma *et al.*, 2015).

Long time scale molecular dynamics (MD) simulations suggested that there are broadly two stable quinone binding sites, one in the deep part of the cavity next to the N2 (Q_d, or MD sites #1-2) and one in the shallow part (Q_s, or MD sites #4-5) next to the ND1 subunit close to the exit from the tunnel (Warnau *et al.*, 2018; Haapanen, Djurabekova and Sharma, 2019). Based on the fact that many complex I preparations have been shown to co-purify with native quinone and the fact that the quinone cavity is so long and narrow, there has been a proposal that there is a stably bound quinone within the cavity that shuttles between the distal and proximal binding sites, where it reduces a second quinone molecule that eventually diffuses into the membrane (Wikström, Sharma, Ville R I Kaila, *et al.*, 2015; Haapanen and Sharma, 2018). Two full-length ubiquinone molecules could simply not fit into the cavity however, and there is at the moment no other direct experimental evidence for this proposal. MD simulations furthermore showed that quinones with long isoprenoid chains actually have fast

binding/unbinding kinetics and that the diffusion of quinone within the cavity is not rate-limiting the overall complex I reaction, casting doubt on this hypothesis (Warnau *et al.*, 2018).

Finally, a purely electrostatic mechanism has been proposed, whereby a wave of electrostatic signals sequentially breaks lysine-glutamate ion pairs along the membrane arm, leading to pK_a changes of the lysines and pumping of protons, until the cascade reverses the direction to re-protonate the lysines (Figure 1.15) (Verkhovskaya and Bloch, 2013; Kaila, 2018). This coupling mechanism could in principle co-exist with many of the postulates of the previous suggestions, especially with regards to the quinone reduction events. In the electrostatic spring model, however, the electrostatic interactions, and not conformational changes, are hypothesized to dominate the mechanism from quinone reduction onwards. Ultimately, it is likely that conformational and electrostatic components both contribute to coupling in complex I.

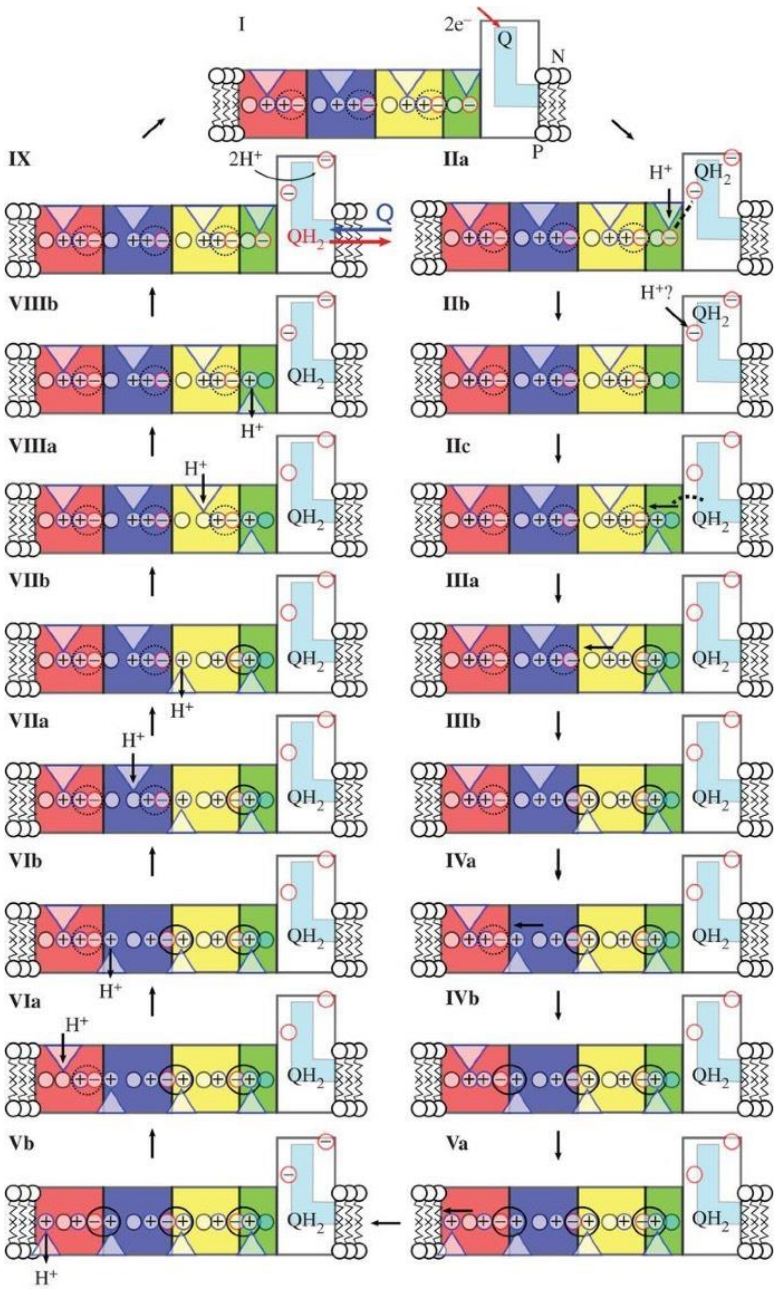


Figure 1.15. Electrostatic spring mechanism of complex I.

Quinone reduction in step I results in an electrostatic wave in steps IIa-Va, whereby proton moves from the connecting to the C-terminal lysine in each antiporter. In step Vb proton gets released from ND5 into the IMS, which enables the connecting lysine to become protonated from the matrix side (Via), starting the electrostatic cascade in the other direction (steps Via-IX), whereby protons get sequentially released from C-terminal lysines into the IMS and the connecting lysines get re-protonated from the matrix side.
Adapted from Kaila, 2018.

1.7. Cryo-EM

Structural biology investigates the structures of biological molecules and until recently, the most common method for obtaining atomic structures of proteins was X-ray crystallography. The technique was developed over a hundred years ago and involves obtaining diffraction images of ordered crystals. Since the structure of myoglobin was solved as the structure of the first protein in 1958, the Protein Data Base (PDB) has expanded to include 145029 crystal structures at the time of writing this thesis (April 2020). This demonstrates the wide success and applicability of the method. However, X-ray crystallography is a laborious method because obtaining regular well-diffracting crystals requires extensive screening of crystallisation conditions. Many proteins do not form regular crystals whatsoever, especially large membrane protein complexes and flexible proteins, which might be too conformationally heterogeneous to pack into a regular crystal lattice.

The most obvious advantage of cryo-EM is the removal of the need of crystals. Furthermore, proteins are imaged in vitreous ice in a state virtually identical to their state in the solution, which eliminates crystal-packing artefacts that have been observed in X-ray crystal structures. Electron cryo-microscopy has been used since the 1970s but its main applicability was low-resolution imaging of large protein complexes and 2D crystallography. Nevertheless, it has been envisaged for a long time that cryo-EM could in principle achieve atomic resolutions (Henderson, 1995). Based on the information content, contrast, electron dose and required resolution, Henderson estimated that in principle particles larger than 10-100 kDa could be aligned to yield an atomic-resolution reconstruction. In fact, because electrons interact much more strongly with matter than X-rays (represented by a higher atomic cross section, Figure 1.16), the information content from electron scattering is much higher than from X-ray scattering per unit of radiation damage.

In the early 2010s, it started to become increasingly obvious that the numerous technological advances have allowed resolutions achievable with cryo-EM to rival X-ray crystallography, an era since termed the resolution revolution (Kühlbrandt, 2014). Since 2012 the number of cryo-EM structures deposited into the PDB has been exponentially increasing. From 439 structures available in 2012 there are now 4884 structures deposited and ever higher resolutions of ever smaller proteins are being achieved. Combined with heavy investments into the cryo-EM infrastructure, cryo-EM is now the method of choice for membrane structural biology. Below I summarize the main principles of cryo-EM and the recent developments which led to the resolution revolution.

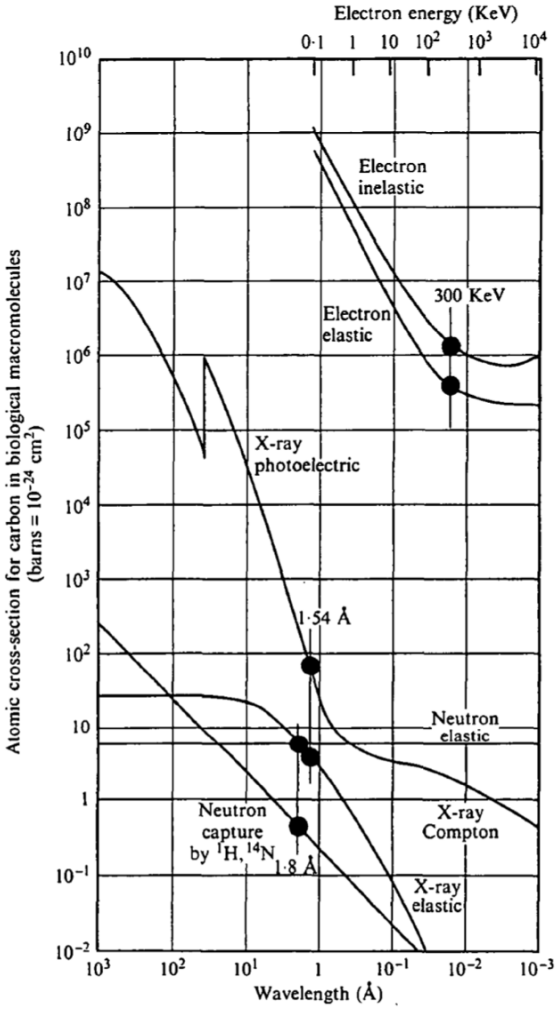


Figure 1.16. Atomic cross-sections for X-ray, neutron and electron scattering.

Adapted from Henderson, 1995.

Sample preparation is a crucial step in cryo-EM. Initially electron microscopy of biological samples was done on stained and desiccated samples which destroys the delicate structure of proteins and limits the amount of information available. Nevertheless, although some insights could be made by studying negatively stained large molecules such as ribosomes and such samples were also instrumental for development of single particle reconstruction algorithms as described below. The main step forward was the development of methods that preserved the biological molecules in more native environments. On the one hand, Henderson and Unwin embedded bacteriorhodopsin crystals in glucose, which preserved the structure of protein much better than heavy metal stains. This led to the first 3-dimensional structure of any membrane protein (Henderson and Unwin, 1975). On the other hand, the group of Jacques Dubochet developed a method of rapidly freezing solutions of protein on the grid such that they become encased in a thin film of amorphous ice (Lepault, Booy and Dubochet, 1983). Two

insights were necessary for the development of this procedure, which remains largely unchanged to this day. Firstly, blotting of a water droplet sitting on a support using a filter paper creates a thin (several tens of nanometers) film that is stable for several seconds during which it can be vitrified. Secondly, flash freezing is required to cool the water film quickly enough to prevent formation of crystalline ice but transform it directly into vitreous ice. This requires plunging the cryo-EM grid into liquid ethane cooled by liquid nitrogen. The use of liquid ethane is crucial as it transfers heat much faster that liquid nitrogen. Since liquid ethane is close to its freezing point, isolating gas bubbles do not form around the grid being cooled, which happens in liquid nitrogen close to its boiling point.

Development of high-end cryo microscopes also contributed significantly to the resolution revolution in cryo-EM. Microscopes need to be manufactured to provide stable parallel illumination with a highly coherent electron beam and the lenses should not introduce aberrations. Microscopes need to operate under a high vacuum to prevent grids from being contaminated. Furthermore, microscopes need to be stable over long periods of time during which automatic image acquisition takes place. Image acquisition must be done in "low-dose" mode, whereby all the alignments are done away from the area that is eventually imaged for collection. This is made possible by automatic collection software such as EPU and SerialEM, which enable automatic collection of images at speeds approaching 200 per hour which can yield a very good dataset in 24h under ideal conditions (Mastronarde, 2005).

Arguably the most important advance was the development of direct electron detectors. Initially the best detector quantum efficiency was that of photographic film. CCD cameras require conversion of electrons into a photon that is detected by the CCD chip. This results in attenuation of the signal at high resolutions and leads to significant blurring. Direct electron detector is able to detect electrons directly on a thin silicon wafer resulting in much more accurate spatial detection of the electron impact event. The main engineering hurdle was making these chips durable enough to be used in cameras. Nowadays, several commercially available cameras exist which have very high detector quantum efficiencies (little or no noise resulting in high contrast and preservation of high-resolution signal) and very fast readouts (Figure 1.17) (McMullan *et al.*, 2014; Nogales, 2015).

Furthermore, during electron beam illumination, the sample moves substantially as a result of charging and simple averaged images are blurred as a result (Glaeser, 2016; Russo and Passmore, 2016). Cryo-EM images are therefore collected as "movies" of about 30-50 frames which are then correcting for beam-induced motion during processing (Brilot *et al.*, 2012).

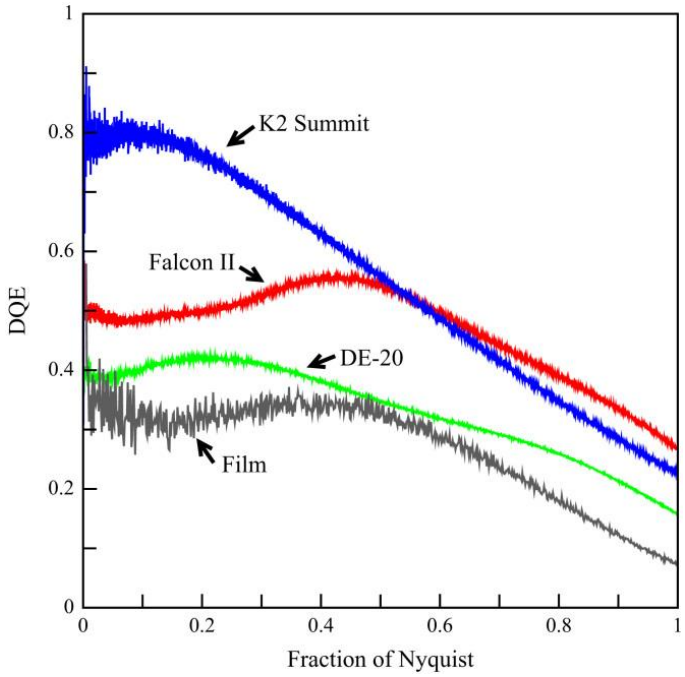


Figure 1.17. Comparison of detector quantum efficiency (DQE) of photographic film and three commercially available direct electron detectors.

Adapted from McMullan et al., 2014.

Because of radiation damage, only small electron doses can be used in cryo-EM image collection (up to ~100 $e^-/Å^2$) before the sample starts to disintegrate. Even then, most high-resolution information is lost after the first 10 $e^-/Å^2$ and the rest of the dose contributes only to low resolution features. This means that particle images are noisy and several thousands of them need to be averaged to achieve an appropriate signal-to-noise ratio. Processing of cryo-EM dataset can be done by several software packages, including RELION, CryoSPARC, cisTEM and others (Scheres, 2012b; Punjani *et al.*, 2017; Grant, Rohou and Grigorieff, 2018). The development of these packages made cryo-EM broadly accessible for the community but the algorithms they employ have been built on the several decades of development of the method of single particles reconstruction.

The mathematical framework for reconstructing 2D projections of particles into a 3D object is the projection-slice theorem which states that a Fourier transform of a projection of an object is a central section of a 3D-Fourier transform of the original object. The main problem to be solved is to find the angles of projection for each of the projection images. The main step forward in the RELION package was the replacement of the discrete assignments of projection orientation by integrating probability-weighted contributions from all the possible assignments

(Scheres, 2012a). This maximum a posteriori (MAP) algorithm uses a Bayesian approach to infer parameters of a statistical model from the data.

The final strength of cryo-EM is the possibility to image heterogeneous protein samples containing multiple conformations of the same protein or even different proteins. Different conformations or molecules can then be separated in silico into the corresponding classes. This complicates particles reconstruction as adding to the lack of information about the relative orientations of the particles, we also lack information about the assignments of the particles to the classes. Perhaps the most powerful feature of cryo-EM image analysis that was indispensable in this work was classification of particles without prior knowledge about their conformations (Scheres *et al.*, 2007). A single low-resolution average structure is used as a reference in the first iteration of likelihood optimisation for K randomly drawn subsets of the data. This results in bias-free seeds, which get updated an improved in each iteration. As the number of different classes in the data is not known, multiple calculations with different values of K need to be performed and carefully compared. Another heuristic value in this process is the T-value, which needs to be carefully adjusted to make use of the available signal in the dataset and prevent overfitting (Scheres, 2012b).

So far many studies have established that occupancies and conformations of different classes of protein molecules in solution can be resolved by 3D classification during processing. One of the recent very convincing examples analysed time-resolved initiation of translation and the populations of 3D classes observed in cryo-EM grids corresponded closely to expected populations based on biochemical and trapping methods (Kaledhonkar *et al.*, 2019). The interpretation that the percentage of different protein classes in the cryo-EM grids are representative of the distribution of those classes in solution is based on the assumption that other factors are not influencing this. One of the factors that could bias occupancies of different classes in cryo-EM is destructive interaction (unfolding/aggregation) of protein molecules at the air-water interface. However, this can be avoided by changing the preparation conditions and any particles that might have undergone air-water interface mediated degradation would have been sorted out by classification and their prevalence would likely be the same across all of the datasets, since grids were prepared under identical conditions.

1.8. Aims of the project

Understanding how the proton gradient is generated, maintained and used for metabolic reactions is a fundamental biological question and arguably the most evolutionarily conserved process in all of life. Using the cryo-EM technique in structural biology means that we can now solve structures of membrane proteins and structurally heterogenous proteins more easily than ever before. Solving structures in different states and at high resolutions (achieved with constantly improving microscopes, cameras and analysis software) allows us to start delineating the chemistry of redox and proton-pumping reactions.

This project aimed to elucidate the redox-coupled proton pumping mechanisms of two fundamentally important but unrelated mitochondrial enzymes, transhydrogenase and complex I, which are discussed in Chapters 3 and 4, respectively. In the initial stages of the project I also investigated *E. coli* complex I and ovine F_1F_o-ATPase as potential model systems but did not pursue them beyond optimising purification and cryo-EM grid preparation protocols. The results on these two proteins are briefly discussed in Chapter 2. Finally, in Chapter 5, I compare the common principles of proton translocation in two different ancient and widespread protein families and what they tell us about the protein function and evolution in general.

The main aims of the project were:

1. <u>To determine the near-atomic resolution structure of intact NNT.</u>

The first aim of the present project was to provide the first, to our knowledge, high-resolution structure of an intact transhydrogenase. The only reported structure of the complete transhydrogenase comes from X-ray studies of the *T. thermophilus* NNT at the resolution of 7 Å (Leung *et al.*, 2015). Failure to crystallise transhydrogenase to high resolution could reflect conformational heterogeneity, hence we chose a cryo-EM approach for this project which confirmed intrinsic conformational heterogeneity but also proved to be less limited by it. Near-atomic resolution structure of an intact NNT presented an important advance for the structural understanding of this essential bioenergetic enzyme. It allowed a structural interpretation of a large amount of biochemical and mutagenesis data on bacterial model enzymes. Furthermore, structure of an intact transhydrogenase allowed a structural and mechanistic interpretation of a number of described mutations from patients. Perhaps further studies will shed some light on the phenotypic heterogeneity in transhydrogenase-deficient patients. As transhydrogenase is an emerging drug target candidate availability of the structure will perhaps enable structure-based

drug discovery targeting transhydrogenase. This is particularly relevant in the light of the fact that no good specific inhibitors of transhydrogenase are currently known.

2. To understand the coupling mechanism of NNT.

Transhydrogenase operates by an entirely unique and poorly understood mechanism among the redox-driven proton pumps. There is evidence that large-scale conformational changes occur during its catalytic cycle. To understand the mechanism of NNT, one or more catalytic intermediates were structurally and biochemically characterised. Since it was shown that conformational changes occur upon binding of different substrates (NADP(H)), we characterised theses states which allowed us to propose a consistent catalytic mechanism for transhydrogenase. It explains all the previously described biochemical properties of the enzyme, its mutants and the structural data generated here.

3. To understand the catalytic mechanism of complex I

The near-atomic resolution structures of complex I have been becoming available from several species since 2013. However, our understanding of the mechanism is poor, firstly because the resolution achieved was not very high (up to 3.3 Å) and not many conditions were investigated. The aim of this project was to improve the resolution to be able to model water molecules, important for the proton pumping process. The active and stable preparation of ovine complex I in LMNG was used to achieve this. This was complemented by comparing cryo-EM structures of ovine complex I in several different conditions and describe the changes induced by substrate binding, turnover, inhibition, and deactivation. Finally, we collected a dataset during turnover and classified different conformations of complex I that exist during the catalytic turnover. This allowed us to describe the exact chemistry of the complex I coupling process and to provide experimental evidence for a coupling process in complex I that depends both on the conformational changes and electrostatic interactions.

2. Materials and methods

2.1.Materials

Materials were obtained from Sigma-Aldrich Corporation (St. Louis, Missouri, USA) unless stated otherwise. All the detergents were obtained from Anatrace, Inc. (Maumee, Ohio, USA).

Thio-$NADP^+$ was purchased from Santa Cruz Biotechnology Inc. (Heidelberg, Germany). Bio-Beads SM-2 were obtained from Bio-Rad Laboratories, Inc. (Hercules, California, USA).

Zeba spin desalting columns and Detergent Removal Spin columns were sourced from Thermo Fisher Scientific Inc. (Rockford, Illinois, USA). Chromatography columns were obtained from GE Healthcare (Chicago, Illinois, USA). Ovine hearts were obtained from a local butcher (Retz, Austria). Vivaspin concentrators were obtained from Sartorius Stedim (Göttingen, Germany).

Electron microscopy grids were obtained from Quantifoil Micro Tools GmbH (Jena, Thuringia, Germany).

2.2.Analytical Methods

2.2.1. Protein concentration measurements

2.2.1.1.BCA assay

Protein concentration was measured using the colorimetric method with the Pierce bicinchoninic acid (BCA) assay kit (Thermo Fisher Scientific Inc., Rockford, Illinois, USA). A standard curve was obtained by preparing bovine serum albumin (BSA) at concentrations between 0.1 and 1 mg/mL in the same buffer as the unknown sample. Upon mixing samples with reagents in 96 well plates, they were incubated at 37 °C for 30 min and then read using a SpectraMax M2 plate reader (Molecular Devices, Wokingham, Berkshire, UK). Protein concentration in the sample was determined by comparison with the standard curve.

2.2.1.2.Spectrophotometric method (A_{280})

Aromatic amino acid side chains absorb UV radiation at 280 nm, a property that can be used to estimate protein concentration in pure samples for which extinction coefficients are known. This method was used because it is relatively precise in the ranges of 1-5 mg/mL required for cryo-EM grid preparation and because only a small volume (1 μL) of the sample is required for a measurement on the NanoDrop 1000 UV-Vis spectrophotometer device

(Thermo Fisher Scientific Inc.). For transhydrogenase, which does not contain any bound cofactors, extinction coefficient at 280 nm in water was predicted *in silico* (with PredictProtein) to be 726601 M^{-1} cm^{-1} (Yachdav *et al.*, 2014). 1 mg/mL NNT thus corresponds to absorption 0.64 AU. For complex I, however, the absorption coefficient had to be estimated empirically due to the large size of the molecule and many cofactors bound. 1 mg/mL complex I corresponded to A_{280} of about 2.3 AU as determined by comparison with the BCA assay.

UV-Vis spectra of protein samples measured on the NanoDrop were also used as a rough estimate of protein aggregation. $A_{260/280}$ ratio is ~0.6 for pure monodisperse protein but rises to ~1 for aggregated protein because of the Rayleigh light scattering from protein aggregates (Fiedorczuk, 2017). Rayleigh scattering increases to the 4th power with dropping wavelength and results in a shift in the peak of absorbance (Penzkofer *et al.*, 2007).

2.2.2. Polyacrylamide gel electrophoresis

SDS-PAGE

Protein samples (~10-50 ng) were mixed in 4:1 ratio with sample buffer consisting of 250 mM Tris pH 6.8, 10% (w/v) SDS, 10% (w/v) DTT, 50% (v/v) glycerol and 0.05% (w/v) bromophenol blue. 10-15 μL samples were loaded into the pre-cast 4-20% gradient Novex Tris-Glycine polyacrylamide gel (Invitrogen, Thermo Fisher Scientific) which was run for 80 min at 150 V in an electrophoresis buffer containing 25 mM Tris pH 8.3, 192 mM glycine and 0.1% SDS. Molecular weights of the resulting bands were estimated by comparison to the Unstained Precision Plus Protein Standards (Bio-Rad) which were run alongside the lanes with unknown samples. Gels were stained using Instant Blue stain (Expedeon, Abcam, Cambridge, UK) for several hours and destained in water. For better resolution of small subunits, Novex Tricine 10-20% gels run according to the manufacturer's protocol proved marginally better. Criterion XT Bis-Tris Precast Gels (Bio-Rad) were not appropriate for studies of multi-subunit protein complexes due to the poor resolution between bands.

Blue native PAGE

Blue native PAGE was done according to the manufacturers protocol (Invitrogen, Thermo Fisher Scientific). Briefly, protein samples were mixed with sample buffer (50 mM BisTris, 50 mM NaCl, 10% w/v Glycerol) and loaded on the NativePAGE 3-12% Bis-Tris Protein Gels. The gels were run at 4ºC for ~1h at 150 V with the dark blue cathode buffer (50 mM BisTris 50 mM, Tricine pH 6.8, 5% Cathode Buffer additive (0.4% Coomassie G-250)) and another 1.5h at 250 V with the light blue cathode buffer (50 mM BisTris 50 mM, Tricine

pH 6.8, 0.5% Cathode Buffer additive). Anode buffer consisted of 50 mM BisTris 50 mM, Tricine pH 6.8.

2.2.3. Detergent concentration measurement

Detergents were measured using a colorimetric assay for quantifying sugars as described before (Urbani and Warne, 2005). A standard curve was prepared with 0-0.03% solutions of a detergent to be measured in 0.005% increments. 50 µl of standards and sample dilutions were mixed with 300 µl 5% Phenol in 2 ml capped tubes, to which 720 µl conc. sulphuric acid was immediately added drop-wise, mixed and cooled. Absorbance at 490nm was measured and compared to the standard curve to obtain the detergent concentration in the sample.

2.2.4. Complex I activity measurements

Complex I catalyses NADH:ubiquinone oxidoreduction in vivo, however, this reaction is not easy to measure in vitro due to the high hydrophobicity of ubiquinone. Potassium ferricyanide can be substituted as an electron acceptor, but the reaction happens at the FMN site, is not inhibited by rotenone and is therefore not physiologically relevant. It is nevertheless useful for determining the presence of complex I in impure samples. Decylubiquinone, a soluble analogue of the natural substrate ubiquinone can be used instead to assay the physiological reaction of complex I – DQ is soluble enough in an appropriate buffer and the reaction is fully inhibited by rotenone.

2.2.4.1.NADH:DQ oxidoreductase activity

Complex I NADH:quinone oxidoreductase activity was measured at 30°C spectrophotometrically by following the NADH ($\varepsilon = 6.1\ mM^{-1}\ cm^{-1}$) oxidation at 340 nm using a Shimadzu UV-2600 UV-VIS spectrophotometer (Shimadzu UK Ltd., Milton Keynes, Buckinghamshire, UK) fitted with a temperature controller and automated sample changer (Shimadzu DPS-240A) with a magnetic stirring unit (Electronic Stirrer Model 300, Rank Brothers Ltd. Cambridge, UK). Reaction buffer was optimized for the ovine CxI purified in LMNG before (20 mM HEPES, pH 7.4, 50 mM NaCl, 1 mM EDTA, 0.1% CHAPS, 0.1% LMNG, 0.25 mg/mL lipids (1:4 mixture of cardiolipin and dioleoyl phosphatidylcholine), 100 µM NADH and 200 µM DQ) (Letts *et al.*, 2016). Complex I was added to the above buffer without NADH and incubated at 30ºC for 5 min before the reaction was started by NADH

addition. Reaction was measured for 120 s and the initial activity was calculated from the slope of the absorbance curve in the interval between 30 and 90 s using a Beer-Lambert law and expressed as μM NADH consumed per mg protein per minute, i.e. U/mg.

When SMPs or non-solubilised samples from various stages of purification were used, 10 μM CCCP (10 mM stock in ethanol) was added to prevent proton gradient formation. When inhibitors were used, they were prepared as stocks in DMSO (100 mM imiquimod or 1 mM rotenone).

2.2.4.2.NADH:Potassium Ferricyanide oxidoreductase activity

Complex I NADH:FeCy oxidoreductase activity measurement was done at 30ºC in the same buffer (20 mM HEPES, pH 7.4, 50 mM NaCl, 2 mM EDTA, 0.1% LMNG, 1 mM $KFe(CN)_6$, 100 μM NADH) as described before (Letts *et al.*, 2016). This reaction was also followed spectrophotometrically by measuring the NADH ($\varepsilon = 6.1\ mM^{-1}\ cm^{-1}$) oxidation at 340 nm.

2.2.4.3.Superoxide production assay

Superoxide production assay was done under the same conditions as the above assays. The buffer used consisted of 20 mM HEPES, pH7.4, 100 mM NaCl, 0.25 mg/mL DOPC:CL 4:1, 0.1% CHAPS, 400 μM KCN. 20 μg/mL complex I was added to this buffer and after 5 min incubation, 240 μg/mL of cytochrome c was added, followed by 100 s measurement of background. 100 μM NADH was added, followed by 200 s measurement of the activity and finally, 40 U/mL of superoxide dismutase were added followed by 120 s measurement of residual activity. If inhibitors were used, they were added as a DMSO stock to 15 nM (rotenone) or 36 μM (imiquimod). KCN was added to the buffer to inhibit cytochrome c oxidation by miniscule amounts of residual complex IV in the purified complex I sample. Our ovine complex I preparation produced ~20nmol/min/mg superoxide.

2.2.4.4.NEM labelling assay

Labelling with NEM was performed by incubating a sample of 1 mg/mL of purified complex I with 1 mM NEM on ice for 5 min, after which NEM was neutralized by the addition of 1 mM DTT. Activity of complex I labelled with NEM was performed as described above.

2.2.5. Transhydrogenase activity measurements

Transhydrogenase can catalyse reverse, forward and cyclic transhydrogenation, which can be followed spectroscopically. Because NAD(H) and NADP(H) both have absorption maximum at 340 nm, different analogues need to be used for these assays. NAD^+ analogue, 3-acetylpyridine adenine nucleotide, is used in the reverse transhydrogenation reaction and thio-$NADP^+$ is used to follow forward transhydrogenation.

2.2.5.1. Optimisation of the transhydrogenation assay

In the literature, several different buffer conditions (Wu, Alberta and Fisher, 1986; Sazanov and Jackson, 1995) have been used in the transhydrogenase assay, which mostly contained Brij-35, a detergent for which we showed that it modestly inhibits transhydrogenase. When solubilized in LMNG or FOM, transhydrogenase has more than twofold higher activity than in Brij-35 which was previously reported as the most activity-compatible detergent (Figure 2.1d) (Hu, Zhang and Rydström, 1998). Hence, we substituted it for CHAPS and determined optimal concentrations for other components as well. Detergent and lipids in reaction buffer significantly increase activity – 0.1% CHAPS and DOPC:CL mixture used routinely. pH of the buffer had a broad maximum pH6.5-pH7.5 (Figure 2.1a). Salt concentration had a similar broad maximum between 50-100mM NaCl (Figure 2.1b). As shown before, buffers can also inhibit transhydrogenase and because MES/HEPES slightly decreased the activity, we used Tris in the activity assay buffer. 100 μM substrates were optimal for activity, and activity declined above and below this value (Figure 2.1c).

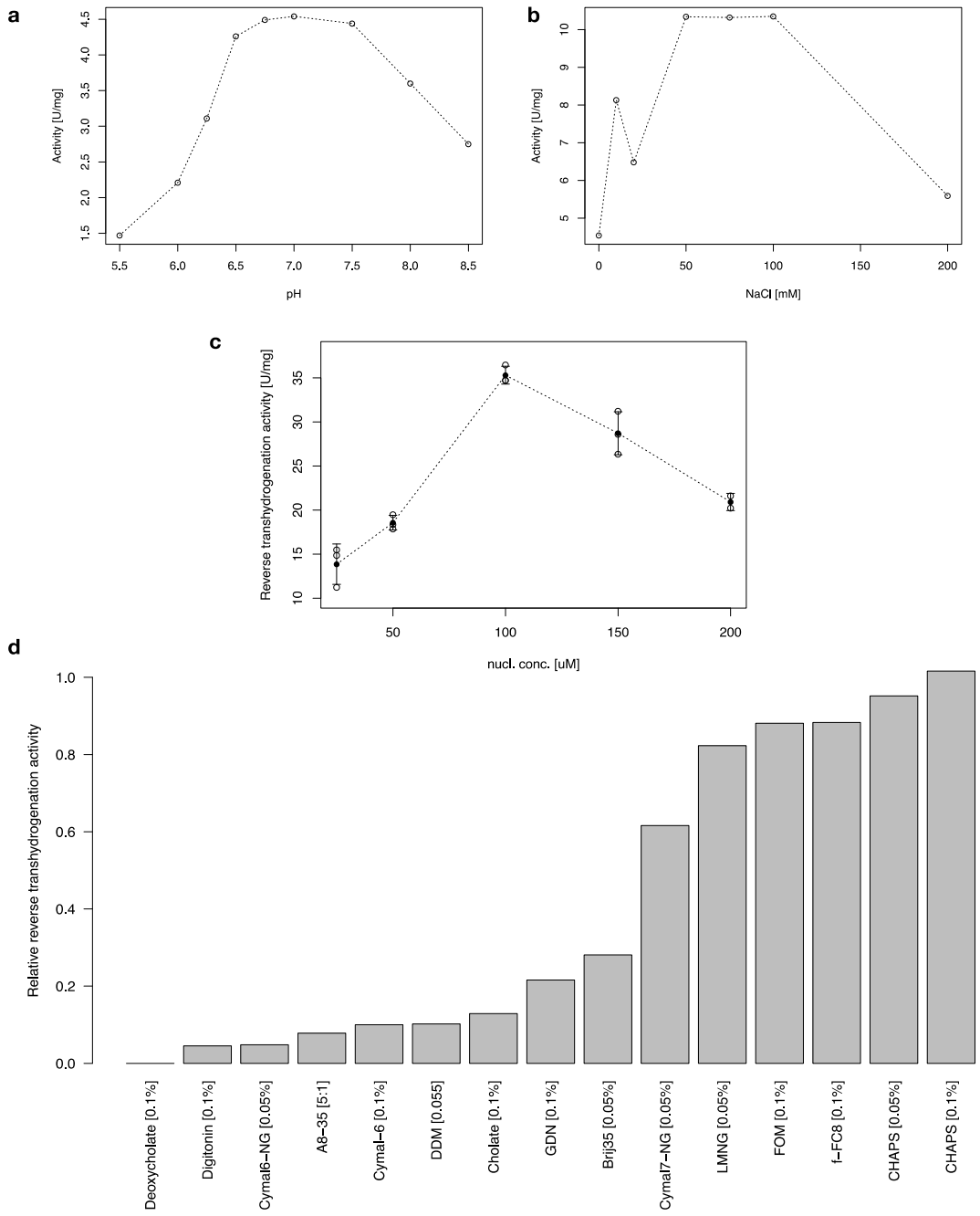


Figure 2.1. Optimisation of transhydrogenase activity assay.

a. *Reverse transhydrogenation activity has a broad pH optimum between 6.5 and 7.5.* ***b.*** *Reverse transhydrogenation activity has a [NaCl] optimum between 50 and 100 mM. c. NADPH and APAD⁺ in equimolar ratio were added to the reaction and the reaction has the optimum at 100 μM each. d. Detergents have a profound effect on reverse transhydrogenation activity. The highest activities were achieved in CHAPS, fluorinated detergents and branched chain detergents, while most other commonly used detergents inhibited NNT strongly.*

2.2.5.2.Reverse transhydrogenation activity

Reverse transhydrogenation activity was measured at 30°C spectrophotometrically by following the $APAD^+$ (3-acetylpyridine adenine dinucleotide; NAD^+ analogue; $\varepsilon = 5.1\ mM^{-1}\ cm^{-1}$) reduction at 375 nm using a Shimadzu UV-2600 UV-VIS spectrophotometer. Background absorption at 455 nm was subtracted from absorption at 375 nm. Final optimised reaction buffer consisted of 20 mM Tris-HCl, pH 6.8, 50 mM NaCl, 0.5 mM EDTA, 0.1% CHAPS, 0.25 mg/mL lipids, 100 μM NADPH and 100 μM $APAD^+$. For lipids we used soybean asolectin or DOPC:CL 4:1 mixture with similar results. When membrane-containing samples were measured, 10 μM CCCP was added to prevent pmf formation. Reaction was started with the addition of NNT.

When done with NNT reconstituted into proteoliposomes, reaction was started by substrate additions (NADPH and $APAD^+$ at 100 μM). CCCP (10 μM) was added after 1 min to decouple the proton gradient across proteoliposomes to assess the dependence of the transhydrogenation reaction on the proton motive force. After another 1 min, CHAPS (0.05%) was added to solubilise proteoliposomes and give a total decoupled reverse transhydrogenation activity of both inward and outward facing transhydrogenases. The degree of reaction acceleration reflects the degree of coupling between the hydride and proton transfer reactions and the leakiness of the proteoliposomes (negligible).

2.2.5.3.Cyclic transhydrogenation activity

Cyclic reaction was measured in the same buffer as above by following $APAD^+$ reduction by NADH at 375 nm. Because our NNT preparation was devoid of nucleotides, 30 μM NADPH was added to the protein sample first and incubated on ice for 10 min. Alternatively, it was added directly in the reaction cuvette and after it was consumed as confirmed by a stable baseline, 100 μM NADH was added to start the cyclic reaction.

2.2.5.4.Forward transhydrogenation activity

Forward transhydrogenation was measured in the above buffer. Instead of NADH, NADPH was added and instead of $APAD^+$, thio-$NADP^+$ was added ($\varepsilon = 11.7\ mM^{-1}\ cm^{-1}$) and the reaction was followed by measuring absorption at 398 nm.

2.2.6. ATP-ase activity measurement

ATP-ase activity was measured as described previously (Runswick *et al.*, 2013). Buffer used consisted of 50 mM Tris-HCl pH 7.4, 50 mM KCl, 2 mM $MgCl_2$, 0.1% CHAPS, 0.4 mg/mL soybean asolectin, 400 μM NADH, 1 mM PEP, 2 mM ATP, 12 U/mL LDH and 12 U/mL PK. Reaction was measured at 37ºC and was started by the addition of ATPase. Consumption of NADH was followed spectrophotometrically at 340 nm ($\varepsilon = 6.1\ mM^{-1}\ cm^{-1}$) using a Shimadzu UV-2600 UV-VIS spectrophotometer. Coupling between F_1 and F_o was probed by measuring sensitivity of this reaction to oligomycin (added to 0.1 mg/mL from an ethanol stock).

2.2.7. Trypsinolysis of NNT

Trypsinolysis of NNT was performed as described previously with a few changes (Yamaguchi, Wakabayashi and Hatefi, 1990). NNT at 0.05 mg/ml in GF buffer with only 1 mM HEPES pH 7.4 was incubated with 0.4 mM nucleotides, trypsin (at different mass ratios) and 60 mM BAT buffer (1:1:1 Bis-Tris, acetate, tricine) at pH 5-8. Due to the pH optimum of trypsin being ~8, trypsin:NNT mass ratio between 1:400 and 1:40 and incubation times between 30 and 60 min were used at different pH values 5-8.

2.2.8. Unfolding assay

It has been demonstrated that optimal buffer conditions are crucial for obtaining stable monodisperse protein samples suitable for high-resolution structure determination by cryo-EM (Chari *et al.*, 2015). An unfolding assay can detect shifts in melting temperature of multi-subunit protein complexes and thus give the most optimal buffer system and pH (Chari *et al.*, 2015). Unfolding at higher temperatures exposes hydrophobic regions of the proteins which can be detected using a fluorescent SYPRO orange dye. Because some of our samples were purified in detergent, the contribution of detergent to the signal had to be subtracted. 0.1mg/mL protein was used for all samples (NNT in FOM, NNT-1E3 or F1Fo in LMNG).

For NNT in FOM, the best five buffer conditions were: HEPES pH 7.2, MES pH 5.7, SPG (succinic acid, NaH_2PO_4, glycine) pH 7.2 and pH 6.8 and Na-phosphate buffer 5.8. The first two conditions were almost identical to the buffers used during purification as described later. For F_1F_o-ATPase in LMNG, the best five buffer conditions were: Na-citrate pH 5.5 or 7.0 or 7.25, Bis-Tris pH 5.8 and MES pH 6.9. Differences in melting temperatures between different buffers were very small, however.

2.3. Sample preparation

All the protein purification procedures described below were done at 4°C. All the proteins were purified from ovine heart mitochondria as described below. During protocol optimisation, several proteins (transhydrogenase, F_1F_o-ATPase and complex I) were sometimes purified from the same starting material. However, during the final stages of sample preparation optimisation and preparation of final grids for high resolution data collection, only one protein was purified from a single batch of mitochondria. This is because the proteins of interest elute over a large range of ionic strengths. Complex I elutes the last at ~250 mM NaCl and is subjected to a long gradient wash required to elute transhydrogenase (at ~110 mM NaCl) and ATPase (at ~150 mM NaCl). Such a long wash actually increased the yield of complex I (due to its release from supercomplexes) but also reduced the activity of complex I, potentially due to the delipidation as a result of extensive washing (Figure 2.2b-c). Avoiding a long gradient elution and eluting ATP-ase and transhydrogenase containing material with a step gradient before complex I elution led to too impure transhydrogenase and ATPase fractions. They would require another loading on the anion exchange column, potentially also resulting in suboptimal final protein samples. The best approach was thus to purify CxI separately and NNT and ATPase together from the same batch of mitochondria.

Purifications 1-6 which were mainly used for protocol optimisation were done on a single batch of mitochondria stemming from sheep hearts obtained in Chelmsford, UK. All the subsequent purifications were done on a single batch of mitochondria from sheep hearts from Retz, Austria. All the samples used for high resolution were therefore collected from Austrian sheep mitochondria.

2.3.1. Mitochondrial isolation

Purification procedures were done as described before (Letts *et al.*, 2016). Briefly, mitochondria were purified from fresh ovine heart tissue; the hearts were cooled on ice within an hour of slaughter. Subsequent transport and mitochondrial isolation took ~10 hours. Mitochondria were purified from fresh ovine heart tissue by differential centrifugation and stored at –80°C according to the Procedure 3 by Smith (Smith, 1967). Briefly, heart muscle tissue was minced and washed in 1.4 L of buffer (250 mM sucrose, 10 mM Tris-HCl pH 7.8, 5 mM 2-mercaptoethanol) per 1 kg of tissue. Washed tissue was blended for 30 s using a Wahring blender and then centrifuged at 1800 g for 15 min to remove large debris and then the supernatant was centrifuged at 20000 g for 27 min to pellet mitochondria. Dry mitochondrial pellets were frozen and stored at -80ºC.

2.3.2. Isolation of inner mitochondrial membranes

On the day of purification, inner mitochondrial membranes were isolated as described before (Letts *et al.*, 2016). Briefly, 10 g of mitochondria were ruptured by homogenisation (Potter-Elvehjem homogeniser, ~50 strokes) in 100 mL milliQ water. Upon addition of 150 mM KCl (from 2 M stock), membranes were separated by 45 min centrifugation at 50000 g, supernatant was discarded and the pelleted membranes were resuspended in 100 mL of buffer M (20 mM HEPES, pH 7.4, 40 mM NaCl, 1 mM EDTA, 10% v/v glycerol, 2 mM DTT and 0.002% PMSF). After another round of resuspension and centrifugation, membranes were resuspended in 50 mL of buffer M.

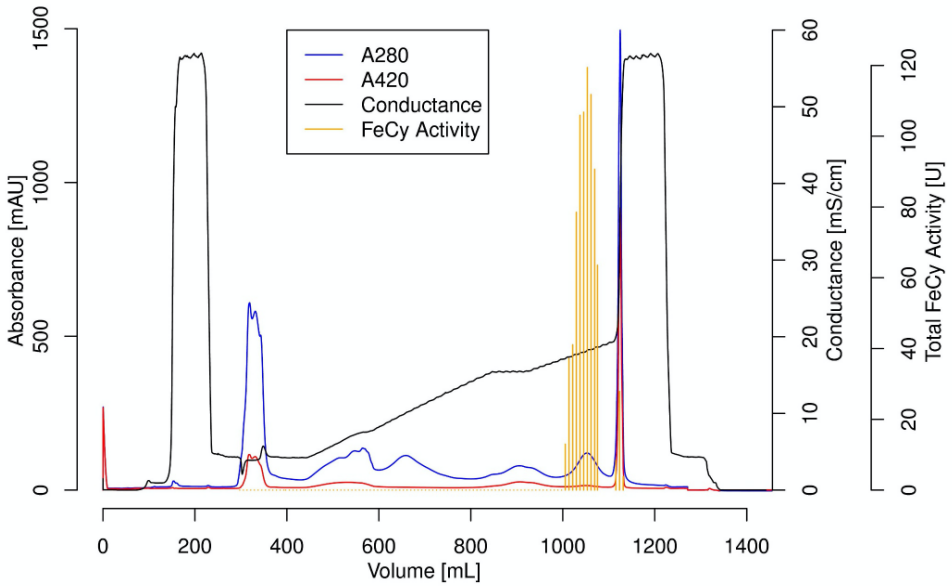
2.3.3. Complex I preparation

2.3.3.1. Anion exchange chromatography

Inner mitochondrial membranes were solubilised in 50 mL of buffer M (20 mM HEPES, pH 7.4, 40 mM NaCl, 1 mM EDTA, 10% v/v glycerol, 2 mM DTT and 0.002% PMSF, 1% LMNG) for 45 minutes. Upon centrifugation (50000 g, 45 min) the supernatant (~50 mL, depending on the purification) was filtered with a 0.22 μm filter and loaded onto a 45 mL Q-sepharose HP anion exchange column equilibrated in buffer Q-A (20 mM HEPES pH 7.4, 40 mM NaCl, 2 mM EDTA, 10% v/v glycerol, 1 mM DTT, 0.05% LMNG). The column was then washed with 75 mL Q-A buffer, 50 mL of 10% Q-B buffer (Q-A with 1 M NaCl), 125 mL of 20% Q-B buffer and finally complex I was eluted by a 200 mL linear gradient with 20-27% Q-B buffer (Figure 2.2a). Complex-I containing fractions (~50 mL) were pooled, concentrated to ~15 mg/mL (~1 to 2 mL) using a Millipore 100-kDa cut-off filter and stored at 30% glycerol under liquid nitrogen in aliquots of 250 μL.

SDS-PAGE profile of selected fractions from the anion-exchange column elution reveals the pattern of elution of the major respiratory chain complexes and some other abundant mitochondrial proteins (Figure 2.3).

a



b

Prep4: long method

	NADH:FeCy activity per mL [U/ml]	total activity [U]	activity per mg protein [U/mg]	protein concentration [mg/mL]
Mitochondria	11.0	1321.6	2.7	4.09
Wash 1	0.2	21.0	N/A	N/A
Membranes	11.0	1104.6	2.3	4.86
Wash 2	0.0	0.0	N/A	N/A
Solubilised membranes	18.5	923.7	2.6	7.11
Qload (supernatant)	13.8	636.0	2.7	5.04
Pellet	18.4	220.9	2.1	8.86
pooled fractions	12.1	676.74	~23	~0.6
	yield	61%		

c

Prep5: short method

	activity per mL [U/ml]	total activity [U]
Membranes	17.3	1731.6
Wash	0.4	95.7
Qload (supernatant)	23.0	1150.4
Pellet	25.4	405.7
pooled fractions	5.97	382.2
high salt wash	8.75	210
	yield	22%

Figure 2.2. Purification of complex I from ovine heart tissue.

a*. Anion exchange elution profile for a typical long-method complex I purification. b. Yield and purification table for long-method complex I purification. c. Yield and purification table for short-method complex I purification.*

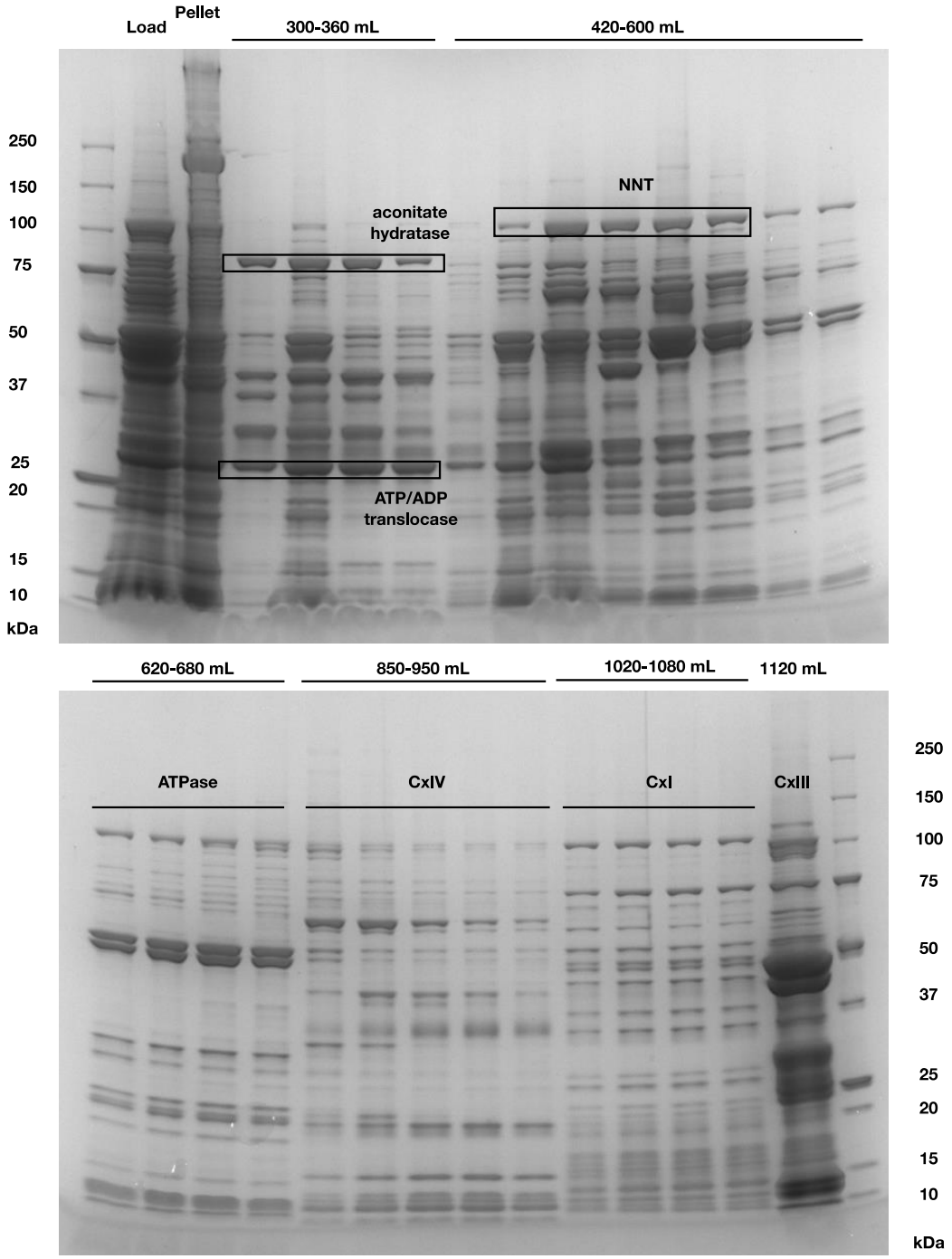


Figure 2.3. SDS-PAGE profile for a typical purification of complex I from ovine heart tissue.

Aconitate hydratase and ATP/ADP translocase were annotated by MS/MS, and the rest of the proteins were annotated based on the SDS-PAGE pattern, A_{280} and A_{420} and activity assays.

2.3.3.2. Size exclusion chromatography

After anion exchange chromatography, complex I was ~50% pure and size exclusion chromatography was used as the last step of the purification mainly to exchange the detergent or to lower the LMNG concentration. It also removed the remaining contaminants: pyruvate dehydrogenase (PDH) and 2-oxoglutarate dehydrogenase (OGDH) complexes (as shown by MS/MS) and complex IV (as determined by SDS-PAGE and activity measurements) (Figure 2.4f).

A 250 μL sample of complex I after anion exchange was thawed on ice on the day of cryoEM grid preparation, loaded onto a Superose 6 10/300 gel filtration column equilibrated in buffer GF (20 mM HEPES, pH 7.4, 50 mM NaCl, 1 mM EDTA, 0.002% LMNG) to remove excess detergent and remaining protein contaminants. The purest and the most concentrated fractions of complex I eluted at ~11.5 mL and were concentrated to 3 mg/mL and used immediately for cryo-EM grid preparation (Figure 2.4a,f).

Additionally, complex I samples were also prepared by gel filtration in the following detergents: Cymal-7 (0.02%), FOM (0.05% or 0.1%), LMNG (0.05%, 0.01%) and sample reconstituted into amphipol A8-35 (Figure 2.4b-d). The gel filtration pattern in these conditions did not change with the exception of cymal-7 in which the peak had a slightly asymmetric shape, potentially indicating breaking of the complex.

Changing glycerol concentration also did not alter the appearance of the gel filtration profile or the SDS-PAGE pattern of the fractions. Glycerol concentrations at 0, 1, 2, and 10% were tested. In the final preparations for high-resolution data collection, glycerol was omitted to decrease the background noise during cryo-EM image collection. Changing NaCl concentration (50, 100 or 200 mM) also did not change the outcome of the gel filtration. Final grids for data collection were prepared at 50 mM NaCl to minimize background noise.

It was noted that to minimize aggregation of particles on the grid, protein sample needs to be applied to the grid and frozen immediately after gel filtration. Freezing sample aliquots before the gel filtration step did not seem to have this effect. Any freeze-thaw cycles after the last step of purification however, increased aggregation and led to inferior grids.

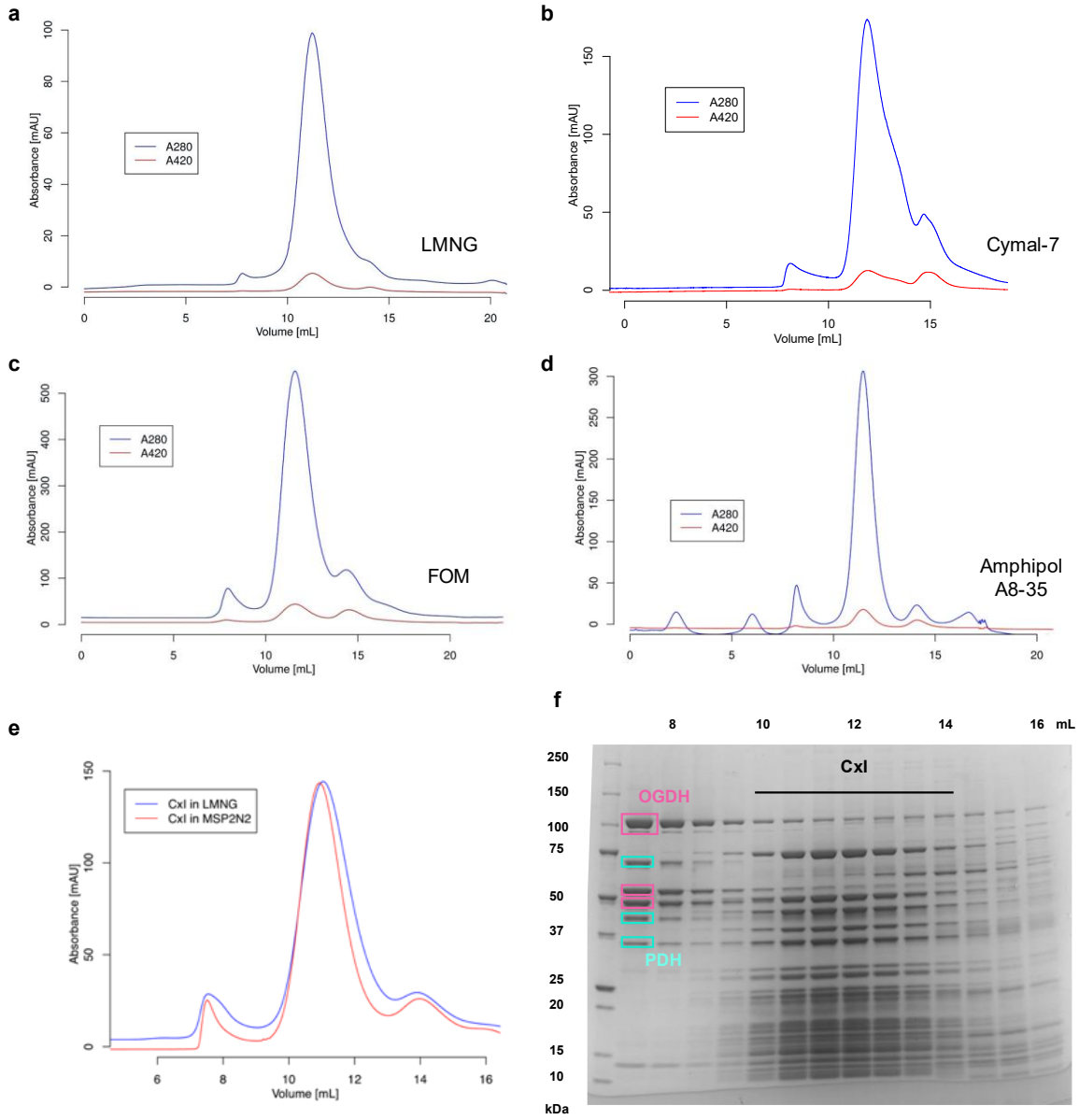


Figure 2.4. Final preparations of ovine complex I.

*Gel filtration profile reveals elution of CxI in (**a**) LMNG-containing buffer, (**b**) Cymal-7-containing buffer, (**c**) FOM-containing buffer, (**d**) A8-35 reconstituted CxI. **e**. Comparison of CxI in naodiscs and reconstituted into MSP2N2 nanodiscs. **f**. SDS-PAGE profile of a typical gel filtration preparation. Void volume contains the remaining contaminants 2-oxoglutarate dehydrogenase complex (OGHD) and pyruvate dehydrogenase complex (PDH) which were identified by MS/MS and labelled on the gel.*

2.3.3.3. Deactive complex I preparation

Deactive complex I was prepared by incubating the concentrated aliquot of complex I after anion exchange at 32ºC for 30 min. To prevent delipidation and proteolysis, 0.5 mg/mL of 1:4 mixture of cardiolipin and dioleoylphosphatidylcholine and 0.0075% PMSF were added to the sample. Afterwards, the sample was cooled on ice and injected to the gel filtration as described above.

2.3.3.4. Preparation of complex I using sucrose gradient

Complex I in digitonin was purified on a hand-made 10-45% sucrose gradient to which a sample of complex I after anion exchange chromatography in LMNG was loaded. The buffer was the same as for gel filtration with additional 0.1% digitonin. Tubes with gradients were centrifuged at 4ºC for 21 h at 130000 g in a swing-bucket rotor before collecting the complex I-containing fractions. On cryo-EM grids, the sample in digitonin looked monodisperse but contained micelles. It also displayed slightly lower activity in digitonin compared to FOM/LMNG, hence this preparation was not optimized further.

2.3.3.5. Nanodisc preparation

Nanodiscs are protein-enveloped lipid bilayer discs into which membrane proteins can be reconstituted to allow their structural characterization in close-to-native lipidic environment (Bayburt and Sligar, 2010). Nanodisc preparation might prove superior to detergent because it allows the protein to reside in the lipid environment which was demonstrated to have a stabilizing effect on membrane proteins and lead to higher resolution structures (Jin *et al.*, 2017).

Complex I purified in LMNG was reconstituted into nanodiscs by mixing complex I, scaffold protein MSP2N2 and 1:4 CL/DOPC lipid mixture in a 1:2:100 stoichiometry. The mixture was incubated on ice for 15 min, upon which 400 mg/mL washed BioBeads SM-2 were added in two batches 2 hours apart. The mixture was incubated overnight with constant stirring and then injected on a Superose 6 column equilibrated in a gel filtration buffer without any added detergent. Complex I eluted in a symmetric peak at around 11.5 mL, almost indistinguishable from detergent-purified sample (Figure 2.4e).

2.3.4. Transhydrogenase preparation

Ovine heart tissue is an excellent native source of mitochondria and has been used in the Sazanov group as a starting material for the purification of the respiratory complex I as described above and elsewhere (Fiedorczuk *et al.*, 2016; Letts *et al.*, 2016). Transhydrogenase is sufficiently abundant in ovine mitochondria to allow purification of adequate amounts of the enzyme for cryo-EM analysis. Purification from a native source is often superior to heterologous expression because it surpasses the need for expression and provides a highly functional and properly post-translationally modified enzyme. Ovine transhydrogenase also exists in only one described isoform and its sequence identity to human NNT is 97%, hence the ovine enzyme can be considered a very good model of the human enzyme.

Previous protocols reported anion exchange chromatography or affinity chromatography using NAD+-linked agarose (Wu, Alberta and Fisher, 1986; Sazanov and Jackson, 1993). However, these procedures resulted in poor yields and activities. Hence, we developed a new procedure that resulted in pure, highly active and stable transhydrogenase from the ovine heart tissue suitable for further biochemical and structural analyses. Choice of detergent was one of the critical factors for stability and activity and the results from activity assay optimization gave initial detergents to screen. LMNG was used in the first purification steps that were adapted from complex I purification protocol. Other detergents that did not inhibit NNT were CHAPS, FOM, fFC8 and Cymal7-neopentylglycol. Detergents which inhibited NNT were DDM, digitonin, A8-35, cymal6, GDN and brij35 (Figure 2.1).

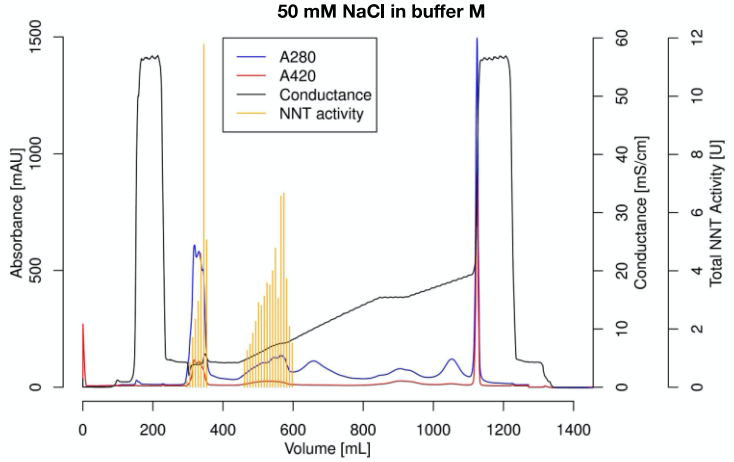
2.3.4.1. Anion exchange chromatography

The first step of purification was adapted from the complex I purification protocol described before (Letts *et al.*, 2016). Two crucial changes were made (Figure 2.5). Firstly, to achieve complete binding of NNT to Q-sepharose NaCl concentration in buffer A was decreased from 50 mM to 40 mM. Secondly, an additional gradient elution step was added between 40 mM and 200 mM NaCl. The final protocol for anion exchange purification of NNT was the following. LMNG (10 %) was added dropwise to this suspension to 1 % and after stirring for 45 min, the sample was centrifuged at 50000 g for 45 minutes. The supernatant was filtered with a 0.22 μm filter and loaded onto a 45 mL Q-sepharose HP anion exchange column equilibrated in buffer Q-A (20 mM HEPES pH7.4, 40 mM NaCl, 2 mM EDTA, 10% v/v glycerol, 1 mM DTT, 0.05% LMNG). Proteins were eluted with a 400 mL linear gradient with 0-17% Q-B buffer (Q-A buffer with 1 M NaCl). NNT eluted as a broad peak (~150mL) around

110 mM NaCl and fractions containing significant amounts of NNT based on the activity assay and the SDS-PAGE profile were pooled.

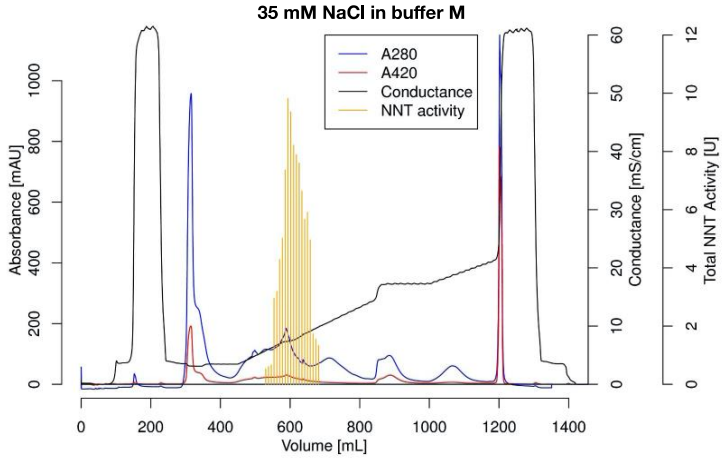
It quickly became apparent that a single round of anion exchange did not yield sufficiently pure NNT for cryo-EM studies. Changing the pH of buffers A and B to pH 8 or pH 8.5 did not change the separation of the components significantly. A second round of purification by anion exchange was tried on partially purified fractions from the first round. Binding of partially purified NNT fractions to MonoQ, ANX-sepharose and DEAE-sepharose columns equilibrated in buffers A and B at pH 7.4 did not significantly improve the purity of the NNT preparation (Figure 2.6a). Hence, a different separation method was needed.

a



	NADPH consumed [μM per min]	activity per mg protein [μmol per min per mg]	protein concentration [mg/ml in stock]	total activity in sample [μmol per min]
Mitochondria	2.00	0.24	4.09	120.00
Wash 1	0.33			3.61
Membranes	3.00	0.31	4.86	150.00
Wash 2	0.31			2.80
Solubilised membranes	7.72	0.54	7.11	193.03
Qload (supernatant)	5.98	0.59	5.04	137.62
Pellet	3.64	0.21	8.86	4.37
Flow-through	4.61	~0.36	~1.28	29.47
Pooled fractions	5.41	~0.77	~0.69	60.41
yield	**29%**			

b



	NADPH consumed [μM per min]	total activity in sample [μmol per min]
Mitochondria	2.15	257.51
Membranes	1.26	126.07
Solubilised membranes	2.36	118.11
Qload (supernatant)	2.50	124.92
Pellet	1.20	14.42
Pooled fractions	8.52	74.99
yield		**60%**

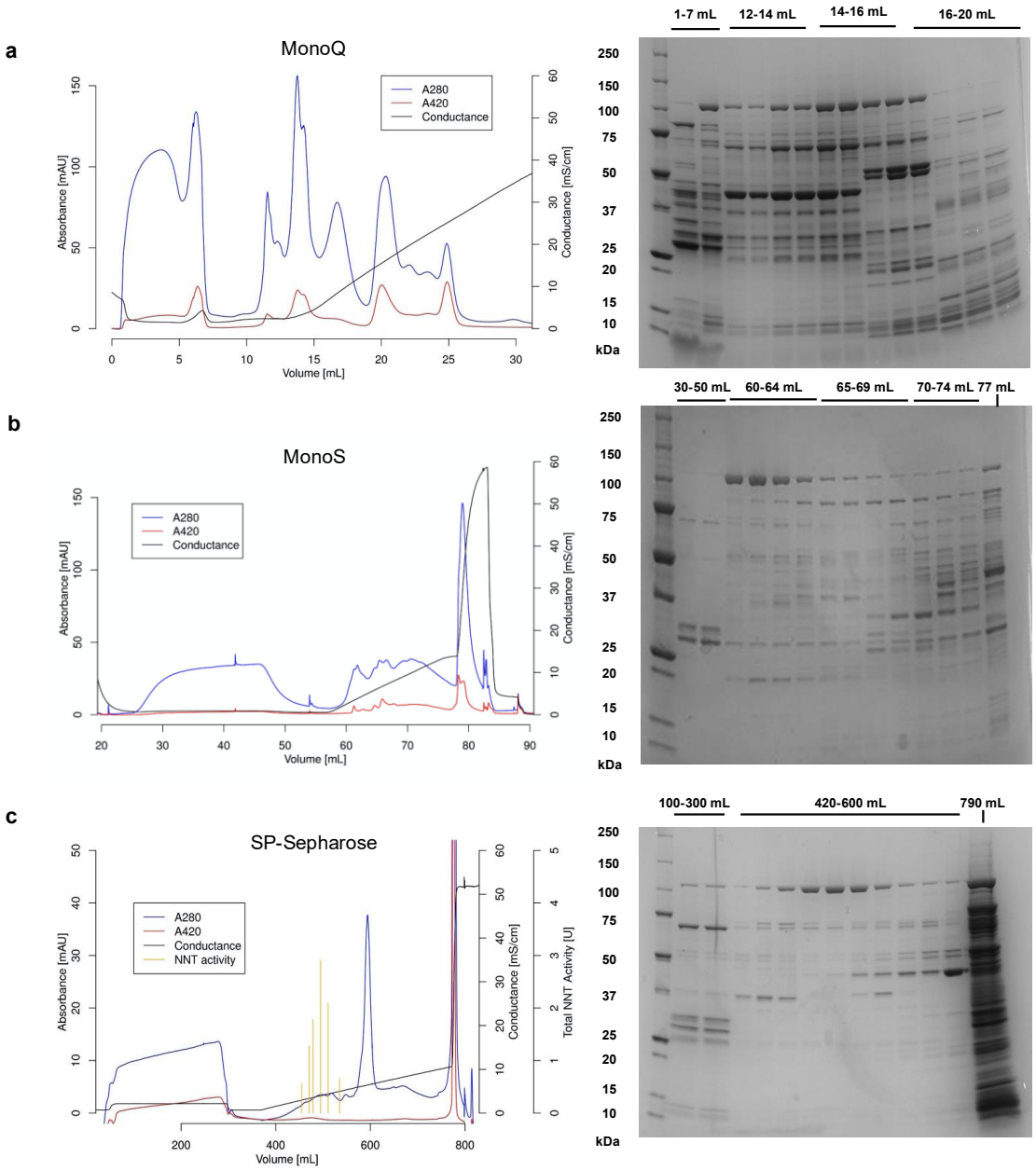
Figure 2.5. Optimising NNT anion exchange step.

***a**. Elution profile of NNT in high salt buffer. **b**. elution profile of NNT in low salt buffer.*

2.3.4.2.Cation exchange chromatography

Purifying NNT-containing fractions by cation exchange chromatography using MonoS column at pH 6.0 gave encouraging results because NNT eluted separately from other contaminants (Figure 2.6b). Because there was some evidence of NNT dissociating on the MonoS column and in particularly in order to scale up the preparation, SP-sepharose was used in further preparations with indistinguishable results (Figure 2.6c,d). Further tests also showed that because NNT binding to cation exchange column is relatively weak, it is important to load NNT in a buffer with NaCl concentration below 10 mM (or conductivity < 2 mS/cm) and at pH less or equal to 6. This was effectively achieved by dilution or dialysis.

The final protocol for cation exchange of NNT involved the following steps. Fractions containing transhydrogenase from Q-sepharose column were dialysed overnight against a 10-fold volume excess of buffer D (20 mM MES, pH 5.8, 2 mM EDTA, 10% v/v glycerol, 1 mM DTT, 0.01% LMNG). This lowered the pH and conductance of the sample below 6 and 2 mS/cm, respectively, and made it suitable for loading onto a 45 mL SP-Sepharose HP cation exchange column equilibrated in buffer SP-A (20 mM MES pH5.8, 10 mM NaCl, 2 mM EDTA, 10% v/v glycerol, 1 mM DTT, 0.025% LMNG). Elution with a 100 mL 0-20% linear gradient with buffer SP-B (SP-A with 1 M NaCl) gave a sharp elution peak of relatively pure NNT at around 35 mM NaCl (conductance 4 mS/cm). Fractions containing NNT were concentrated to 4 mg/mL using a 100 kDa cut-off concentrator and stored at 30% glycerol under liquid nitrogen.



d

	NADPH consumed [µM per min]	total activity in sample [µmol per min]
load	0.99	12.36
flow-through	0.08	1.02
pooled fractions	5.64	13.40

Figure 2.6. Optimising NNT second ion exchange step.

***a**. Elution profile of pooled fractions containing NNT from Fig. 2.5. when loaded on a second anion exchange column MonoQ. **b**. Elution profile of pooled fractions containing NNT from Fig. 2.5. when loaded on a second cation exchange column MonoS at pH 6.0 gives much better separation. **c**. Scaled-up cation exchange step using a 50 mL SP-sepharose column. **d**. Yield table for SP-sepharose preparation.*

2.3.4.3.Size exclusion chromatography and polishing

Gel filtration of NNT after the cation exchange step was crucial to exchange the buffer and adjust the pH and to lower the LMNG concentration or to exchange the sample into a different detergent. On the day of cryo-EM grid preparation, a sample of NNT was thawed on ice and loaded onto a Superose 12 10/300 gel filtration column equilibrated in buffer GF (20 mM HEPES, pH 7.4, 50 mM NaCl, 1 mM EDTA, 0.002% LMNG) to remove excess detergent and remaining protein contaminants (Figure 2.7b,c). The purest and the most concentrated fractions of NNT eluted at ~10.7 mL and were concentrated to 5 mg/mL using a Millipore 100-kDa cut-off filter and used immediately for cryo-EM grid preparation.

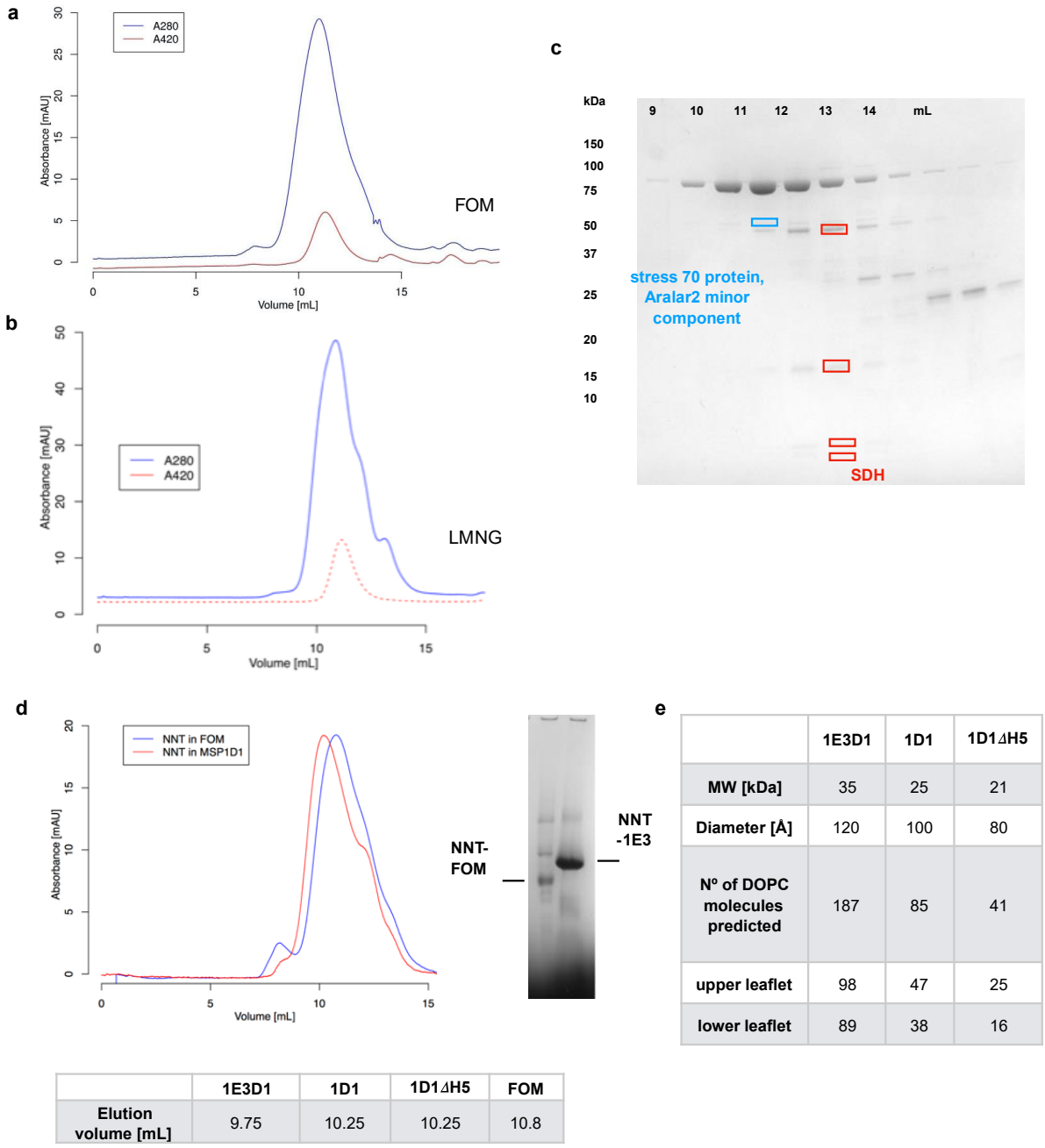
In some purifications, 0.002% LMNG in GF buffer was substituted by 0.05% FOM, which gave a monodisperse and highly active NNT but grids suitable for cryo-EM data collection could not be prepared from FOM preparations as the protein started to aggregate on the grid at concentrations higher than 2 mg/mL (Figure 2.7a). If NNT was purified using Superose 6 column, the peak of elution was at ~14 mL.

Preparations in FOM had a slightly higher activity (~20 U/mg) compared to LMNG (~15 U/mg) but were considerably less stable: after 1 week at 4ºC almost 90% of activity of LMNG sample was retained, while the FOM sample lost any detectable activity.

NNT prepared in this way contained a small but persistent amount of complex II contaminant (as determined by SDS-PAGE and UV-Vis spectrum), which due to similar mass (~160 kDa) did not separate completely from transhydrogenase (~220 kDa) on the gel filtration column and required one to carefully exclude CxII-containing fractions from samples used for cryo-EM grid preparation (Figure 2.7c).

To remove complex II contamination additional purification steps were tried. Adapting the SP-Sepharose preparation by eluting with a pH gradient instead of salt gradient was tried but did not work well. Loading of SP-Sepharose in buffer with pH 6.1, collecting the flowthrough and reloading it a pH 5.8 gave slightly higher purity at the expense of a lower yield. Similarly, NNT seemed to bind to reactive red agarose preferentially over complex II, but the preparation was not easily reproducible, required high salt concentration for elution (500 mM NaCl) and resulted in a decreased yield. Furthermore, F1Fo also appeared to bind to this substrate. Overall, any additional purifications steps did not increase the purity significantly to justify prolonged purification times and lower yield.

Transhydrogenase prepared in this way was devoid of nucleotides, as confirmed by the absence of cyclic reaction without the addition of exogenous $NADP^+$ (data not shown).



	1E3D1	1D1	1D1ΔH5	FOM
Elution volume [mL]	9.75	10.25	10.25	10.8

e

	1E3D1	1D1	1D1ΔH5
MW [kDa]	35	25	21
Diameter [Å]	120	100	80
Nº of DOPC molecules predicted	187	85	41
upper leaflet	98	47	25
lower leaflet	89	38	16

Figure 2.7. Final NNT purification step and reconstitution into liposomes.

a. Gel filtration of NNT in two different detergents, FOM and LMNG. Absorbance at 420 nm is due to the succinate dehydrogenase (complex II, SDH) as shown by SDS-PAGE. c. SDS-PAGE shows the presence of a ~110 kDa polypeptide of NNT. Highest purity fractions around ~10.5 mL were pooled and concentrated for cryo-EM sample preparation. A major contaminant SDH (bands in red) was identified based on A420 and SDS-PAGE profile. Minor contaminants (bands in blue) were identified by MS/MS. d. Comparison of gel filtration elution profiles for NNT in FOM and several nanodiscs. BN-PAGE shows a difference in the apparent molecular weight as well. e. Total number of DOPC lipids predicted in silico to assemble into differently sized nanodiscs with bacterial NNT (PDB 4O9U).

2.3.4.4.Nanodisc reconstitution

Nanodisc preparation was done as described above for complex I. We varied the time of reconstitution (speed of detergent removal), and different ratios of components. Overnight reconstitution (two additions of 200 mg Biobeads per 400 μg NNT) and a unit stoichiometric ratio between NNT and scaffold protein was the best. At ratios of NNT to MSP to lipids 1:2:200 we observed many empty nanodiscs, at ratio 1:1.2:100 less so and at 1:1:100 almost none and we also did not observe any decrease in NNT yield or increased void volume elution that would signal too little MSP or lipids were added.

The choice of membrane scaffold protein depends on the size of the transmembrane region and should be large enough to allow at least two layers of lipids to bind between NNT and MSP. On the other hand, it should be small enough not to add significant additional heterogeneity and unstructured density into the sample, which would complicate subsequent image analysis. The size of nanodiscs was modelled using CHARMM-GUI online server (Figure 2.7e) (Jo *et al.*, 2008).

The choice of lipids should resemble natural membrane lipid composition and should have a melting temperature below 4 ºC at which the protein was purified. DOPC:CL in 4:1 w/w ratio was used, which was determined to be adequate based on activity measurements.

NNT was reconstituted into MSP1E3D1, MSP1D1 and MSP1D1ΔH5 nanodiscs and injected on the Superose 12 gel filtration column equilibrated in gel filtration buffer without any additional detergent. NNT-1E3 eluted at ~9.7 mL and had activity ~20 U/mg, while NNT-1D1 and NNT-1ΔH eluted at ~10.25 mL and had activities ~12 U/mg. NNT-1E3 migrated significantly higher on BN-PAGE than NNT-FOM and based on the gel filtration profile, the nanodisc adds about 200 kDa to the overall apparent molecular weight of the particles (ferritin which is 440 kDa elutes at 9.7 mL from Superose 12 column) (Figure 2.7d).

2.3.4.5.Proteoliposome reconstitution

NNT was reconstituted into proteoliposomes using a detergent dilution procedure as described before (Sazanov and Jackson, 1995). Briefly, NNT purified in FOM or in LMNG was mixed with a 500-fold molar excess of DOPC solubilised in 1% CHAPS. Following a 10 min incubation on ice, the solution was diluted 100-fold in GF buffer and incubated on ice for 3-5 hours before used for measuring the activity directly. Activity measurements confirmed that proteoliposomes prepared in this way were well coupled (~10x stimulation of transhydrogenase activity by the addition of protonophore). To assess the effect of lipids on the

NNT activity, NNT was also reconstituted into proteoliposomes with different lipid compositions, in particular POPC, asolectin, and DOPC:CL, DOPC:PS, DOPC:PE, DOPC:PI in 4:1 w/w ratios, respectively. Activity was the highest in DOPC/POPC liposomes, slightly lower in DOPC/CL, DOPC/PS or DOPC/PI and the lowest in DOPC/PE. POPC liposomes were slightly leakier than the others (Chapter 3).

2.3.5. ATPase purification

2.3.5.1.LMNG extracts ATPase dimers

F_1F_o ATPase is another abundant inner mitochondrial protein that was observed to elute from the Q-Sepharose anion exchange column as a broad but relatively pure peak at a slightly higher NaCl concentration (140 mM) than NNT during the purification of the latter. Curiously, when the sample of ATPase was gel filtrated in the following buffer, 50 mM Tris-HCl, pH 7.4, 50 mM KCl, 2 mM $MgCl_2$, 0.05% LMNG, the resulting profile displayed a main peak at ~12 mL and a distinct shoulder at ~10.7 mL. Blue-native PAGE confirmed that the low-molecular weight peak corresponded to a monomer of ATPase, while the high-molecular weight corresponded to dimeric ATPase. Previously, it was reported that only weak detergents such as digitonin are able to extract ATPase from mitochondria in the dimeric form.

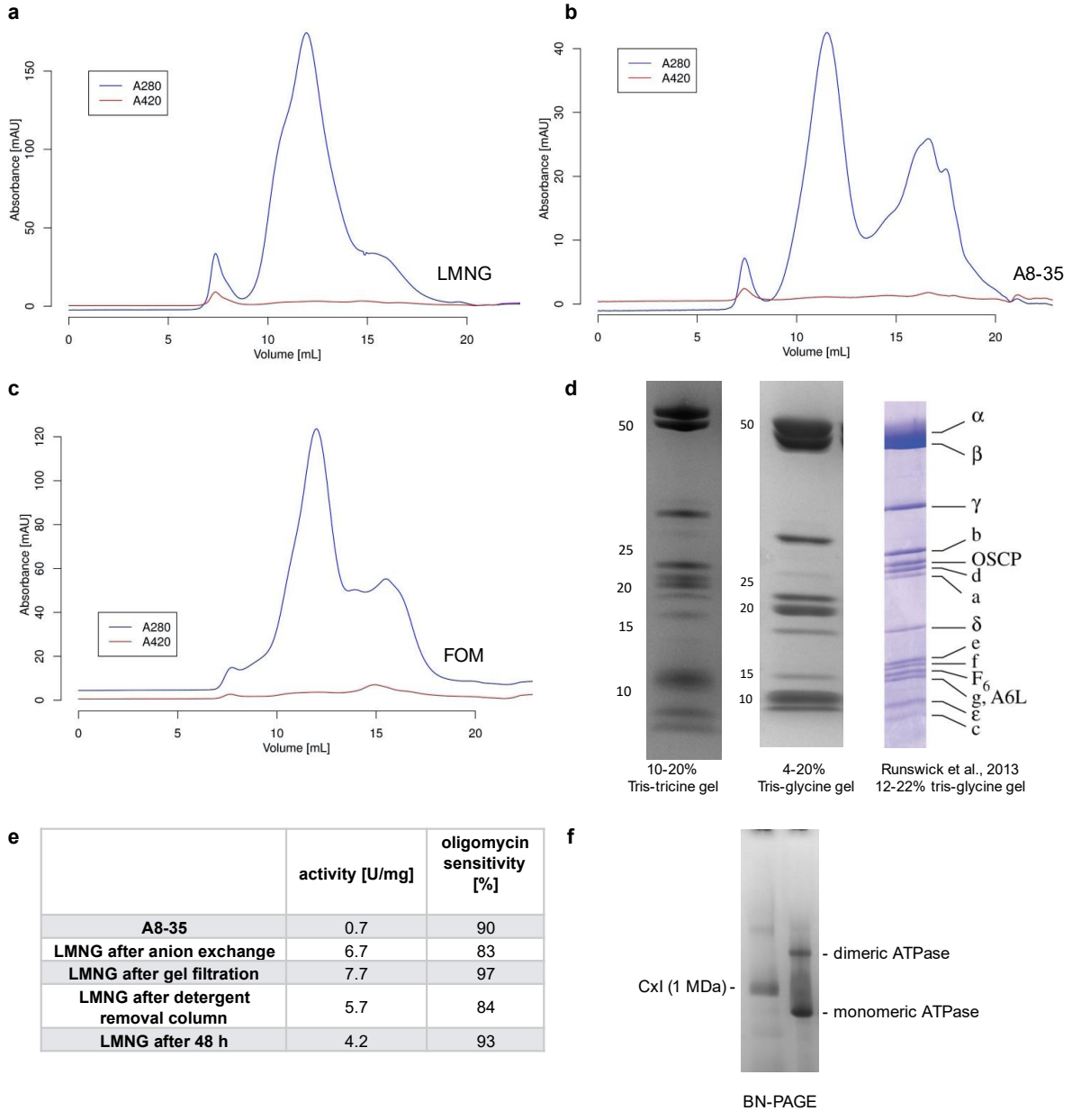
Gel filtration was done with similar results also in 0.02% LMNG and 200 mM KCl. When LMNG was substituted for FOM or ATPase was exchanged into amphipol A8-35, the high-molecular weight peak practically disappeared.

To improve separation between monomer and dimer, several purification methods were tried. Density gradient purification was done in 5-45% sucrose gradient along with an opposite 0.003-0% LMNG gradient as described before in the GRADER approach (Hauer *et al.*, 2015). Similarly, the gradient was also done with glycerol instead of sucrose. The separation was comparable to gel filtration. Finally, separation by MonoQ column at different pH values (pH7.4, 8, 9.2) was tried, but dimer and monomer peaks always overlapped. Gel filtration of the dimer-containing frozen fractions results in a monomer peak only.

The final optimised GF buffer based on the results of ProteoPlex consisted of 50 mM Na-citrate buffer pH 7.0, 100 mM NaCl, 1% glycerol, 0.05% LMNG, 1 mM EDTA.

ATPase was reconstituted into MSP2N2 and DOPC:CL 4:1 nanodiscs according to the procedure described above for complex I with a 72 h reconstitution procedure. F_1F_o-ATPase, MSP, lipid ratio was 1:3:300, respectively. Only monomers reliably reconstituted, possibly because the complex has begun to dissociate during the long reconstitution procedure.

F1Fo prepared this way displayed activity between 4 and 8 U/mg and between 83 and 97% oligomycin sensitivity. Reconstitution in to amphipols A8-35 and 10 mM $CaCl_2$ fully inhibited the ATPase activity.



	activity [U/mg]	oligomycin sensitivity [%]
A8-35	0.7	90
LMNG after anion exchange	6.7	83
LMNG after gel filtration	7.7	97
LMNG after detergent removal column	5.7	84
LMNG after 48 h	4.2	93

Figure 2.8. Summary of ATPase purification.

a-c. Gel filtration profiles in (a) 0.05% LMNG, (b) A8-35 reconstituted sample and (c) 0.05% FOM. d. SDS-PAGE analysis of purified ATPase using Tris-tricine and Tris-glycine gels and compared to the SDS-PAGE profile of bovine ATPase published in literature. e. Activity and oligomycin sensitivity of several different preparations as described. f. BN-PAGE analysis reveals a mixture of monomeric (~600 kDa) and dimeric (~1.2 MDa) ATPase. In the left lane, complex I in LMNG is run for comparison.

2.4.Electron cryo-microscopy

2.4.1. Grid preparation

Unless stated otherwise, grids were prepared by applying 2.7 µL protein sample to a freshly glow-discharged ($7x10^{-1}$ mbar and 30 mA for 2 min in the ELMO Glow Discharge unit, Agar Scientific, Stansted, UK) Quantifoil 0.6/1 copper grid and blotted for 4-10 s using a blotting force of 25 at 4°C and 100% humidity in an FEI Vitrobot Mark IV. Grids were flash-frozen in liquid ethane and stored in liquid nitrogen until data-collection.

2.4.2. General lessons on grid preparation

Cryo-EM grids suitable for high-resolution data collection must satisfy a number of criteria and achieving grids suitable for high resolution cryo-EM structure determination is currently the most important bottleneck of many projects. Most importantly, protein needs to be pure, monodisperse, sufficiently concentrated, show random orientations and not be disassembled or aggregated (Orlova and Saibil, 2011). Moreover, the ice thickness needs to be low – both to minimize background noise and to limit distribution of protein particles along the z-axis.

To achieve this grid preparation protocol needs to be iteratively optimized and a number of different conditions screened. The first crucial component is isolating the protein of interest under the right conditions, especially with respect to the choice and concentration of detergent, buffer and protein concentration. The second crucial component is optimization of conditions during the vitrification process itself, e.g. the grid type and blotting time. Many optimised parameters will be unique for each protein examined, as confirmed during the preparation of this thesis which included cryo-EM grid optimisation for four different proteins and several different preparations of each of them. However, some patterns in the grid optimisation process did emerge which will be briefly discussed here. The most common problems encountered included: aggregation, problems with ice thickness, breaking of particles, presence of detergent artefacts and heterogeneity within or between grids.

2.4.2.1.Achieving consistent results

The first requirement for consistency is a pure, stable and fresh protein sample that behaves in a predictable way when concentrated and applied on a cryo-EM grid. Grid preparation was always done from freshly gel-filtrated sample. Diluting and then re-concentrating protein samples from thawed glycerol-containing stocks resulted in aggregation

of protein on the cryoEM grids (Figure 2.9). In general, the fewer the handling steps of protein between gel filtration and grid freezing, the better the grids.

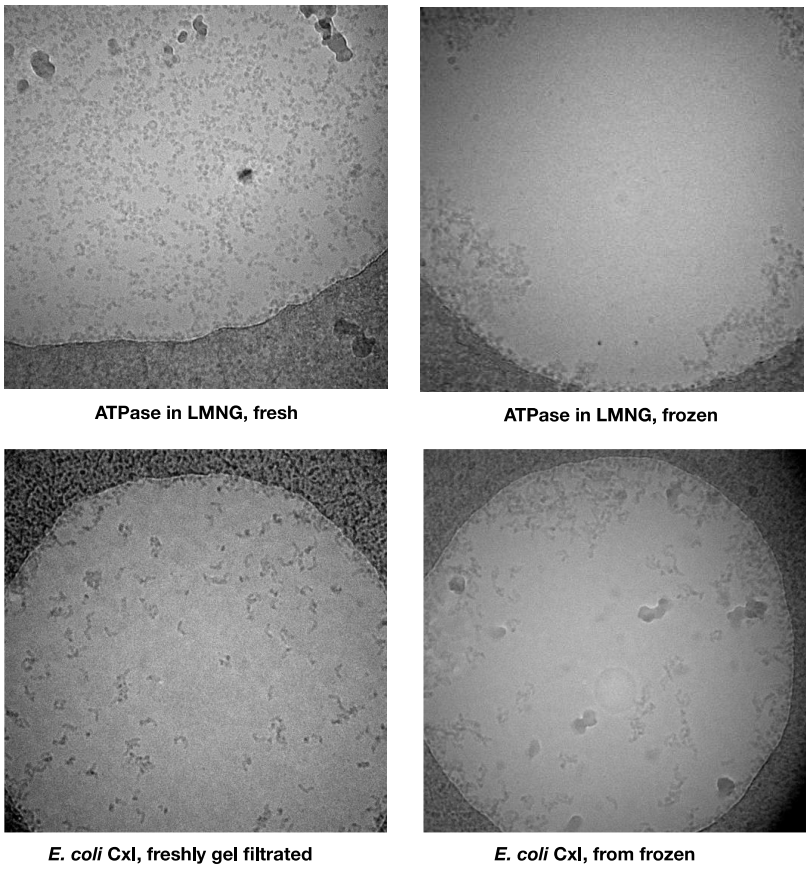


Figure 2.9. Effects of gel filtration on grid preparation in ATPase (top) and E. coli complex I (bottom).

ATPase in LMNG, fresh grid was prepared immediately after gel filtration without any further concentration or additons. E. coli complex I, fresh grid was prepared by gel filtration in DDM and concentration to ~5 mg/mL. ATPase and CxI frozen grids were prepared by diluting-concentrating a sample of protein frozen with 30% glycerol.

The other critical step for achieving consistency are blotting conditions. Care was taken to operate the Vitrobot as consistently as possible on different sessions. This included setting the temperature, blotting "force" and blotting "time" parameters, maintaining the same orientation of the tweezers and insertion of blotting paper right before the actual grid preparation and not during Vitrobot equilibration. This achieved a satisfactory reproducibility but a representative grid from the same batch was always still checked in-house before the other grids were used for high-resolution data collection elsewhere.

Sometimes it was observed that particle coverage varied quite strongly within the same grid. Ultimately, this seems to be an effect of too low protein concentration and higher protein concentration solved the variability.

2.4.2.2.Aggregation

Aggregation was one of the most persistent problems encountered during cryo-EM grid preparation. Membrane proteins tend to aggregate easily because of the exposed hydrophobic regions normally embedded in the membranes. The most obvious solution for aggregation is therefore a change of detergent. Typically, detergents in which proteins easily aggregate can be discarded already during the purification optimization step, by stability and activity screening. However, that is not always the case as demonstrated by NNT in FOM, as described later. Another possible solution is including a secondary detergent, which was beneficial in many cases, especially if the primary detergent had a low critical micelle concentration (CMC) and the secondary detergent had a high CMC (e.g. FOM, CHAPS, OG) (Figure 2.10). High-CMC detergent molecules orient themselves at air-water interface and prevent protein from interacting with it, but some partitioning of secondary detergent into primary detergent micelles also probably happens and contributes to different sample properties (Chen *et al.*, 2019).

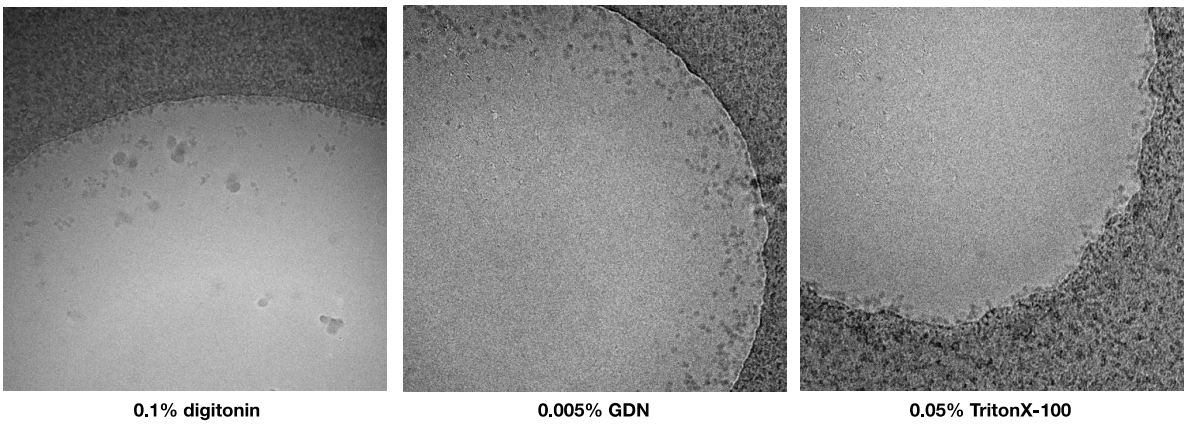


Figure 2.10 Effects of secondary detergents on aggregation of ATPase on cryoEM grids.

Secondly, salt concentration determined the aggregation of particles in solution. Salt shields charges as the cations arrange themselves around negatively charged protein surfaces and anions arrange around positively charged protein surfaces. Salt can have two opposing effects on aggregation behavior. Simple colloidal particles can be induced to aggregate by increasing salt concentration as the dissolved ions shield repelling charges on their surfaces

(Dill and Bromberg, 2010). Proteins, however, have large and heterogeneous surfaces with both positively and negatively charged patches. Opposing charges on the surfaces will also be shielded by increased salt concentrations and associate less readily. The effect of salt on protein aggregation was documented for *E. coli* complex I, which was very prone to aggregate at 50 mM NaCl, but when this was increased to 250 mM NaCl, aggregation was more or less gone and increase to 400 mM did not improve monodispersity further (Figure 2.11).

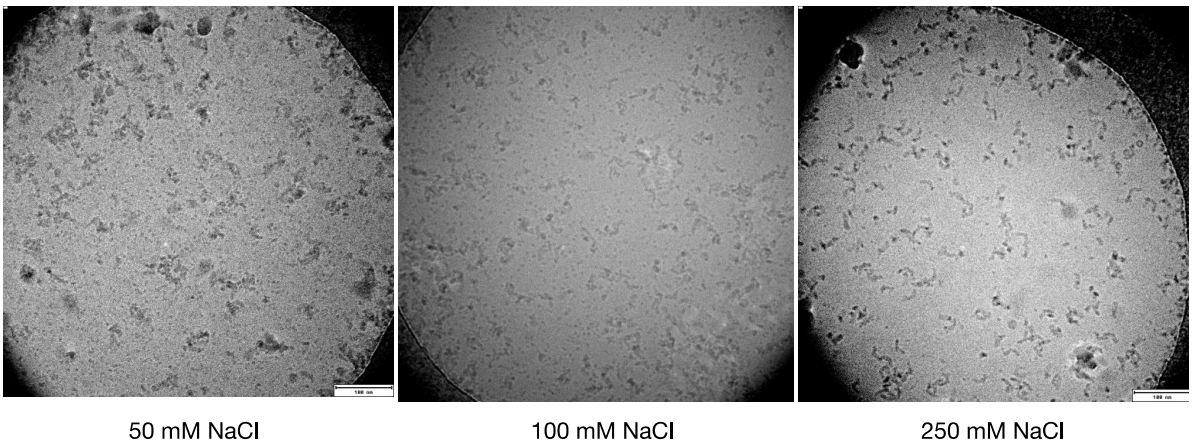


Figure 2.11. Effects of salt on aggregation of E. coli complex I on cryoEM grids.

An alternative method for modulating surface charges is a change of pH. This can modulate the protein surface charges by changing protonation states of amino acid residues. The effects of pH change on the aggregation of F_1F_o-ATPase sample are not as easy to explain, however (Figure 2.12). Surface charge properties of large protein assemblies are complex and consist of regions of opposing charges which will behave in unpredictable ways at different pH values. Varying pH could also affect protein activities and conformations in unpredictable ways and was hence abandoned as an optimizable parameter.

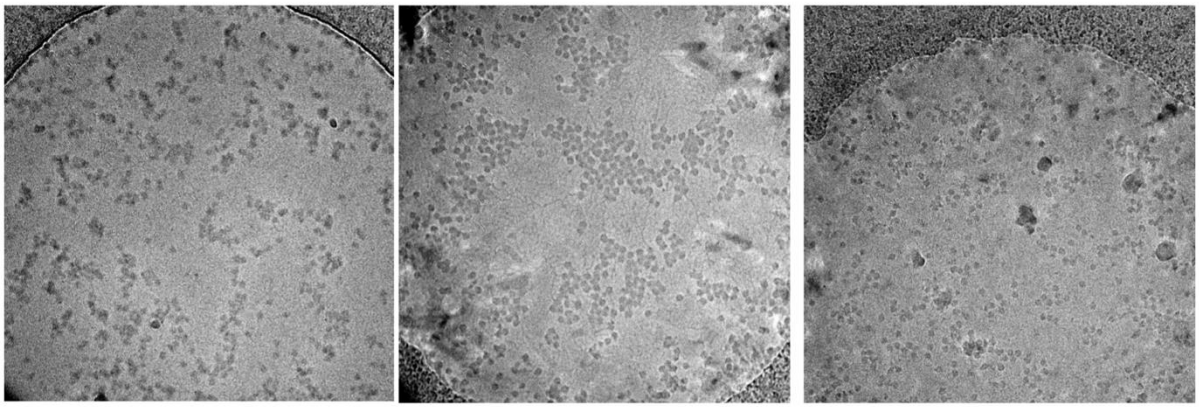


***Figure 2.12.** Effects of pH on aggregation of ATPase on cryoEM grids.*

2.4.2.3.Detergent artefacts

Protein samples after gel filtration are rather diluted and need to be concentrated several fold before they can be used for grid preparation. Transhydrogenase in particular, for which the yield was not high, had to be concentrated 20-30 fold after gel filtration to reach ~5 mg/mL suitable for grid preparation. For samples without detergent this is not a problem (in amphipols or nanodiscs) but concentrating of samples with detergent also concentrated detergent, because detergents exist in a form of large molecular weight micelles in aqueous solutions.

The optimal concentration of total detergent for cryo-EM grids is probably around 0.2-0.4%. Anything above this results in artefacts: either free micelles visible or long threads (Figure 2.13). For some detergents, such as LMNG, artefacts were apparent even at lower concentrations, above 0.1%. Too low concentration of detergents, or more precisely, too low concentration of free detergent (not in micelles), resulted in denaturation of protein particles at the air-water interface.

For samples purified in LMNG, the final concentration of detergent was controlled by gel filtration at 2xCMC = 0.002%, so even several-fold concentration of samples did not result in [LMNG] > 0.1%. For NNT purified in FOM, however, excess detergent was removed using Detergent removal columns (ThermoFisher Ltd.) that bind detergents and other hydrophobic contaminants. Typically, 750 μL NNT was concentrated to 20uL ([FOM] ~ 2%), and after a single passage through a detergent removal column according to the manufacturer's instructions, 0.075% remaining FOM was measured. Detergent removal by dilution and re-concentration or by using spin desalting Zeba columns did not remove significant amounts of detergent.

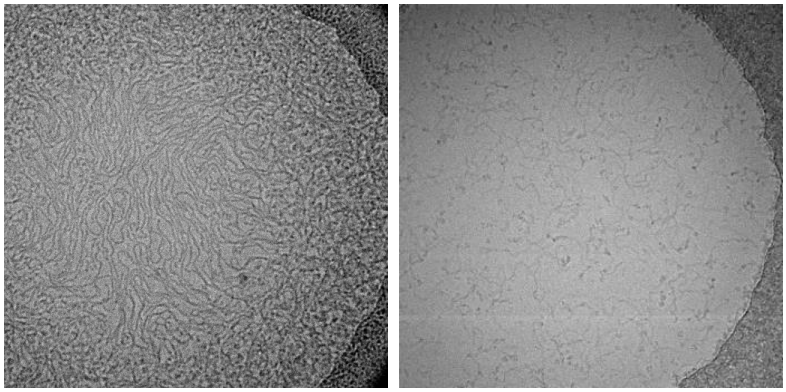


Figure 2.13. Detergent artefacts observed at high detergent concentrations.

2.4.2.4.Breaking of particles

Particles were sometimes observed to break or disintegrate into smaller pieces. It has been described that this is due to the denaturing effects of the air water interface on proteins (Glaeser *et al.*, 2016). This can be avoided by using a high-CMC secondary detergent, as described above, which binds at the air-water interface and prevents protein adsorption and denaturation there. This was observed to be particularly useful for samples without primary surfactant (nanodiscs) or with primary surfactants with very low CMC values (amphipols, LMNG) (Figure 2.14 and 2.15). Addition of secondary detergent also seemed to make the ice thickness less variable across the grids. On the negative side, as described below, higher concentrations of detergent also require higher concentrations of protein to achieve the same particle coverage.

2.4.2.5.Thin ice

The most important factors to tune for achieving thin ice is blotting time. This varied between 5 and 10 s for most of the samples. Blotting times up 20-30 s also worked sometimes, especially for samples with high concentrations of detergent.

Samples without or with little detergent were more prone to "over-blotting", as were the samples set up on gold grids. A persistent problem of many preparations was the partial overblotting, where the centres of the holes became too thin and excluded the particles from that region. This was partly avoided by increasing protein concentration and concentration of the secondary detergent. Furthermore, I observed that a consistently thin ice was achieved most

easily using Quantifoil 0.6/1 grids, which were less prone to overblotting than Quantifoil 1.2/1.3 grids with larger hole size.

Ice thickness was qualitatively assessed based on the contrast between the particles and the background in the acquired images. Another semi-quantitative method for estimating ice thickness is measuring electron fluence over an empty hole and a full hole, the ratio of which is approximately related to the ice thickness according to $d/\lambda = \ln (I_i / I_o)$, where d is the mean electron path in vitreous ice at 120kV (~600 Å), λ is the ice thickness, I_i is the incident fluence and I_o is the fluence after the sample (Feja and Aebi, 1999). Another semi-quantitative method for measuring ice thickness is the channel-method from tomography, which involves burning a narrow channel into the vitreous ice when the grid is tilted at 45º (Feja and Aebi, 1999). When the sample is not tilted anymore, the edges of the channel reveal the thickness of the sample. The method was useful for ice thicknesses above 50 nm but became increasingly imprecise below that as the vitreous ice around the channel became less stable. Ultimately, each cryo-EM grid displayed a range of ice thicknesses at different grid squares and also within the same grid square and it was important to choose the suitable holes during the collection itself.

2.4.2.6.Achieving adequate particle coverage

Ideal outcome of cryo-EM grid preparation is a single layer of tightly packed particles embedded in thin ice (Figure 2.14 and 2.15). Packing can be as tight as possible, to the point where proteins seemingly touch, as long as they are not overlapping. The resulting dataset will thus contain the highest number of particles per image. This required a high concentration of protein, reaching 5 mg/mL in most conditions and seemed to be dependent on the concentration of detergent present in the sample: the more detergent, the higher the protein concentration required to reach the desire coverage. As concentrating the protein to such high concentrations may not always be achievable due to propensity for aggregation or to avoid concomitant concentration of detergent, we investigated the use of graphene-oxide covered grids. As protein molecules get adsorbed on the solid graphene oxide support, an order of magnitude lower concentrations are required to achieve the same coverage. This comes at the expense of lower signal-to-noise ratio as graphene oxide contributes to the background. This has been reported to be quite minimal compared to amorphous carbon covered grids and should not interfere with image analysis of large (>0.5 MDa) protein complexes. Presence of support could even help with initial CTF parameters estimation and should limit beam-induced movement (Russo and Passmore, 2016).

Graphene oxide-covered grids were prepared as described previously. Graphene oxide solution (3 μL, 0.5 mg/mL) was applied to freshly glow discharged Quantifoil grids, incubated for 1 min and then washed with 20 μL water, twice from the front and once from the backside of the grid and air-dried (Bokori-Brown *et al.*, 2016). Grids prepared in this way were however, not covered very evenly and because of the fear that multiple layers of graphene oxide over some squared would detrimentally affect image quality, efforts were focused on optimising grid preparation in ice. Nevertheless, complex I from a small dataset on graphene-oxide-covered grids reached 4 Å. A further complication from supported grids is preferential orientation of particles because molecules tend to adsorb to the film with the largest or the most charged surface. Additionally, overblotting such grids results in complete dehydration of protein – freeze drying – precluding high resolution data reconstruction.

2.4.3. Optimized grid preparation: Transhydrogenase

2.4.3.1.FOM preparation

For FOM-solubilised transhydrogenase, several ten-fold concentration of fractions after gel filtration was required, which increased FOM concentration above 1%. This was not compatible with preparation of grids for cryo-EM, for which optimal FOM concentration was determined to be between 0.2 and 0.4%. As described above excess FOM was removed using detergent removal columns (Thermo-Fisher Scientific Inc.) which typically removed ~90% of free detergent. It was important to measure detergent concentration of each preparation prepared this way and adjust FOM concentration back to 0.2-0.4%. Detergent removal columns did not decrease activity or cause aggregation of the sample. Zeba columns, dilution-reconcentration and microdialysis on the other hand did not remove more than ~25% of the detergent.

Ultimately it was discovered that concentrating NNT in FOM above ~2.5-3 mg/mL results in aggregation – seen on cryoEM grids and UV-Vis spectra ($A_{260/280} > 1$) (Figure 2.14). This aggregation preceded detergent removal and seemed to happen very fast. At 2 mg/mL before the protein started to aggregate, a small dataset was collected which showed 2D classes, but contained too few particles to allow 3D reconstruction and the FOM preparation had to be abandoned.

2.4.3.2.Nanodiscs

Cryo-EM grids were prepared from NNT in MSP1E3D1 nanodiscs (large nanodisc) and MSP1D1ΔH5 nanodisc (small nanodisc). Due to the lack of detergent in these samples, addition of secondary detergent was necessary to prevent aggregation of particles and denaturation at the air-water interface (Figure 2.14). Both samples were prepared in the presence of 5 mM $NADP^+$. NNT- MSP1E3D1 sample included 0.05% CHAPS and NNT-MSP1D1ΔH5 nanodisc included 0.1% OG and was prepared using Ultrafoil Gold 0.6/1 grids. Despite excellent cryo-EM grids prepared with nanodisc sample, this preparation was also abandoned due to the fear that the large nanodisc density around the membrane domain is interfering with cryo-EM image alignment.

2.4.3.3.LMNG preparation

Ultimately, NNT in LMNG was used directly for grid preparation. It was critical to perform gel filtration in minimal (0.002%) LMNG and not to concentrate fractions more than 30-fold to avoid too high LMNG. Care was also taken to prepare a fresh LMNG stock on the day of purification. 0.05% CHAPS was added to the protein sample before grid preparation to improve the ice quality and the particle distribution (Figure 2.14). For samples with substrates added, these were added at 5 mM concentration to the protein sample 15-30 min before the grid preparation and the samples were incubated on ice.

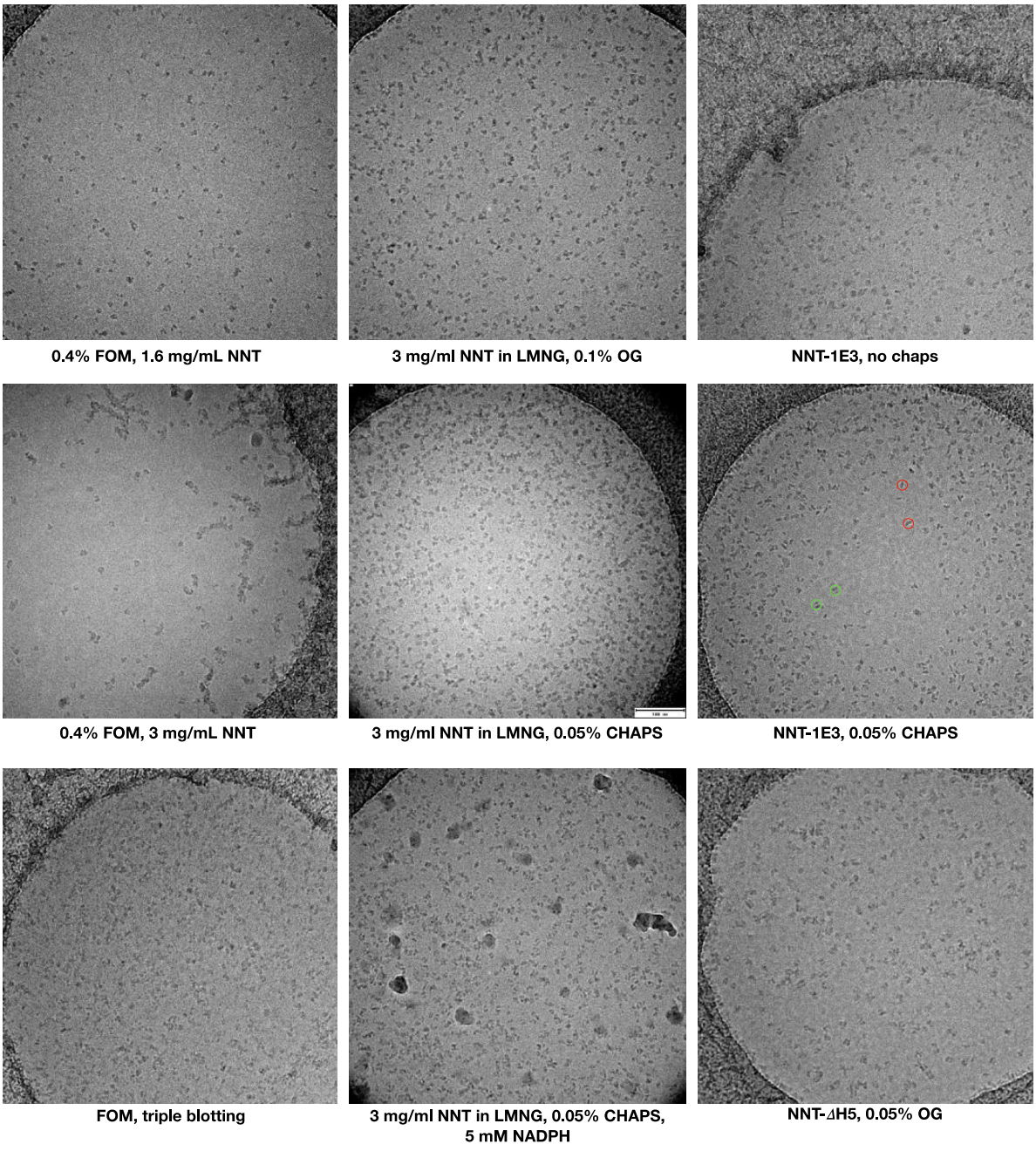


Figure 2.14. Summary of NNT grid optimisation.

2.4.4. Optimised grid preparation: complex I

As described in chapter 4, high resolution datasets for complex I were also collected on a sample purified in LMNG. Nevertheless, cryo-EM grids of complex I were prepared also with complex I reconstituted into amphipols and nanodiscs or dissolved in FOM, digitonin or Cymal-7 detergents and all of these preparations could in principle be used for data collection (Figure 2.15). Grid preparation procedure for CxI-LMNG was ultimately relatively similar to NNT. After gel filtration in 0.002% LMNG-containing buffer, complex I containing fractions were concentrated to ~3 mg/mL. 0.2% CHAPS was added to the protein sample before grid preparation to improve the ice quality and the particle distribution (Figure 2.15). The native and the deactive complex I samples were frozen without further substrate additions. For NADH-only dataset NADH at 5 mM was added to the protein 20 min before the grid preparation and the samples were incubated on ice. For NADH and rotenone dataset, NADH (5 mM) was added to complex I sample, followed by 300 μM rotenone (dissolved in DMSO) and 20 min incubation on ice. For NADH&DQ turnover dataset, a mixture of substrates and CHAPS was prepared at 2x working concentration (2 mM DQ, 2 mM NADH and 0.4% CHAPS) which was mixed in 1:1 ratio with 6 mg/mL complex I at room temperature immediately before applying on grid, blotting and freezing. With this simple procedure, complex I was active for 20 s before being frozen and less than 10% of substrates would have been consumed in this time.

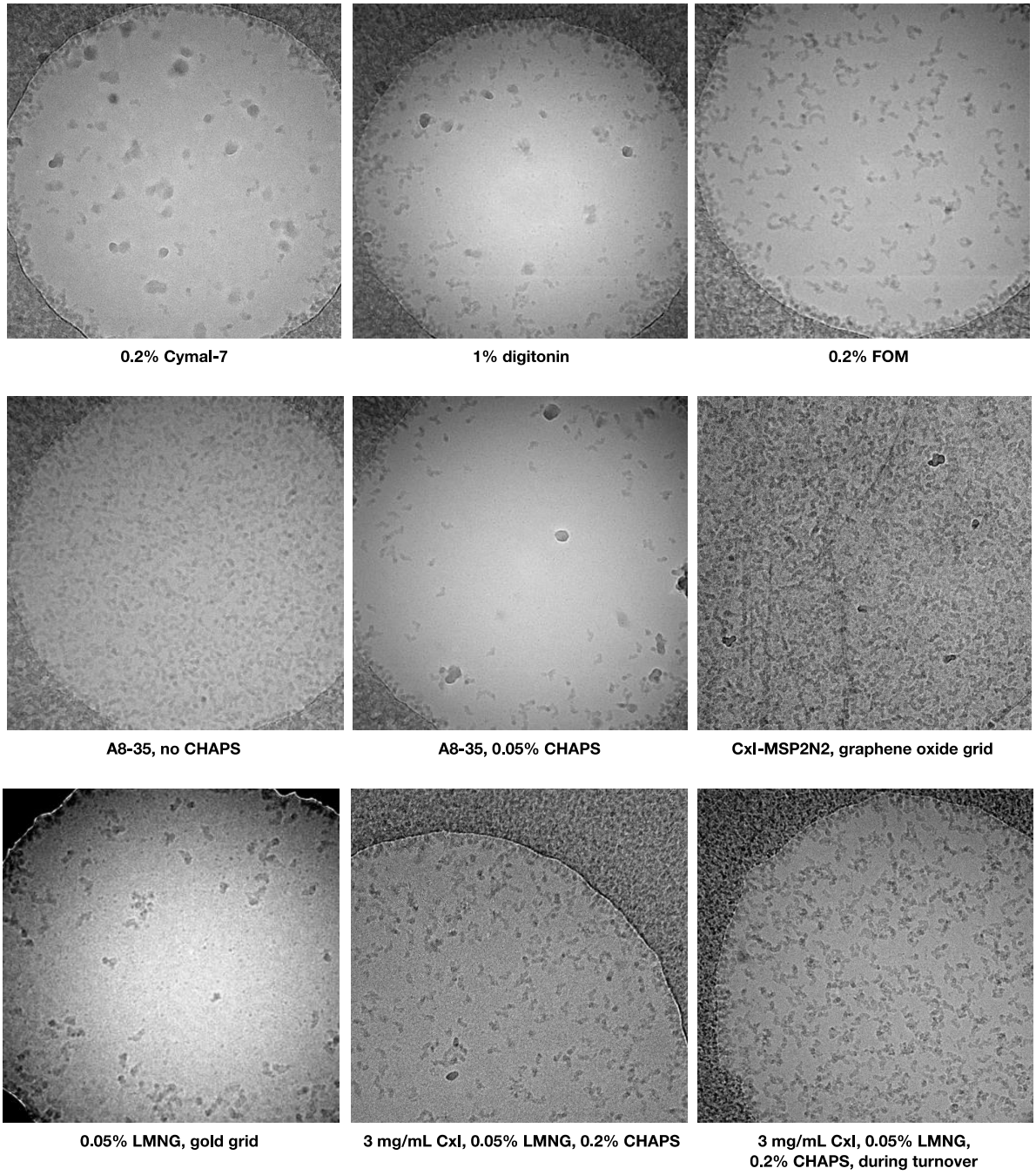


Figure 2.15. Summary of complex I grid optimisation.

2.4.5. Initial grid screening

Initial images for screening purposes were collected on a Tecnai T12 electron microscope (FEI) operating at 120 kV in a low-dose mode. Images were recorded using either a CCD camera or a TemCam-XF416 CMOS camera (TVIPS, Gauting, Germany). Grids were mounted on a pre-cooled single tilt side-entry cryo-transfer holder (Fischione Instruments, Inc., Export, PA, USA). Images were recorded at a nominal magnification of 42000x. Defocus was set for -3.5 μm, electron dose 20-50 $e^-/s/Å^2$ and exposure time 1-3 s.

2.4.6. Data collection: Transhydrogenase

NNT-$NADP^+$ grids were imaged using a 300 kV Titan Krios electron microscope equipped with a Gatan energy filtered K2 summit camera with a slit width of 20 eV at The Astbury Centre for Structural Molecular Biology of the University of Leeds.

NNT-NADPH grids were imaged using a 300 kV Titan Krios electron microscope equipped with a Gatan energy filtered K2 summit camera with a slit width of 20 eV at the CM01 beamline, ESRF, Grenoble.

Images were collected with EPU in a K2 super-resolution mode with a nominal magnification of 130000x and a physical pixel size of 1.065 Å for NNT-$NADP^+$ and 1.07 Å for NNT-NADPH dataset. Total dose of 72 $e/Å^2$ (NNT-$NADP^+$) or 70 $e/Å^2$ (NNT-NADPH) was fractionated into 40 frames of 250 ms each.

Apo-NNT grids were imaged using a 300 kV Titan Krios electron microscope equipped with a Gatan energy filtered K3 camera with a slit width of 20 eV at the Institute of Science and Technology Austria. Physical pixel size was 0.84 Å and the total dose of 90 $e/Å^2$ was fractionated into 88 frames of 36 ms each.

NNT samples in nanodiscs were imaged using a 300 kV Polara electron microscope equipped with a K2 camera at the Vienna Biocenter Electron Microscopy Facility. Images were collected with SerialEM in a super-resolution mode with a nominal magnification of 39000x and a physical pixel size of 0.98 Å. Total dose of 80 $e/Å^2$ was fractionated into 40 frames of 250 ms each. NNT-1E3D1 samples were collected from Quantifoil Carbon grids and NNT-ΔH5 from Quantifoil Ultrafoil gold grids.

2.4.7. Data collection: complex I

Native, NADH, rotenone and turnover grids were imaged using a 300 kV Titan Krios electron microscope equipped with a Falcon III camera operating in an integrating mode at

Cryo-Electron Microscopy and Tomography Core Facility in CEITEC, Brno. Turnover, NADH and rotenone datasets were imaged at a nominal magnification of 130000x with a physical pixel size of 1.061 Å. Total dose of 89 e/Å^2 was fractionated into 40 frames of 21 ms each. Native dataset was imaged at a nominal magnification of 165000x with a physical pixel size 0.83 Å. Total dose of 98 e/Å^2 was fractionated into 40 frames of 25 ms each. Images were collected from 3x3 patches of holes using image shift – resulting beam tilt error was small and refined during processing and this mode of acquisition allowed for a high speed of collection.

Deactive complex I was imaged using a 300 kV Titan Krios electron microscope equipped with a K3 camera operating in a super-resolution mode at the Institute of Science and Technology Austria. Grids were imaged at a nominal magnification of 81000x and physical pixel size of 1.10 Å. Total dose of 78.5 e/Å^2 was fractionated into 50 frames of 88 ms each. The deactive dataset was collected using SerialEM with an active beam tilt compensation, hence beam tilt error was not refined separately for each of the nine holes imaged using the image shift method.

Complex I samples in nanodiscs were collected using a 300 kV Polara electron microscope equipped with a K2 camera at the Vienna Biocenter Electron Microscopy Facility. Images were collected with SerialEM in a super-resolution mode with a nominal magnification of 39000x and a physical pixel size of 0.98 Å. Total dose of 80 e/Å^2 was fractionated into 40 frames of 250 ms each.

2.5.Image processing: Transhydrogenase

We collected 2272 movies for the NADP$^+$ dataset, 1722 movies for NADPH dataset and 786 movies for the apo-NNT dataset. The processing was done in RELION 2.1 (Scheres, 2012b) unless otherwise stated. Movie frames were aligned using MotionCor2 (Zheng *et al.*, 2017) and initial CTF parameters were estimated from averaged images using CTFFIND (Rohou and Grigorieff, 2015). Autopicking with 2D class averages as references resulted in 516395 particles for the NADPH dataset, 1076677 particles for the NADP$^+$ dataset and 500001 particles for the apo dataset. Multiple rounds of 2D classification and 3D classification were performed to classify considerably heterogeneous particles. CTF parameters and per-particle trajectories in the best two classes from the NADP$^+$ dataset were estimated and corrected (particle polishing) using RELION 3.0 (Zivanov *et al.*, 2018). Local resolution was estimated using Resmap 5.0 (Kucukelbir, Sigworth and Tagare, 2014).

For the NADP$^+$ dataset we performed one round of 2D classification followed by two rounds of 3D classification with 3x and 2x binned particles, respectively, which resulted in

eight classes containing 342284 particles from which it was obvious that there is a considerable heterogeneity in the conformations of dI_2 and dIIIb. To address this, the particles were re-extracted without binning and refined with a dII_2dIII_2 mask to 3.5 Å and then classified into ten classes without searches and without a mask to separate the particles based on the relative position of the dI_2. By combining similar classes and auto-refining them, we separated six different states of NNT which differ from each other with respect to the positions of dI_2 and dIIIb. Two of these classes were resolved to 13 Å, further two to 7 Å and the best two classes to 3.7 and 3.2 Å. From the latter two maps it was obvious that the density for dI_2 was still weaker than for the rest of the particle, indicating remaining heterogeneity in the positions of dI_2, hence the final maps for these two classes were produced by combining focus-refined maps of dII_2dIII_2 and dI_2 parts. This gave the final map for class 1 (closed) with 2.9 Å resolution in the dII_2dIII_2 and 3.2 Å resolution in the dI_2 and for class 2 (open) 3.5 Å resolution for the dII_2dIII and 4.1 Å resolution for the dI_2. Due to the poor quality of the map for the open class in some regions of dIIb (CL6,7,8), we hypothesised that there is still remaining conformational heterogeneity present in this class and we performed another round of classification of these particles with a mask around dIIb to sort out any particles with dIIIb attached/semi-attached to dII. This procedure removed 52% particles and we performed an auto-refinement with a dII_2dIII mask of the remaining 28498 particles with dIIIb unambiguously detached. This resulted in a 3.7 Å map with dIIb density much better resolved and clearly different from the density of dIIa (which has dIII attached) and was used for model building.

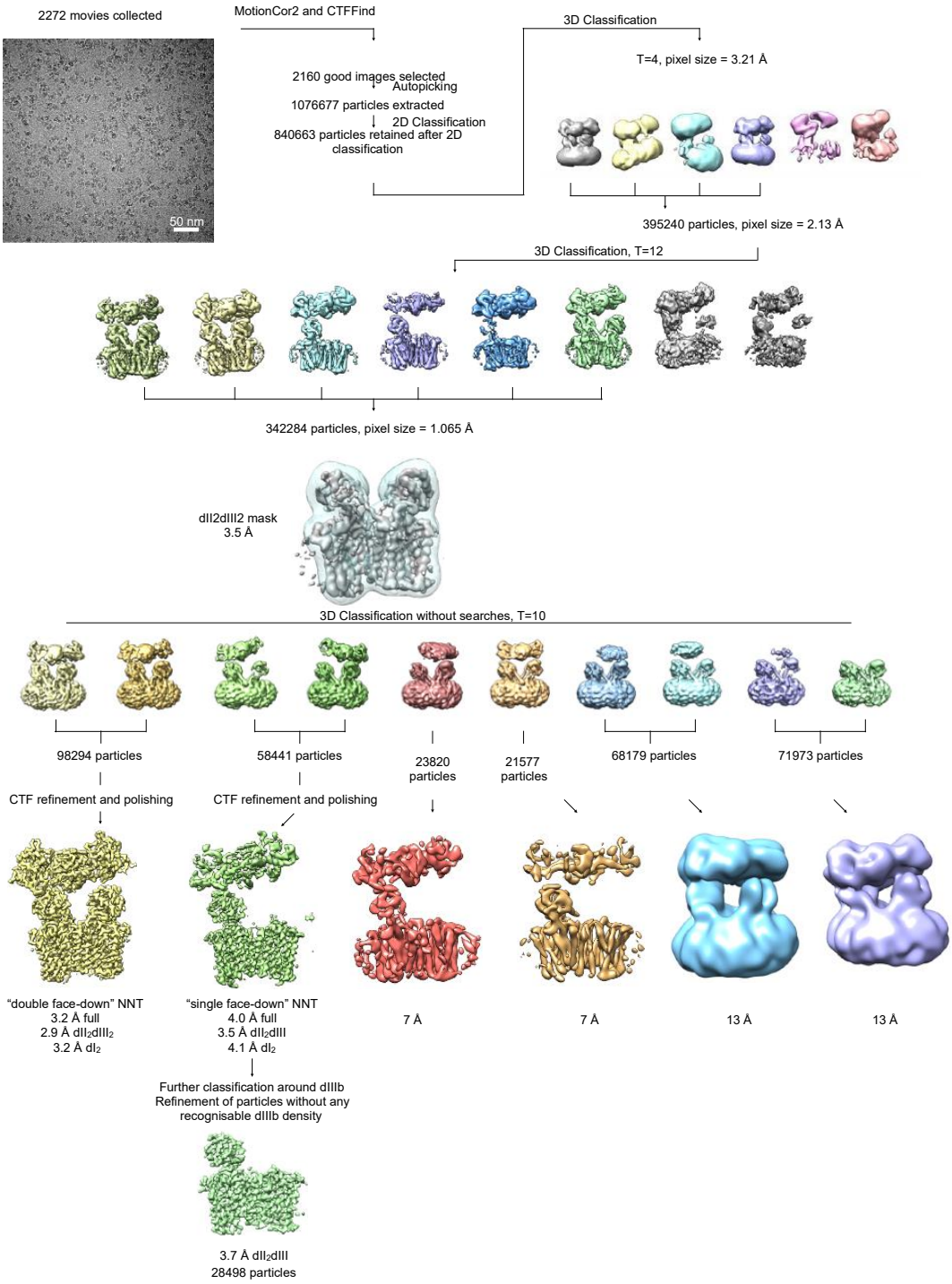


Figure 2.16. Processing of NNT-NADP+ dataset.

Thorough classification of particles resulted in six distinct classes with different domain orientations and resolutions. Best-resolution class ("double face-down" NNT) is almost symmetric with both dIIIs bound in the 'face-down' position while the other five have only one dIII bound in the face-down position and the other dIII detached and the dI2 tilted at different angles. Simultaneous opening of dI2 and dissociation of dIIIb is probably necessary to permit full dIIIb rotation during the catalytic cycle. Monomers with dIII detached from dII also show weak or no density for dI-dII linker which suggests that it detaches from dIII to allow dI2 to open.
Adapted from Kampjut & Sazanov, 2019.

For the NADPH dataset, we performed one round of 2D classification followed by one round of 3D classification on 510073 3x binned particles with the same parameters as in the $NADP^+$ dataset, which failed to yield any class with visible secondary structural features. We performed extensive classifications of these particles into 4-20 classes with T values between 4 and 12 and different filtering of the reference volumes which all failed to give a single well-defined class with secondary structural features. Nevertheless, we selected 405141 particles from the best 4 classes from a classification with T=8 and k=6 (same as for $NADP^+$ dataset) for further classification attempts. Another round of extensive classifications was performed on this particle stack with the best results obtained from a T=8, k=20 run. Even in this run only ~6 classes had well-defined features corresponding to the domains of NNT. These 6 classes were re-extracted and refined which gave structures at 8.3, 9.2, 13, 16, 16, and 16 Å and revealed a high degree of dI_2 and dIII mobility (Extended Data Fig. 4a). To address this, we tried to refine the 405141 particles together into one consensus structure at 13 Å, followed by focused refinement of either dI_2 or dII_2 but neither of these strategies yielded structures better than 9 Å, probably due to the too large conformational heterogeneity within the consensus refinement. Furthermore, dI_2 and dII_2 are only 80 kDa each and dII_2 is obscured by a micelle which further complicates the alignment of such a heterogeneous dataset. Finally, we performed multi-body refinement in RELION 3.0 starting from the consensus refinement of 405141 particles and using masks for the two bodies: dI_2 and dII_2. This yielded a 9 Å reconstruction of dI_2 and 10 Å reconstruction of dII_2 but more importantly, the principal component analysis showed the striking degree of dI_2 mobility in these particles (Fig. 2de). This is not unexpected as upon dIII dissociation, dI_2 remains only loosely tethered to the rest of the particle by a flexible dI-dII linker.

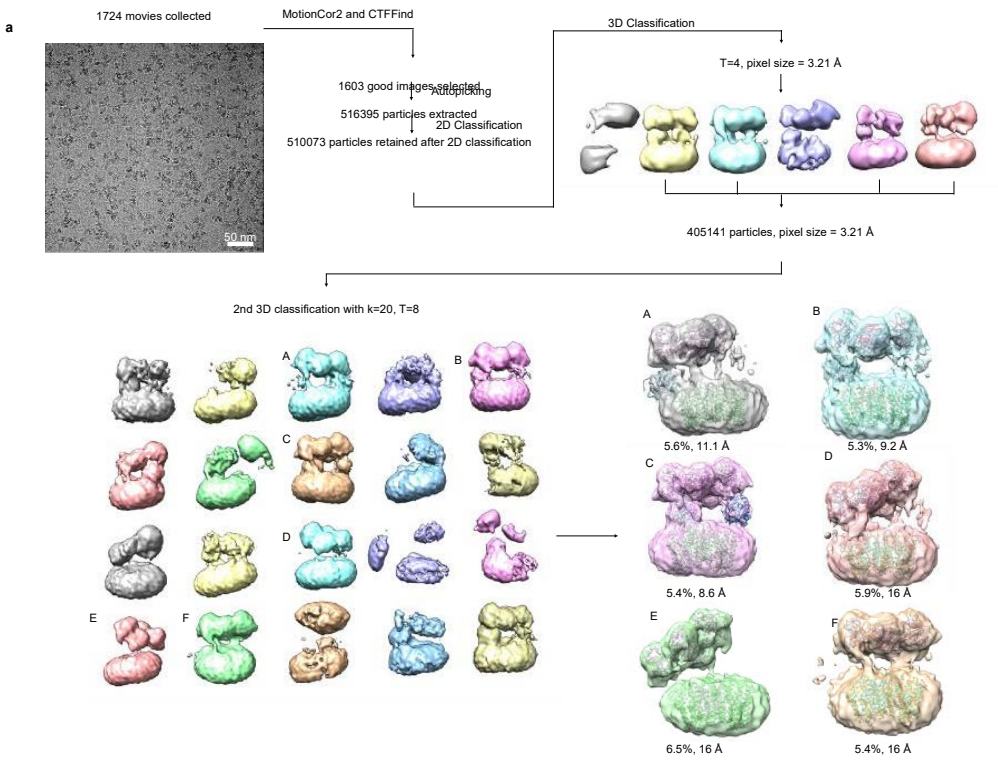


Figure 2.17. Processing of NNT-NADPH dataset.

NNT bound to NADPH exhibited a large degree of conformational flexibility, which prevented a high- resolution (beyond 8 Å) refinement of any class of particles. Classification into 20 classes revealed that only a small proportion (~5%) of particles have one dIII bound to dII in the face-down orientation (class C). Another ~10% have partially detached dIII (classes A and B). A vast majority, however, have both dIIIs dissociated from the other domains leading to a large degree of freedom of movement of dI2 and both dIIIs independently of each other (see classes D, E and F).
Adapted from Kampjut & Sazanov, 2019.

The apo-NNT dataset was fully processed in RELION 3.0. We manually selected 643 good images and extracted 3x binned particles. Only one conformation of NNT could be observed using various 3D classification schemes, hence the 283706 particles corresponding to it were extracted without binning and auto-refined to 4.5 Å. From this structure it was clear that dI_2 exhibits the same heterogeneity as in the $NADP^+$ dataset. To improve the resolution in the dII-dIII, we focus-refined around this region and classified the particles without searches into 6 classes. The best class with 67908 particles was auto-refined and polished to reach the resolution of 3.7 Å. Further polishing or subclassification of this class did not improve the resolution further. Attempts at focus-refinement of dI_2 did not improve its resolution compared to the polished full structure of the final selected particles in which local resolution in dI_2 was ~5 Å. We conclude that dI_2 without bound nucleotide exhibits even more disorder than when NAD^+ is bound, possibly allowing NAD(H) exchange and dI_2 binding/release.

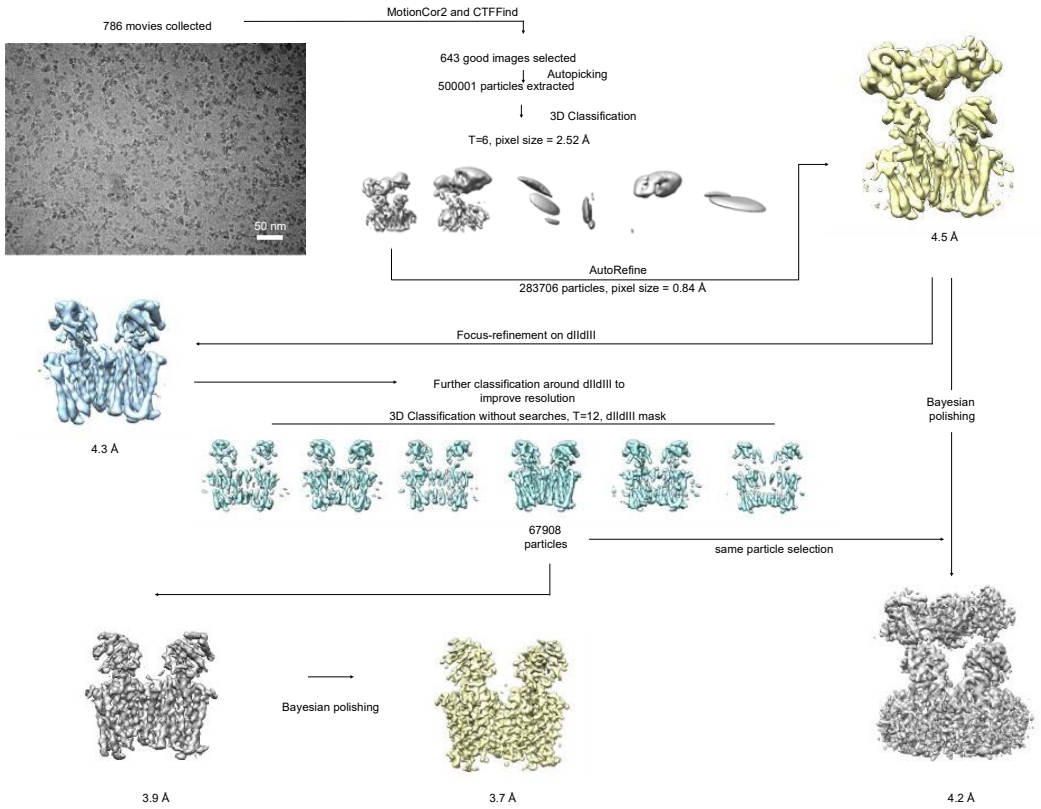


Figure 2.18. Processing of Apo-NNT dataset.

Apo-NNT exhibited a single conformational class similar to the "double face-down" NNT class in the presence of NADP+.
Adapted from Kampjut & Sazanov, 2019.

NNT in MSP1E3D1 nanodisc sample (NNT-1E3D1) was processed in RELION 2.0. Collection yielded 2153 micrographs, of which 1664 displayed Thon rings beyond 6 Å and were used for further processing. Most micrographs displayed a signal cut-off in one direction in Fourier space, which later became apparent was a defect of the microscope. 230466 particles were autopicked. After several rounds of classification, two distinct classes were separated, class 1 contained 40670 particles (10.3 Å) and class 2 contained 34247 particles.

NNT in MSP1D1ΔH5 nanodiscs (NNT-ΔH5) was processed in RELION 2.0. 879 micrographs were collected, but only 272 gave Thon rings beyond 6 Å because of the microscope defect. 138515 particles picked and classified. 3D classification resulted in one best class containing 38912 particles (12 Å) similar to Class 1 from 1E3D1 nanodiscs, apart from a noticeably smaller nanodisc density. No minor (class2-like) conformation could be classified from this dataset.

Particles from the best classes of nanodisc datasets were combined and several focus refinements tried: whole, without nanodisc, dI2dII2dIII, dI2dIIdIII, dI2dIII, dI2, dI2dIII2,

dII2dIII. None of this worked, either because of technical problems with Polara at VBC or nanodisc density interference. In fact, only a few examples of small proteins in nanodiscs have so far been solved. Finally, conformational heterogeneity combined with nanodisc density could have been the problem. Refinements and classification were also tried in cisTEM software (Grant, Rohou and Grigorieff, 2018), but the results were almost the same as from RELION.

2.6. Image processing: Complex I

We started with 2284 movies for the native dataset, 1443 movies for NADH dataset, 4239 movies for the turnover dataset, 2323 movies for the rotenone dataset and 2065 movies for the deactive dataset. The micrographs were included in the further processing steps if the Thon rings extended to 4 Å (turnover dataset) or 5 Å (all other datasets) according to CTFFIND maximum resolution estimate. Processing was done in RELION 3.0 and 3.1 (Zivanov *et al.*, 2018) unless otherwise stated. Movie frames were aligned using MotionCor2 (Zheng *et al.*, 2017) and initial CTF parameters were estimated from averaged images using CTFFIND 4.1.13 (Rohou and Grigorieff, 2015). Autopicking with 2D class averages as references was done in RELION and resulted in 295630 particles for the native dataset, 364499 particles for NADH dataset, 938197 particles for the turnover dataset, 354690 particles for the rotenone dataset and 740026 particles for the deactive dataset.

For the sake of consistency processing steps were done as similarly as possible for all the datasets. Particles were extracted at 3x the physical pixel size and processed using one round of 2D classification and one round of 3D classification (Supplementary Figures 1-3). Good classes were then extracted at full pixel size and autorefined into a single “consensus” structure. CTF parameters and per-particle trajectories were then refined in an iterative manner (one round of CTF refinement followed by Bayesian polishing and another CTF refinement). This led to a consensus refined complex I structure of all particles, reaching nominally high resolutions (up to 2.5Å in turnover dataset) but with blurring around the edges and some weaker subunits (42 kDa and B14.7) suggesting a large degree of conformational heterogeneity and presence of partially disassembled particles. Particles were therefore 3D classified without searches to remove damaged particles and then focus-reverse-classified to sort out the heterogeneity as done earlier (Letts *et al.*, 2019). Briefly, all the intact particles were focus-refined using a peripheral arm mask and then classified without searches using a loose membrane mask to finely classify states with slightly different angles between the two arms. This procedure resulted in a single closed class and 3-4 open classes from each dataset, based

on which atomic models were built, analyzed and compared. In most datasets finer sub-classification of classes with different masks, in particular around the quinone binding cavity, was tried but did not result in finding additional conformational states of complex I. Further classifications of consensus refinements of MD or PA of all particles also did not result in finding additional conformational classes.

In the native dataset, the first round of 3D classification was done with k = 6 classes and regularization parameter T = 6, which resulted in 148307 good particles. Upon duplicates removal, consensus refinement and two rounds of CTF refinement and Bayesian polishing, 128964 particles were refined around the matrix arm and subjected to final classification without searches around the membrane arm with k = 6 and T = 20. This resulted in four good classes with 91979 particles in total and two partially broken classes. The good classes were refined and post-processed separately to give final maps for model building.

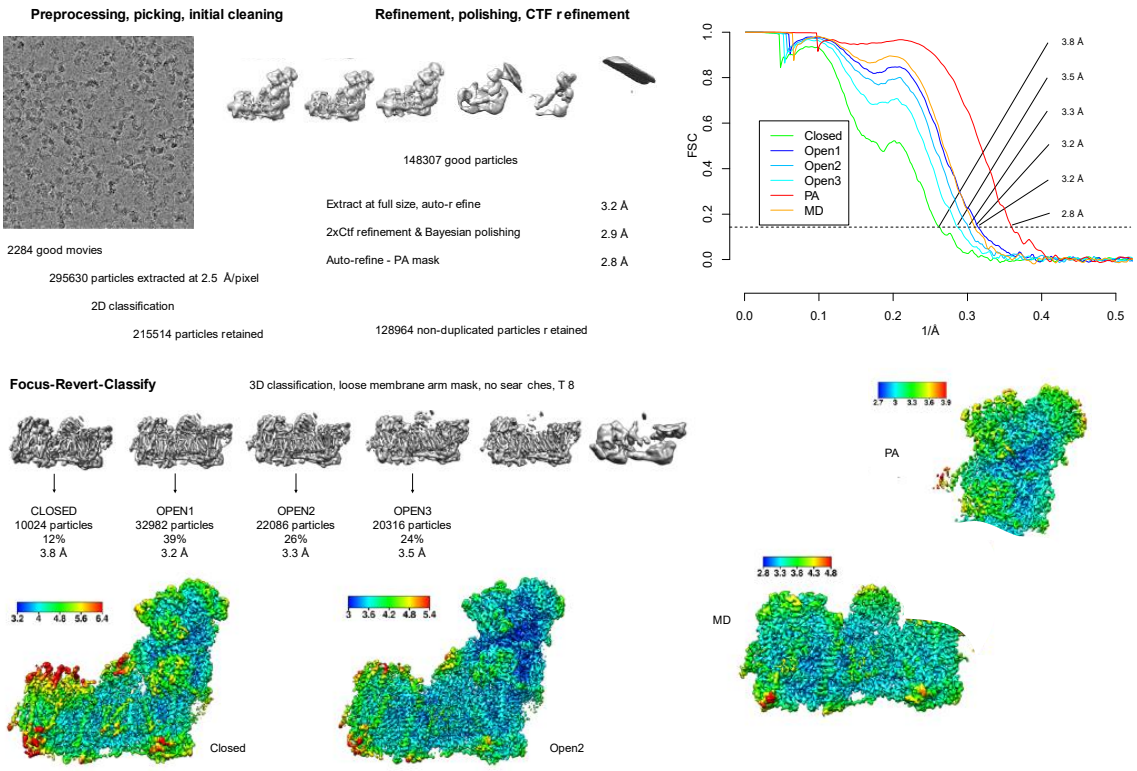


Figure 2.19. Processing pipeline of the native dataset.

FSC plots and local resolution of the closed and open2 classes are shown as well as local resolution plot of the focus-refined PA and MD.
Adapted from Kampjut & Sazanov, 2020.

In the NADH dataset, the first round of 3D classification was done with k = 6 and T=4, which resulted in 276169 good particles. After duplicates removal, consensus refinement, two rounds of CTF parameter refinement and Bayesian polishing, another 3D classification with k = 6 and T=8 without searches was performed to remove damaged particles. Remaining 182053 particles were refined around the PA and subjected to classification around the membrane arm with k = 6 and T = 8. This resulted in 4 good classes with a total of 140630 particles and two classes with bad density and partially disassembled complex. Particles from the four good classes were refined, post-processed and the maps were used for model building.

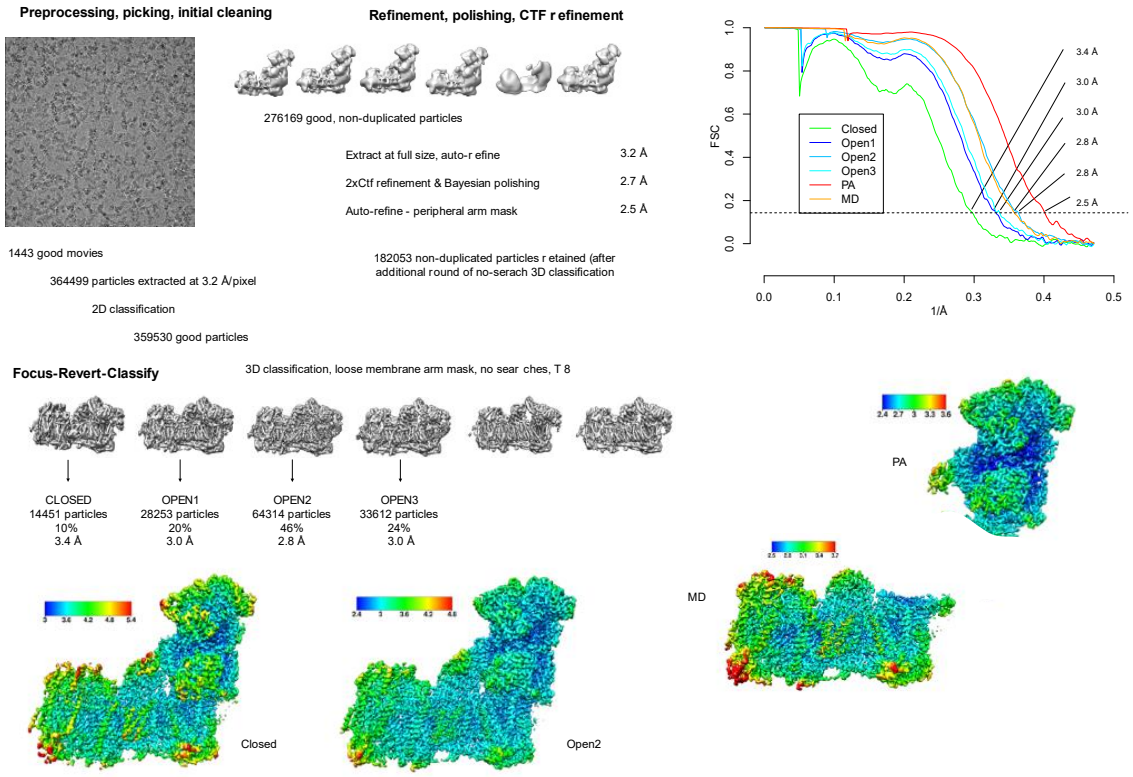


Figure 2.20. Processing pipeline of the NADH dataset.

FSC plots and local resolution of the closed and open2 classes are shown as well as local resolution plot of the focus-refined PA and MD.
Adapted from Kampjut & Sazanov, 2020.

In the deactive dataset, the first round of 3D classification was done with k = 6 and T = 4 and resulted in 458556 good particles. After a consensus refinement, another 3D classification step without searches and with k = 6 and T = 10 was performed to remove further damaged particles. This resulted in 272990 particles, which were refined around the PA and

classified around the membrane arm using 3D classification without searches and k = 4 and T = 10. This resulted in four classes that differed from each other by the degree of openness. They were subjected to two rounds of CTF parameter refinement, Bayesian polishing, refinement and post-processing to obtain the final density maps used for model building.

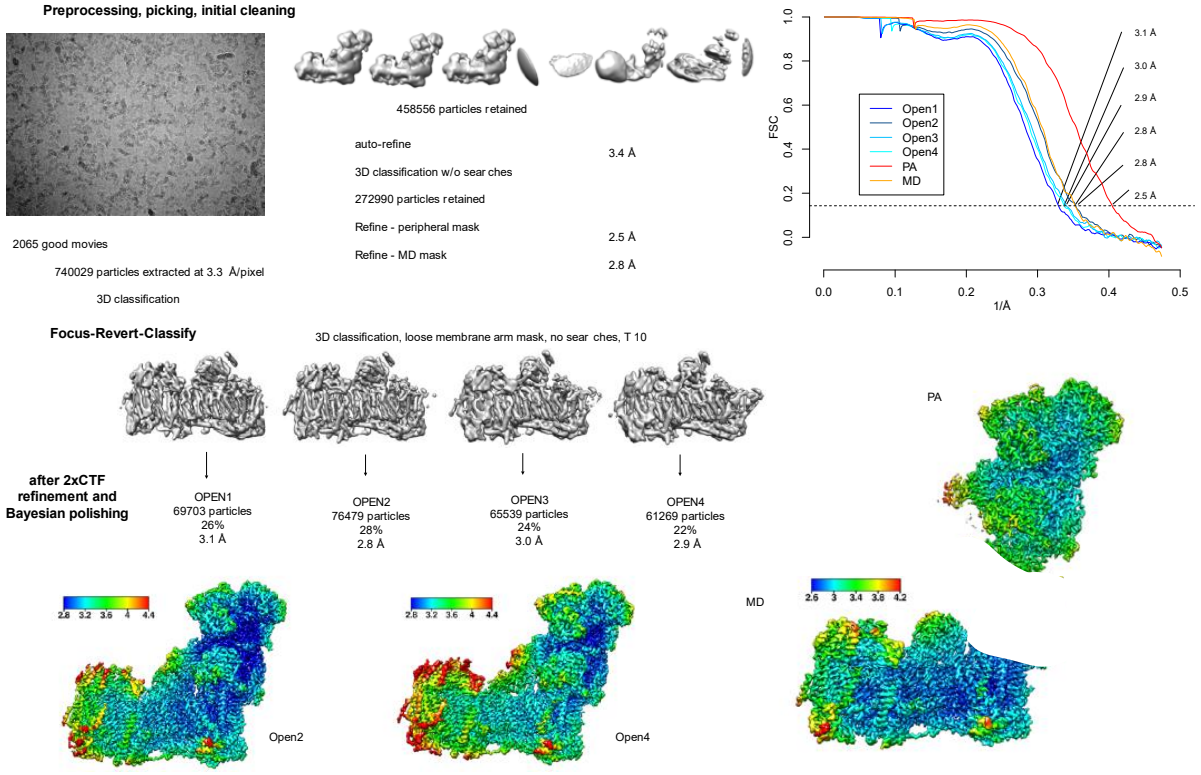


Figure 2.21. Processing pipeline for the deactive dataset.

FSC plots and local resolution of the closed and open2 classes are shown as well as local resolution plot of the focus-refined PA and MD.
Adapted from Kampjut & Sazanov, 2020.

In the rotenone dataset the first round of 3D classification was done with k = 6 and T = 4 and resulted in 175658 good particles after duplicate removal. After two rounds of CTF parameter refinement, Bayesian polishing and auto-refinement, another 3D classification without searches and with k = 6 and T = 8 was performed to remove bad particles, which resulted in 142565 good particles, which were focus-refined around the membrane domain. Because of the low abundance of the closed class in this dataset, the closed class was separated by 3D classification around the membrane arm with k = 6 and T = 10. The rest of the particles were 3D classified separately with k = 6 and T = 6 to better separate different open classes. The resulting particles were refined, post-processed and used for model building.

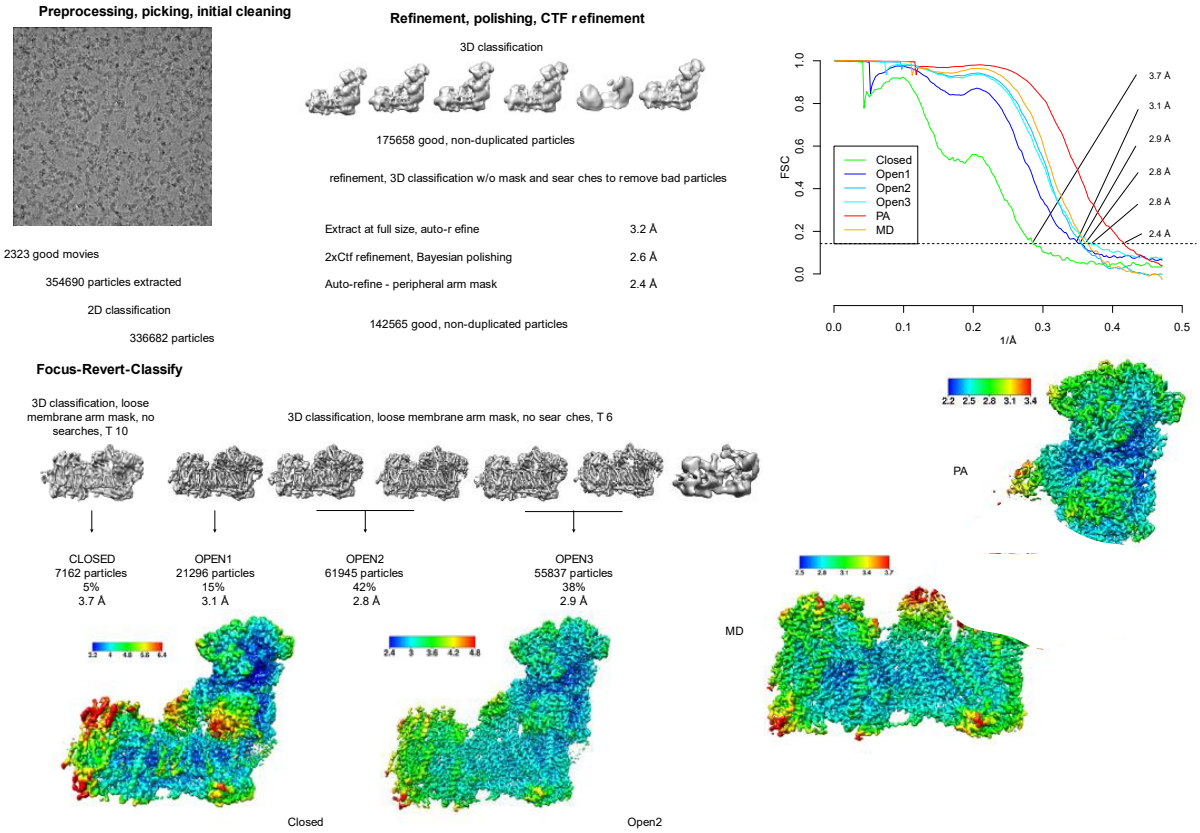


Figure 2.22. Processing pipeline for the rotenone dataset.

FSC plots and local resolution of the closed and open2 classes are shown as well as local resolution plot of the focus-refined PA and MD.
Adapted from Kampjut & Sazanov, 2020.

In the turnover dataset, the initial 3D classification was done with k = 6 and T = 4 and resulted in 518117 good particles after duplicated removal. After two rounds of CTF parameter refinement, Bayesian polishing and auto-refinement, another 3D classification without searches and k = 6 and T = 8 was performed to remove damaged particles, resulting in 344541 good particles retained. These were then 3D classified around the membrane arm without searches and with k = 6 and T = 8. Four of the resulting classes contained good particles with a total of 257669 particles, which were refined and post-processed for model building. The density within the quinone cavity of the closed class appeared broken, hence this class was further classified without searches and with k = 3 and T = 50. This resulted in one major class with about 55% particles that had better defined density in the quinone cavity and elsewhere and another minor class with about 30% of particles that had worse density overall but

otherwise no significant differences to the major class. Only the major closed class was used for model building.

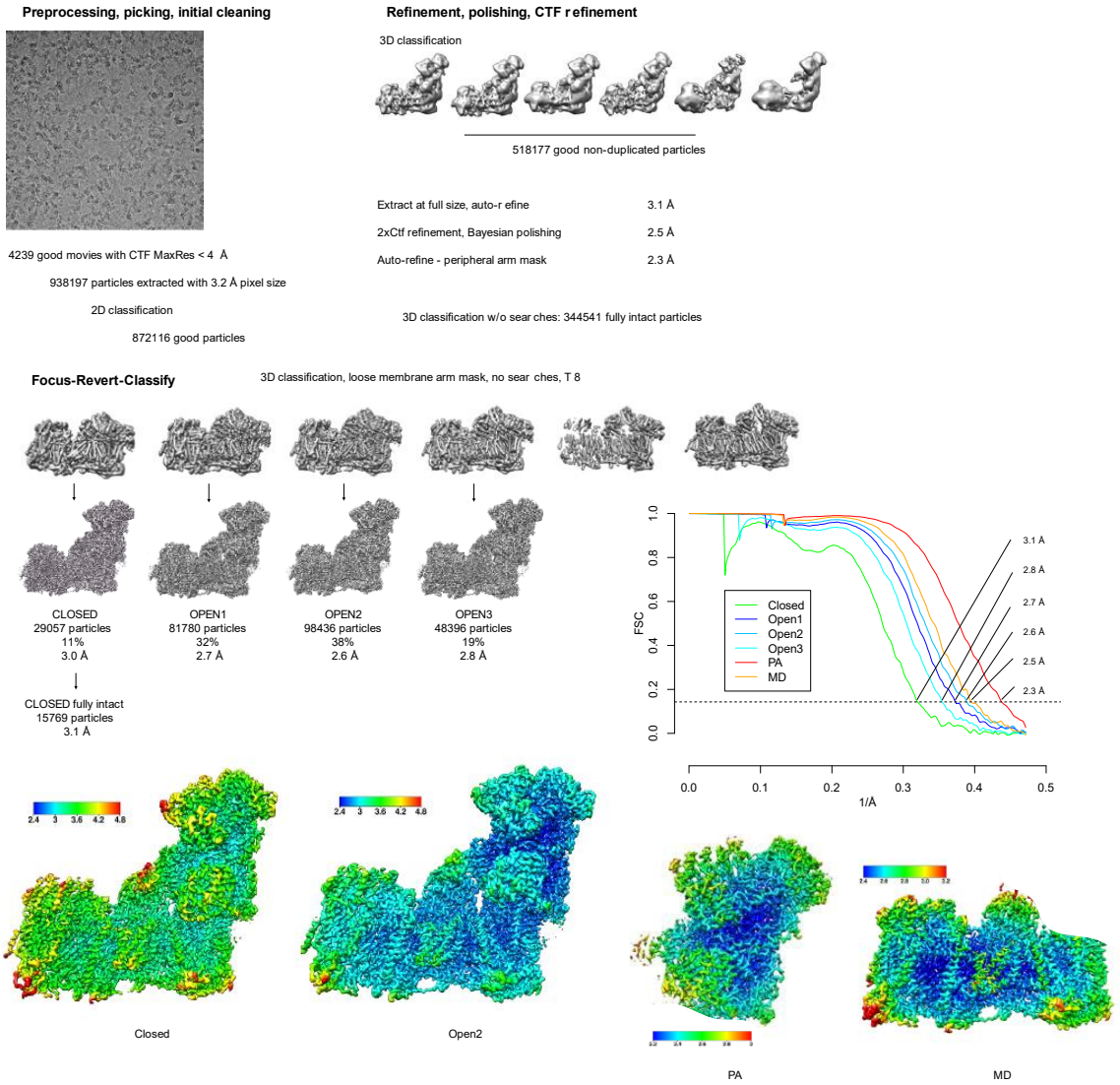


Figure 2.23 Processing of the turnover cryo-EM dataset.

Processing pipeline of the turnover dataset. FSC plots and local resolution of the closed and open2 classes are shown as well as local resolution plot of the focus-refined PA and MD.
Adapted from Kampjut & Sazanov, 2020.

In the complex I in nanodiscs dataset, 1080 images were collected and 629 images displayed Thon rings beyond 5 Å and were used in further processing. 189035 particles were picked in total, subjected to one round of 2D classification and the resulting 170091 particles were 3D classified into 6 classes (T = 4). One of the classes (54233 particles) displayed high resolution features and was refined to 4.5 Å (Figure 2.24a). This was further classified into 23532 particles in the open3 state without the B14.7 subunit (4.45 Å) and 30791 particles in the the open state with the B14.7 subunit (4.26 Å). The open state was focus refined around PA and MD separately. Counterintuitively, comparison of the angular distributions of this dataset which was collected on the graphene oxide support with the LMNG-datasets without support revealed better angular distribution with support (Figure 2.24b). The thickness of the nanodisc belt was barely larger than the thickness of the LMNG micelle (Figure 2.24c).

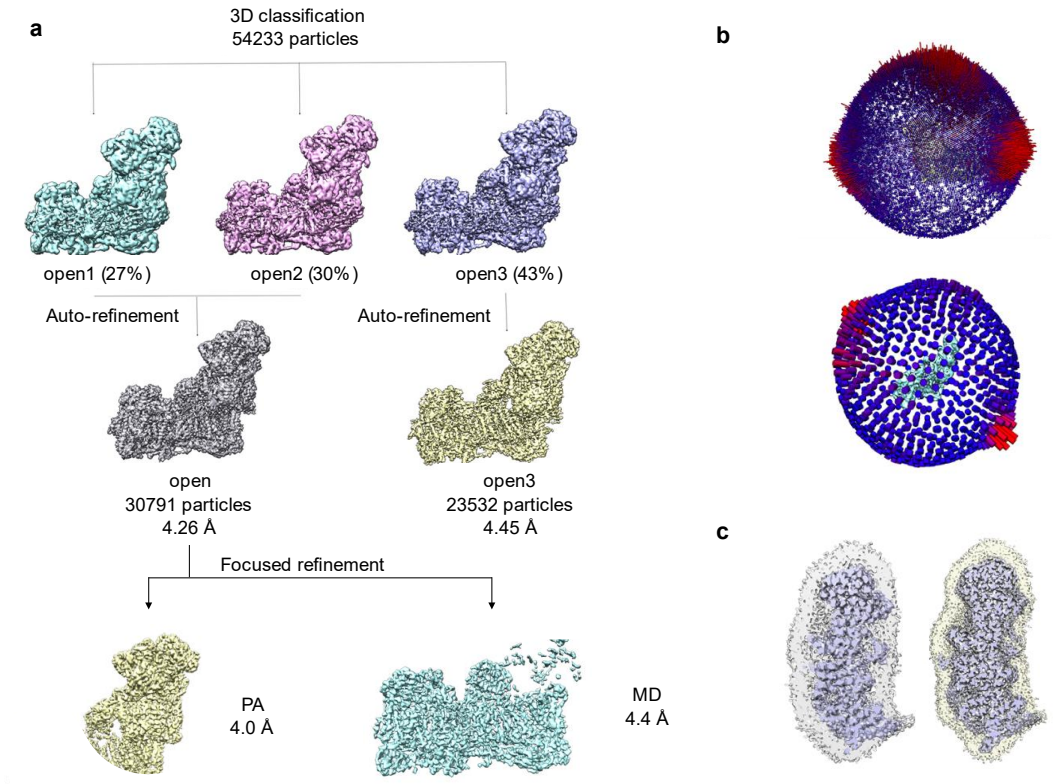


Figure 2.24 Processing of complex I in nanodiscs.

a, *3D classification reveals different conformations of complex I in the sample. Open and closed state were refined separately. There appears to be some additional flexibility in each state as focused refinement of the matrix and membrane arm of the complex reach slightly higher resolution.* ***b,*** *Comparison of angular distribution on graphene (top) and in ice (bottom). Angular distribution of particles is slightly improved on graphene support.* ***c,*** *Comparison of micelle thickness in graphene (left) and in LMNG (right). Density in nanodiscs is only slightly thicker and the overall structure of membrane arm of complex I is the same in both conditions.*

2.7. Model building and refinement: transhydrogenase

The initial model was built into the "double face-down" $NADP^+$ class density by starting with the bovine dIII crystal structure (PDB 1D4O) with substituted ovine sequence and the dI homology model that was generated using Phyre2 server (Kelley *et al.*, 2015). dII model and the linkers were built manually in *Coot* (Emsley *et al.*, 2010). Manually adjusted models were refined using *PHENIX* software (Adams *et al.*, 2010) using an adapted method for refining cryo-EM structures which uses two rounds of a single cycle of group ADP refinement followed by three cycles of global minimization. This allows optimization of 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). The final model for NNT is of very high quality in terms of geometry and fit to density (Appendix 1).

Densities for several lipid molecules were observed in each of the half-closed cavities formed by helices TM1, TM2 and TM6. They were modelled in the structure as phosphatidylcholines based on their density appearance and the fact that only phosphatidylcholine could reactivate lipid-depleted insect NNT (Vandock *et al.*, 2011).

A density corresponding to the adenylate part of NAD^+ was observed in the NAD(H) binding site of the dI. We propose that this density is NAD^+ which remained bound to the enzyme during purification or, more likely, was introduced into the sample as a contaminant from the $NADP^+$ solution. Our $NADP^+$ was only 98% pure and at 2% NAD^+ contamination, NAD^+ concentration would be 100 μM which is in the range of the K_d for the enzyme (Yamaguchi and Hatefi, 1993; Bizouarn *et al.*, 2005). The latter is supported by the fact that in our structure the conserved R139 is in the same extended conformation as in the *E. coli* dI_2 with bound NAD^+ and differs significantly from the apo or NADH-bound forms (Johansson *et al.*, 2005). Nevertheless, our structure is a bona fide oxidised ground state conformation of NNT and earlier biochemical work suggests that nucleotide exchange in dI subunit can happen during any stage of the catalytic cycle and does not influence the global conformation of the enzyme (Yamaguchi and Hatefi, 1989; Yamaguchi, Wakabayashi and Hatefi, 1990; Tong, Glavas and Bragg, 1991; Hu *et al.*, 1995). Models for the open $NADP^+$ class and the apo NNT were based on the closed NNT as the starting model. Due to the lack of reliable amino acid side chain information in the cryo-EM density, apo-dI was modelled as a polyalanine model based on the 6QTI dI and fit into the density as a rigid-body.

2.8. Model building and refinement: complex I

The models presented here are based on the initial 3.9 Å structure of ovine complex I determined previously (PDB 5LNK) (Fiedorczuk *et al.*, 2016). Membrane domain and peripheral arm were built on the basis of 2.5 Å and 2.3 Å focus-refinements of all the particles belonging to the open classes from the turnover dataset. This allowed us to improve the geometry, fit and completeness of the models significantly (Appendix 2-6). We were also able to model additional lipid molecules, as well as several known post-translational modifications, including N-terminal acylations of subunits B18, B14.5b and B14.7, N-formyl methionines in ND6, ND5 and ND4L, symmetric dimethylation of 49 kDa Arg85 and phosphorylation of 42 kDa Ser36 (Carroll, Fearnley and Walker, 2006) (see Chapter 4). For accurate modelling of water molecules, particularly to avoid false positives, we filtered the maps by local resolution and resampled them at 0.5 Å per pixel in Relion. After this procedure, water molecules displayed strong signal (~2 rmsd), had nearly spherical densities, were not clashing with other atoms and participated in hydrogen bonds, which are all strongly indicative of real water molecules (Prisant *et al.*, 2020). This allowed automatic placement of water molecules in COOT, which were then all checked and corrected manually, to leave only waters with clear density and fulfilling geometry criteria. Using this procedure, we could reliably place water molecules not only into 2.3-2.8 Å structures but also into the closed MD focus refined map (overall resolution 2.9 Å, higher in the center). The two highest resolution open PA and open MD models were then rigid-body fit into all the classes, corrected in *Coot* (Emsley *et al.*, 2010) and refined using *PHENIX* software (Adams *et al.*, 2010). We used two rounds of a single cycle of group ADP refinement followed by three cycles of global 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 as described previously (Letts *et al.*, 2019). To maintain consistency, we performed this procedure first with rotamer fit option enabled, checked and corrected the model in coot and then finally repeated the procedure without rotamer fit.

2.9. Model analysis and figure preparation

Evolutionary conservation scores were projected onto the transhydrogenase structure using ConSurf (Landau *et al.*, 2005). Buried water molecules in transhydrogenase were modelled using the original Dowser software (Zhang and Hermans, 1996). Proton translocation pore was predicted and analysed using a MOLEonline server (Pravda *et al.*, 2018). Interdomain contacts and binding energy of nucleotides were analysed using the PISA server (Krissinel and Henrick, 2007; Evgeny, 2010). Models were evaluated using MolProbity and EMRinger (Chen *et al.*, 2010; Barad *et al.*, 2015). Clustal Omega was used for sequence alignment (Larkin *et al.*, 2007; Goujon *et al.*, 2010). Models and density maps were visualised for analysis and figure preparation purposes in PyMOL 2.2.3 and UCSF Chimera (Pettersen *et al.*, 2004).

3. Structure and mechanism of transhydrogenase

3.1.Biochemical characterisation

3.1.1. Catalytic activity

A novel way of purification of transhydrogenase from ovine mitochondria described in Chapter 2 led to a highly active preparation. Crucial for the development of this was screening detergents based on their effects on enzyme activity (Figure 2.1) and optimising the buffer conditions by a ProteoPlex method (Chari *et al.*, 2015).

Activity depended on the type of detergent used during the final size exclusion chromatography step. We prepared transhydrogenase in two different detergents, LMNG and FOM and three different nanodiscs. In FOM and the largest nanodiscs (1E3) the activity of the enzyme was the highest at ~20 U/mg (Figure 3.1). In smaller nanodiscs or LMNG, the activity of NNT was around 10 U/mg. These values are comparable to previous measurements on the bovine NNT, which also range between 10 and 25 U/mg (Wu, Alberta and Fisher, 1986; Sazanov and Jackson, 1995).

The fact that nanodiscs displayed one of the highest activities is not surprising as they provide a more native lipidic environment than detergent. It was also demonstrated that transhydrogenase requires specific phospholipids for its function (Rydström *et al.*, 1976; Vandock *et al.*, 2011). The fact that the activity decreased slightly in smaller diameter nanodiscs could mean that the transmembrane region was becoming constricted in those nanodiscs, interfering with the protein dynamics required for reaction catalysis.

The fact that NNT-FOM had high activity can be explained by FOM being a uniquely mild and non-disruptive detergent having a relatively low affinity for proteins and lipids due to its fluorinated acyl chain (Frotscher *et al.*, 2015). This could on the other hand lead to the low stability of NNT purified in FOM (Chapter 3.1.2). The fact that LMNG did not inhibit activity, while DDM was a strong inhibitor can potentially be explained by three facts (Figure 2.1). Firstly, LMNG has two acyl groups and two maltoside groups, making it very bulky and potentially unable to access an inhibitory site to which DDM binds. Secondly, DDM being smaller could better solubilise some tightly bound lipids in the tight binding sites. Thirdly, DDM inhibition could mimic palmitoyl-CoA inhibition, and the acyl groups of palmitoyl-CoA and DDM could access the same inhibitory site (Rydstrom, 1972). DDM inhibition was transient and NNT activity was recovered by exchanging DDM-purified NNT into LMNG.

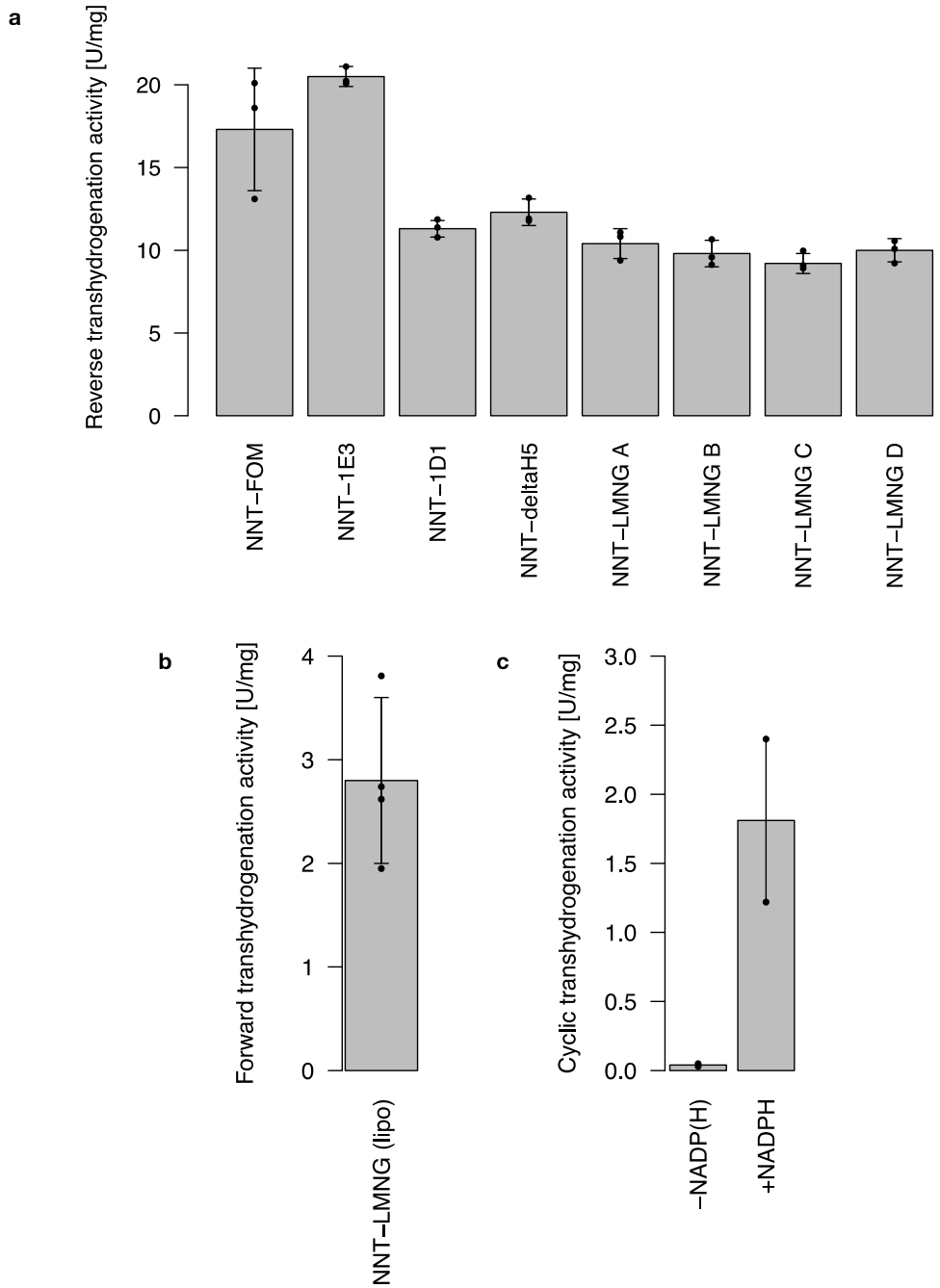


Figure 3.1. Catalytic activities of different NNT preparations.

Ovine NNT catalyses all the known activities: (a) reverse activity, (b) forward activity and (c) cyclic activity. Reverse activity is generally slightly higher in large nanodiscs and FOM than in LMNG. No cyclic activity in the absence of added NADP(H) indicates absence of co-purified NADP(H) with the NNT.

We also measured the cyclic transhydrogenation reaction, i.e. the reduction of $APAD^+$ by NADH (Sazanov and Jackson, 1995). This reaction is not believed to be of physiological significance and is thought to be a result of tightly bound NADP(H) nucleotide in the dIII domain. After NADH reduces $NADP^+$, the resulting NADPH stays bound to the dIII domain long enough to reduce $APAD^+$ that will have bound to the same site as the NADH. This reaction

is not coupled to proton transfer and suggests that NADP(H) binding or release is coupled to the proton pump. The reaction rates of around 2 U/mg are similar to those reported in the literature for the bovine NNT (Figure 3.1) (Sazanov and Jackson, 1995). The fact that we could record no cyclic transhydrogenation in the absence of added NADP(H) also confirms that ovine NNT purified by the method described here does not contain any bound nucleotides and can be considered in the true apo form.

Finally, we measured forward transhydrogenation activity, which was several fold lower than reverse transhydrogenation activity, consistent with the previous measurements and the fact that reverse transhydrogenation is coupled to proton pumping, while forward transhydrogenation is coupled to proton motive force consumption (Figure 3.1) (Eytan, Eytan and Rydström, 1987).

3.1.2. Stability

Transhydrogenation activity was highly stable at 4ºC in LMNG. After 14 days at 4ºC, NNT-LMNG samples still retained ~75% activity, indicating that the preparation in this detergent is highly stable (Figure 3.2). In contrast, in FOM, the preparation lost almost all the activity after 48 h.

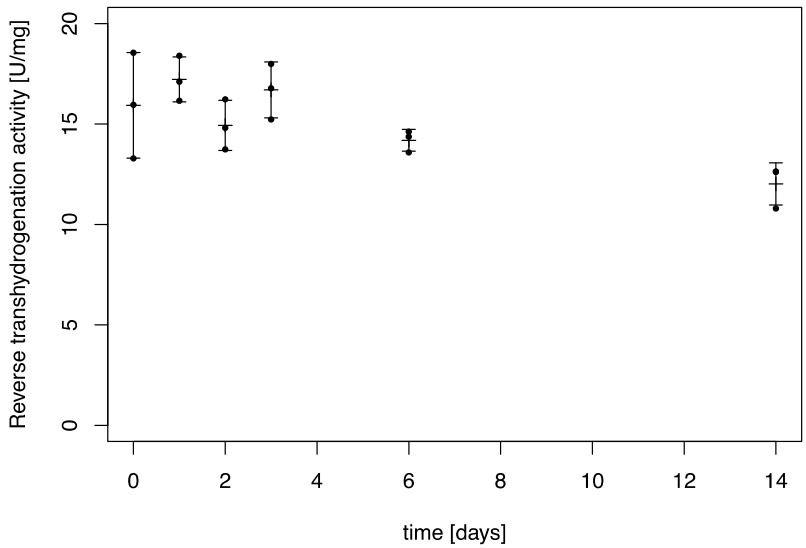


Figure 3.2. NNT is highly active and stable at 4ºC when purified in LMNG.

Activity of NNT-LMNG remains undiminished over several days when stored at 4°C. Error bars represent standard deviations based on n=3 independent measurements.
Adapted from Kampjut & Sazanov, 2019.

3.1.3. Proton-pumping

Proton pumping activity coupled to reverse transhydrogenation was measured to assess the coupling of the NNT preparation. NNT was reconstituted into proteoliposomes as described in Chapter 2 and reverse transhydrogenation activity was measured before and after the addition of the protonophore (CCCP). About 10-fold acceleration of the reaction by the addition of protonophore indicates that NNT generates a large proton motive force which inhibits transhydrogenation reaction, but this inhibition is released by the protonophore (Figure 3.3). It also indicates that the proteoliposomes were well-coupled and not leaky.

To assess the influence of lipid composition on the activity of NNT, we reconstituted it into liposomes of varying lipid compositions and measured the activity. The highest activity was recorded in POPC and DOPC liposomes with or without the addition of cardiolipin, while the lowest activity was recorded in asolectin liposomes and liposomes containing phosphatidyl ethanolamine (Figure 3.3). This is consistent with mitochondrial inner membrane lipid composition as it includes large amounts of phosphatidyl choline and cardiolipin. Recently it was also demonstrated that NNT from cestodes requires PC for optimal activity, although PS and PE could also save a bit of activity (Fu, Ma and Fioravanti, 2019). The downside of this study is, however, that the assay was done on DDM-purified NNT and DDM is a known inhibitor of NNT.

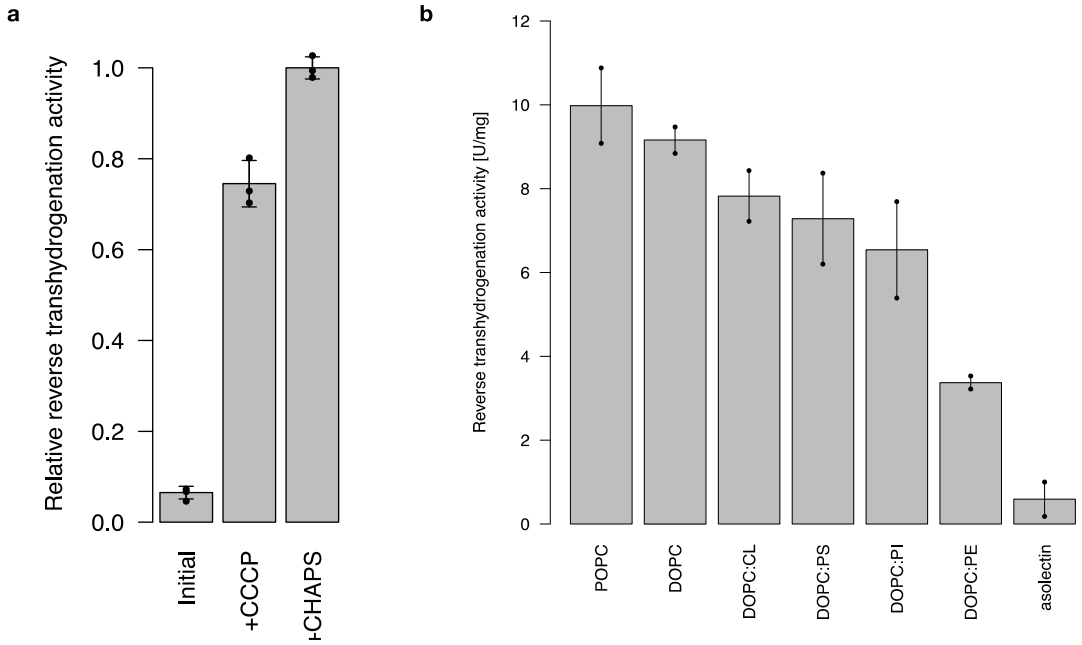


Figure 3.3. Proton pumping assay.

Reconstitution of purified NNT into DOPC liposomes shows that the reverse transhydrogenation reaction is tightly coupled to proton transfer as it is stimulated ~10-fold by proton gradient uncoupler CCCP and by solubilisation in CHAPS detergent. Error bars represent standard deviations based on n=3 independent measurements. Panel (a) adapted from Kampjut & Sazanov, 2019.

3.1.4. Inhibitors of NNT

There are few known inhibitors of NNT and essentially none very specific micro-molar small molecule inhibitors (Rydstrom, 1972; Bragg and Hou, 1999). It was described that divalent cations are inhibiting transhydrogenase, potentially by binding at the proton-translocation site (Sazanov and Jackson, 1993; Whitehead *et al.*, 2005). With the ovine NNT-LMNG preparation we were able to reproduce this effect in 10 mM $MgSO_4$, which inhibited NNT by about 50% (Figure 3.4). Furthermore, we observed a slight inhibition of NNT-LMNG by prolonged incubation in NADPH (5 mM, on ice), an effect which is easy to explain by the structural changes induced by NADPH. Such inhibition was not observed with $NADP^+$.

Finally, because NNT is involved in redox regulation and because cysteine-893 was shown to be susceptible to alkylation before we hypothesised that glutathione could regulate NNT activity by glutathionylation (Yamaguchi and Hatefi, 1989). We could observe that oxidized glutathione, but not reduced glutathione, inhibits NNT by about 50%.

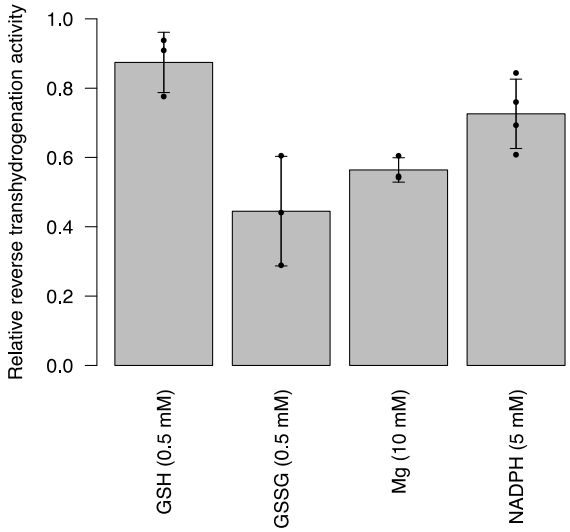


Figure 3.4. Inhibitors of NNT.

The strongest inhibitors were oxidised glutathione (GSSG) and Mg. Reduced glutathione did not inhibit NNT, while prolonged incubation with high concentration inhibited NNT slightly.

3.2. Overall structure of transhydrogenase

3.2.1. Model of intact mammalian transhydrogenase

We solved the structure of complete and intact ovine NNT at up to 2.9 Å resolution using cryo-EM (Appendix 1). The highest resolution was reached with the double face-down conformation in the presence of $NADP^+$ (Figure 3.5). Excellent EM maps resulted in high quality full atomic models of mammalian NNT including modelling of almost entire polypeptide chain, bound dinucleotides and several lipid molecules (Figure 3.6).

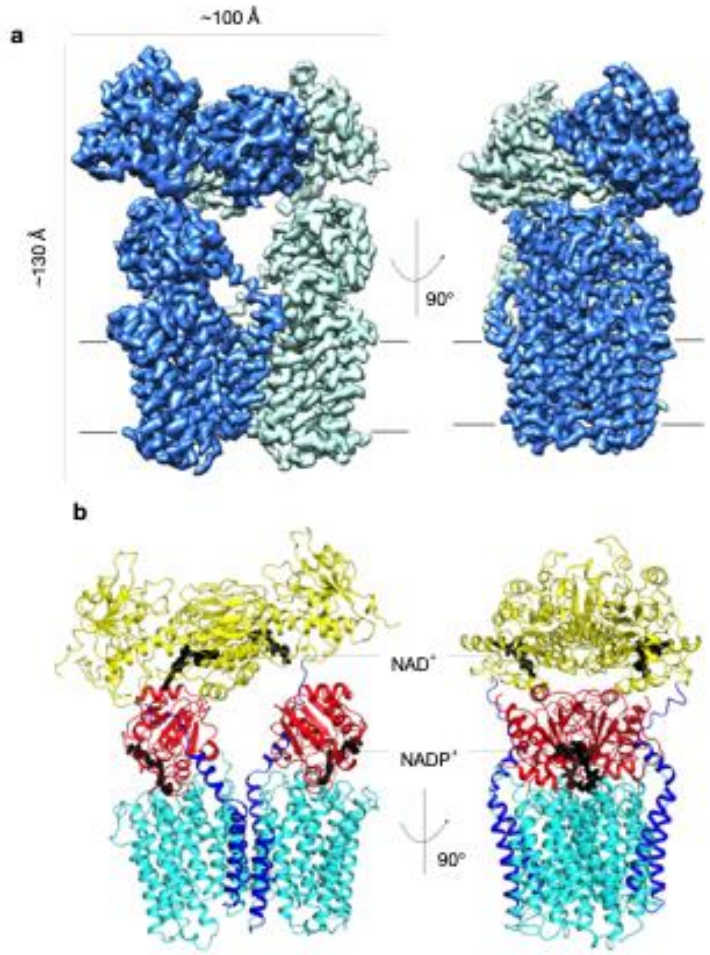


Figure 3.5. Model and density of double-face down NNT.

***a**, The cryo-EM density of transhydrogenase in the "double face-down" $NADP^+$ state, with two monomers coloured differently. **b**, Atomic model of mammalian NNT with domains coloured as in panel a, and mammalian-specific elements dI-dII linker, TM1 and TM5 highlighted in blue. Bound nucleotides are shown in black.*
Figure adapted from Kampjut & Sazanov, 2019.

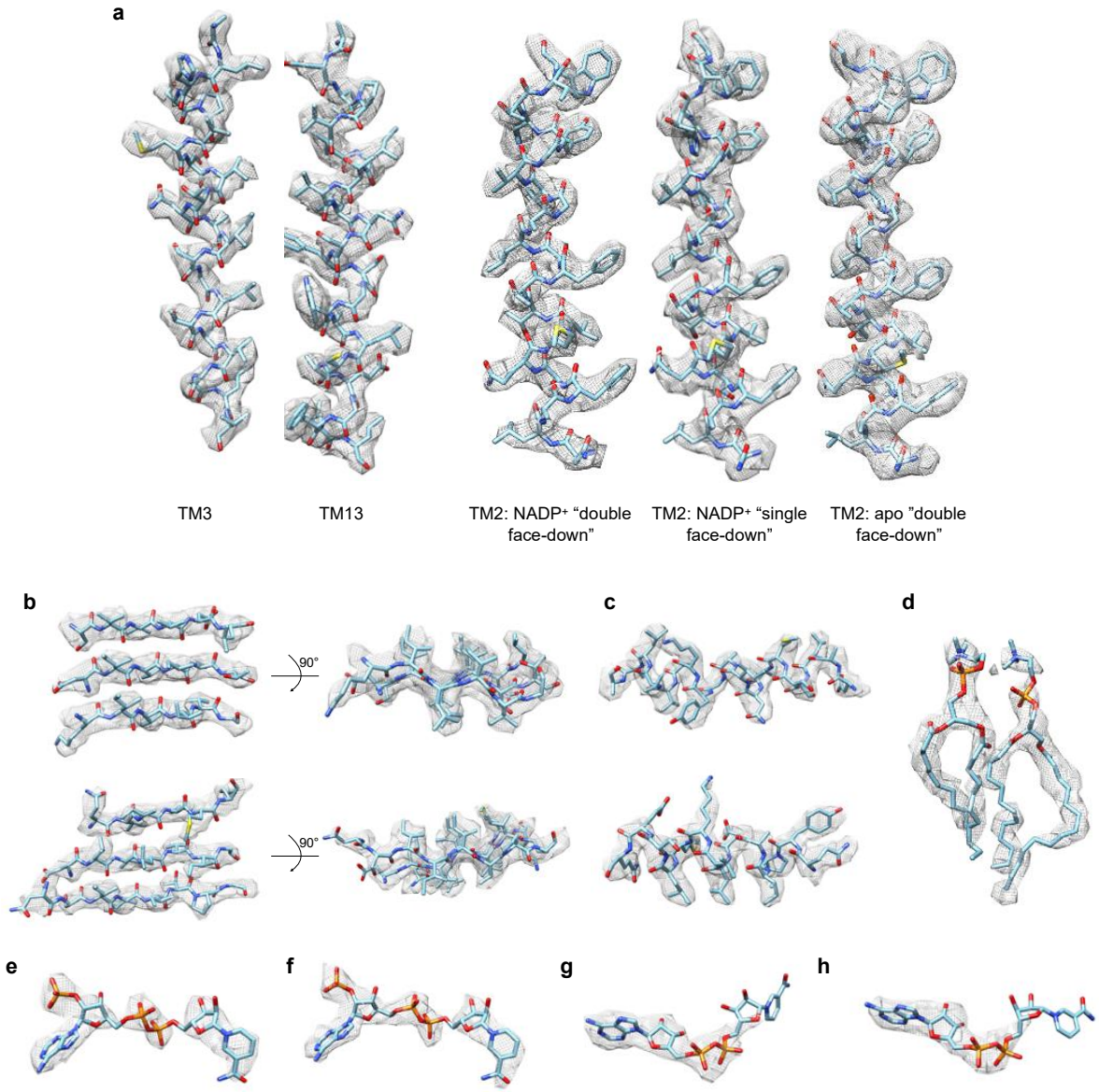


Figure 3.6. Cryo-EM density examples of the "double face-down" class in the presence of NADP$^+$.

***a**, Density of the transmembrane helices lining the proton channel (TM3 and 13). Density comparison of the TM2 between the "double face-down" and "single face-down" (dIII-detached monomer) NADP$^+$ classes, as well as apo class, which remains in the same conformation in all states. **b**, Beta sheet density of dI (top) and dIII (bottom). **c**, Alpha-helical segments of dI (top) and dIII (bottom). **d**, Two phosphatidyl-cholines bound in the cavity enclosed by TM1, 2 and 6. **e**, dIIIa-NADP$^+$ density. **f**, dIIIb-NADP$^+$ density. **g**, Partial NAD$^+$ density in dIa. **h**, Partial NAD$^+$ density in dIb. The density for lipids and putative NAD$^+$ is discussed in Methods.*
Adapted from Kampjut & Sazanov, 2019.

The arrangement of the domains and organisation of the dimer in ovine NNT is similar to the *T. thermophilus* NNT, apart from the fact that both dIIIs are bound to dII, resulting in an almost symmetric structure. The symmetry is broken due to the slight tilting of dI_2, which interacts loosely with dIII in one monomer (dIIIa) and not at all in another (dIIIb) (Figure 3.7). This tilting of dI_2 also resulted in the local resolution of dI_2 to be worse than in dIIdIII part of the molecule. The final structure is thus composed of two focus-refined maps (Figure 3.8).

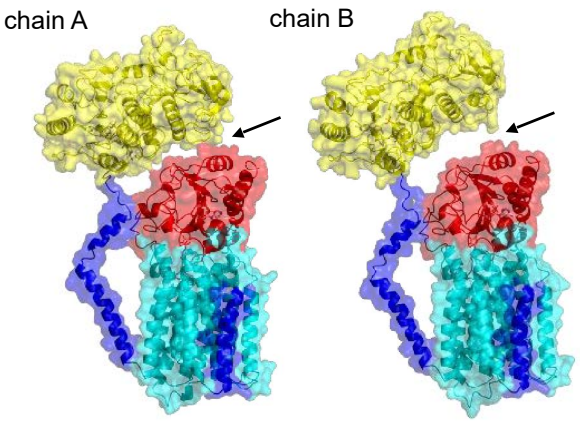


Figure 3.7 Comparison of the two monomers in the double-face down NADP+ class.

The view from the dimerization interface reveals a tilt of dI_2 and asymmetry in dI-dIIIa/b contacts. Adapted from Kampjut & Sazanov, 2019.

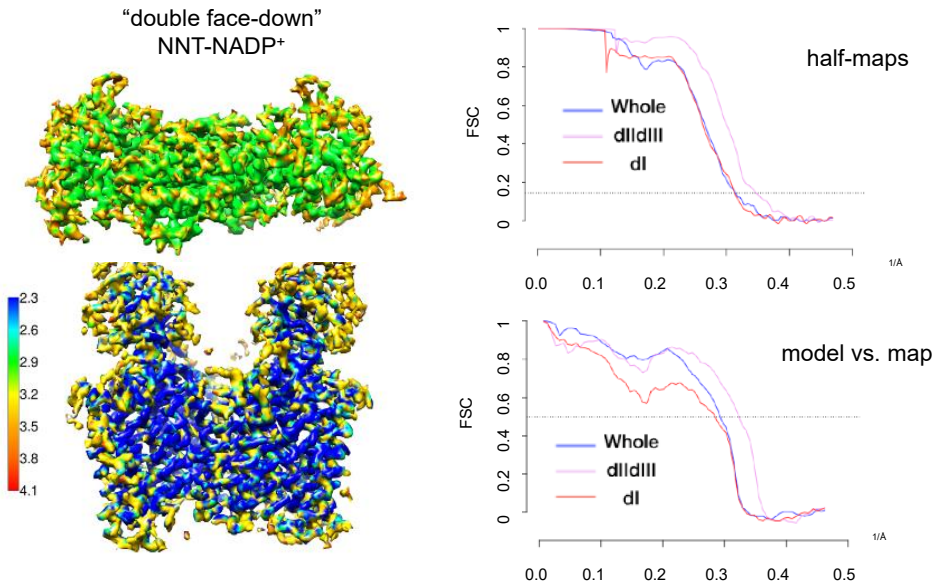


Figure 3.8. Local resolution and FSC curves for the "double face-down" NNT-NADP+ structure.

Adapted from Kampjut & Sazanov, 2019.

3.2.2. Comparison of conformations in different datasets

To define the mechanism, we solved the structures of ovine NNT at pH 7.4 in the apo, $NADP^+$-bound and NADPH-bound forms. We focused on the comparison between NADP+ and NADPH bound states because these two nucleotides were shown to promote distinct conformations in early biochemical work. The coupling mechanism of transhydrogenase likely involves large-scale conformational changes that were shown to be induced by NADPH (but not $NADP^+$ or NAD(H)) in the mammalian enzyme. Upon incubation with NADPH, some parts of the protein (dI-dII linker and TM6-7 loop on the matrix side of dII, as now shown by our structure) become susceptible to tryptic hydrolysis and Cys-893 becomes more susceptible to alkylation (Yamaguchi and Hatefi, 1989; Yamaguchi, Wakabayashi and Hatefi, 1990). In contrast to the mammalian enzyme, both $NADP^+$ and NADPH can induce these long-range conformational changes in *E. coli* (Tong, Glavas and Bragg, 1991; Fjellström, Johansson and Rydström, 1997). Furthermore, the Jackson mechanism did not explain why $NADP^+$ and NADPH would promote different conformations of the entire enzyme. NADH and NAD^+ on the other hand do not lead to different global conformations of the NNT.

The only observed 3D class of particles in the apo-NNT dataset adopted an unexpected and previously unobserved "double face-down" conformation with both dIII domains tightly attached to dII in the "face-down" orientation, as described above (Figure 3.9). This means that the overall architecture of the dimer does not allow for tight interactions between dI and dIII in both monomers, whether dIII is facing "up" or "down". On the other hand, dIII-dII interaction is permitted in both monomers simultaneously.

In the $NADP^+$ dataset the main class showed a similar conformation, but with higher resolution, possibly due to the stabilising effects of the bound nucleotide (overall resolution 3.2 Å, focused refinement of dII_2dIII_2 to 2.9 Å and dI_2 to 3.2 Å) (Figure 3.9). $NADP^+$ dataset also resulted in the structure of the less populated class ("single face-down") in which one dIII is detached from dII, dI_2 is more tilted, probably to allow detachment and rotation of dIII (Figure 3.10).

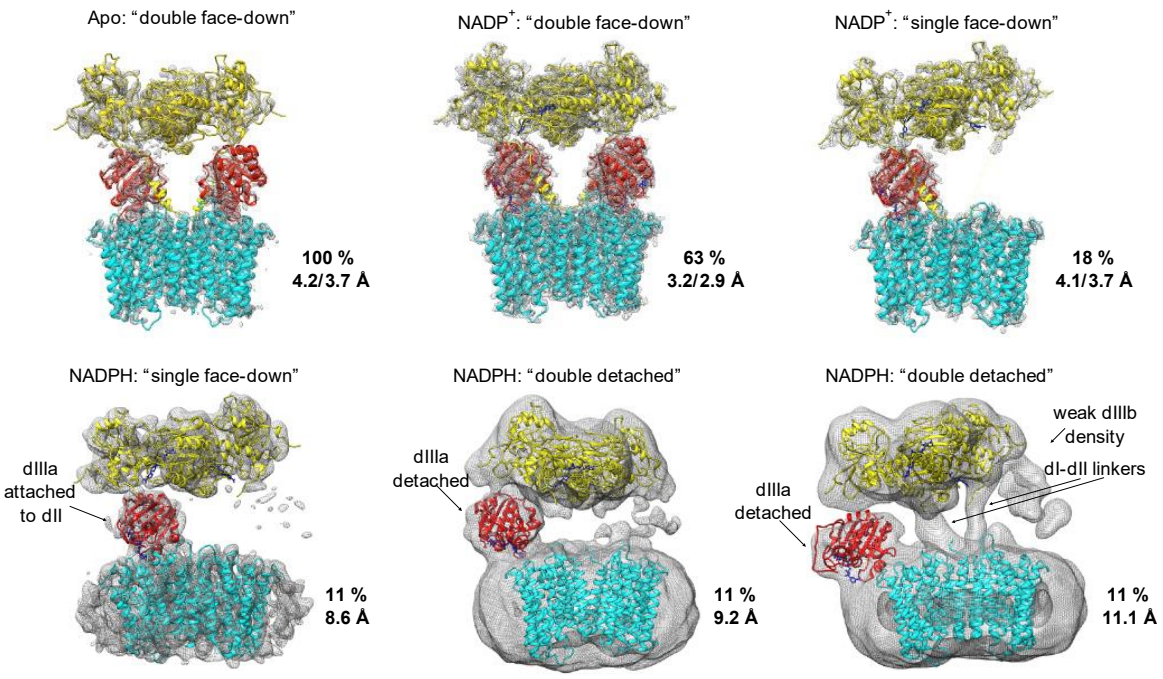


Figure 3.9. Structures and relative abundances of different conformations of NNT observed in the datasets.

"Double face-down" conformation was the only form observed in apo-NNT and the major form in the NADP+ datasets, "single face-down" or "one dIII-detached" conformation was a minor form in NADP+ and NADPH datasets and "double dIII-detached" was a major form in NADPH dataset.
Adapted from Kampjut & Sazanov, 2019.

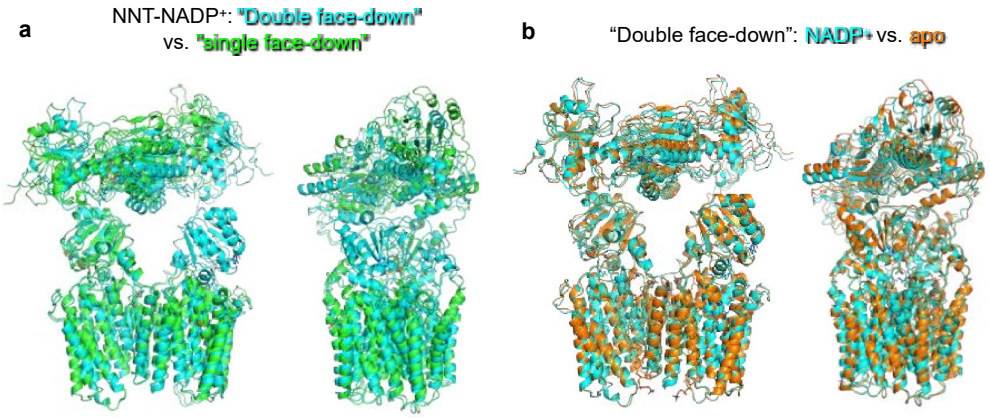


Figure 3.10. Comparison of NNT conformations

***a**, Overall comparison of "double face-down" (cyan) and "single face-down" (green) classes of NNT. An increased dI tilt, dIIIb detachment and dIIb conformation change are visible. **b,** Overall comparison of "double face-down" NNT-NADP+ (cyan) and apo NNT. dII-dIII are in the same conformation and dI2 is slightly more tilted in the apo-NNT.*
Figure adapted from Kampjut & Sazanov, 2019.

Finally, in the presence of NADPH, NNT underwent a drastic global conformational change, as predicted before (Yamaguchi and Hatefi, 1989; Yamaguchi, Wakabayashi and Hatefi, 1990). We did not observe any "double face-down" class and only ~11 % of particles had one dIII weakly bound to dII and the rest of the particles corresponded to many different "double-detached" classes with both dIII domains detached and tethered only by a dII-dIII linker (Figure 3.9). In particles with both dIIIs detached, dI exhibited a particularly striking degree of tilting in all directions as revealed by multibody refinement and principal component analysis in RELION 3.0 (Figure 3.11). None of these NADPH classes reached resolution better than 9 Å.

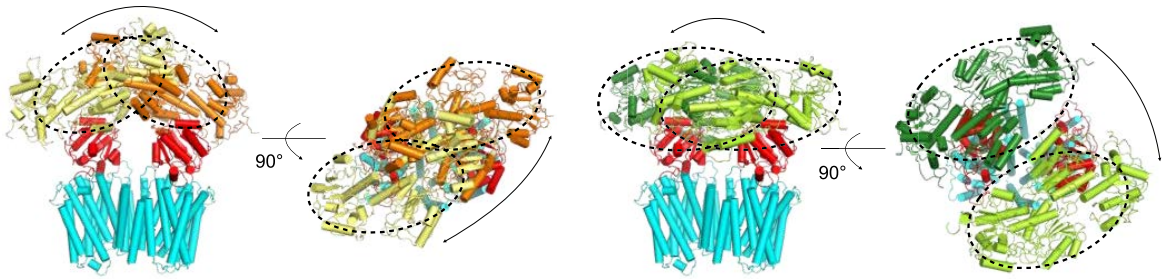


Figure 3.11. High degree of freedom of dI$_2$ in double-detached NNT.

NNT-NADPH particles were analysed using multibody refinement in RELION 3.0 using dI2 and dII2 masks which revealed a high degree of tilting of dI2 domain in all directions, consistent with both dIII being dissociated from the rest of the particle and disrupting the overall solidity of the enzyme. Two bins of particles from the extreme ends of the distribution along two principal components are aligned by dII2 and the difference in dI2 tilt is depicted in different shades of yellow and green.
Adapted from Kampjut & Sazanov, 2019.

Interestingly, we did not find any evidence for an asymmetric conformation as observed in *T. thermophilus*. In none of the classes did we observe the face-up dIII conformation, although such a conformation is clearly necessary for hydride transfer to take place during catalysis. Previous studies suggest that the dI-dIII interface formation requires binding of the corresponding nucleotide to dI and, once formed, has a very short life span as the hydride transfer is fast (Hutton *et al.*, 1994; Venning *et al.*, 1998, 2000). Alternatively, the architecture of the mammalian NNT with its dI-dII linker may prevent the observation of a stable face-up dIII conformations, although this remains to be established in future studies.

Because we did not observe a distinct asymmetric class with one dIII in face-down and another one in face-up orientation, we conclude that NADPH on its own is not sufficient for a stable dI-dIII (face-up) interface formation and promotes dIII detachment. $NADP^+$ and NADPH on their own promote dIII detachment to various degrees but do not lead to a stable

dI-dIII (face-up) interface formation. Most of these conformations (apart perhaps from apo-NNT, where we could only observe one class) would be expected to be transient during the turnover as dIII must shuttle between dI and dIII.

To summarize, transhydrogenase can exist in other conformations with different combinations of dIII in "up", "down" or detached positions (Figure 3.12).

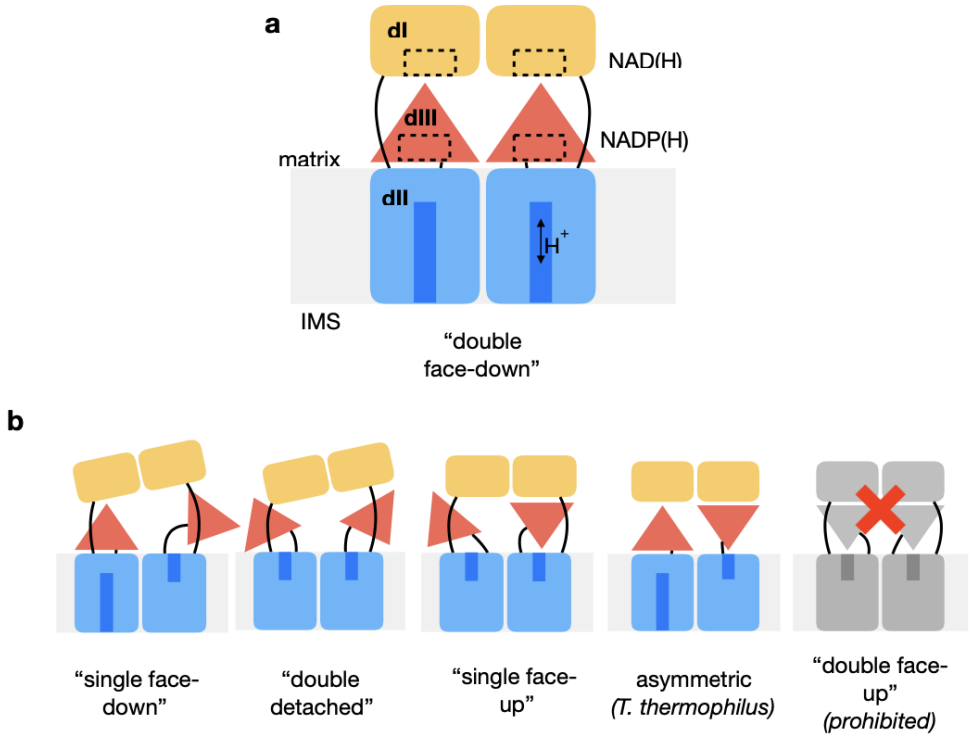


Figure 3.12. Conformations of transhydrogenase.

***a**, Schematic representation of the domains, linkers, binding sites and conformation of the mammalian transhydrogenase in the "double face-down" (both dIII attached to dII) conformation (apo and $NADP^+$ datasets). Other possible conformations of NNT include one or both dIII detached, the asymmetric conformation in T. thermophilus (N.B. the absence of dI-dII linker) and the sterically prohibited double-face-up conformation. Adapted from Kampjut & Sazanov, 2019*

3.2.3. Changes of domain interfaces and comparison to *T. thermophilus* NNT

We modelled the ovine asymmetric class by rigid-body fitting separate domains of ovine NNT to the *T. thermophilus* architecture of NNT to understand and compare interfaces that dIII forms with dI and dII (Figure 3.13). dIII forms the interface with dII and dI with roughly the same surface – helix 4 of dIII makes contacts with both dI and II (lower left and upper right panels of Figure 3.13). The size of the buried interface and free energy of binding

of dIII interfaces is smaller than of dimer-forming interfaces dII-dII and dI-dI in both *T. thermophilus* and *O. aries* (Table 3.1). dIII-dII interfaces seems stronger than dI-dIII, probably because it is slightly larger and contains a salt bridge R544-D942, which is an important stabilising element (lower left panel of Figure 3.13). This suggests that a conformational or electrostatic signal is needed to break the dIII interfaces during the catalytic cycle, which will be discussed in Chapter 3.3.

In contrast to what has been suggested for the *T. thermophilus* enzyme, dIII attachment to dII does not change the conformation of the short lateral helix CL8 nor does it disrupt the conserved salt bridge R839-D787 (Leung *et al.*, 2015). The importance of this salt bridge could be purely to stabilise the dII or the interface, or to transmit conformational changes from dIII to dII via the peptide linker. Either way, the detrimental effects of mutating either of the participating residues are expected.

	O. aries		***T. thermophilus***	
Interface	**Area [Å^2]**	**ΔG [kcal/mol]**	**Area [Å^2]**	**ΔG [kcal/mol]**
dI-dI	1527.5	-16.3	1867.1	-20.1
dII-dII	2602.9	-65.2	1369.2	-27.3
dIII-dI	738.4	-9.0	794.8	-11.8
dIII-dII	1225.5	-21.7*	1004.2	-16.5

Table 3.1. Surface area and free energy of binding of the inter-domain interfaces in ovine and T. thermophilus NNT.

**Could be overestimated, because the interface is dissolved by mutating a single salt bridge.*

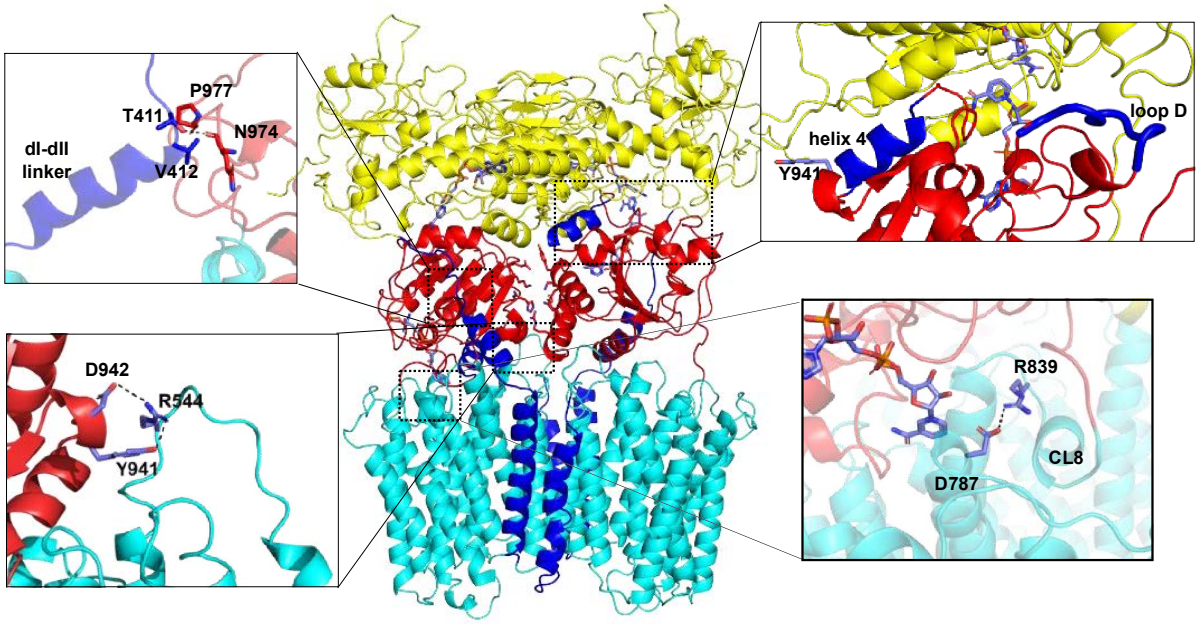


Figure 3.13. Various interfaces of dIII in the asymmetric model of transhydrogenase.

A homology model of the asymmetric ovine NNT based on the T. thermophilus dI2dIII heterotrimer structure (PDB 4J16). N.B. putative interacting residues at the dIIIup-dIIIdown interface. dI-dII linker contacts the loop D of dIII in the face-down conformation. D942 and Y941 from dIII form hydrogen bonds with R544 on CL2 loop stabilising the dII-dIII interface. Helix 4 and loop D also contribute towards formation of the dI-dIII interface but Y941 is too far to interact with dI. R839 and D787 always form a salt bridge, regardless of dIII conformation.
Adapted from Kampjut & Sazanov, 2019.

3.2.4. Mammalian-specific features

The fold of the three domains in ovine NNT is conserved with the bacterial structures, as expected based on the sequence similarity of ~ 40% and conserved function. The mammalian-specific supernumerary helices TMH1 and TMH5 form an unexpected extended dII dimer interface (dark blue in Figures 3.5, 3.7 and 3.13). Relatively large tilt of the TMH1 allows it to contact the entire TMH5 of the opposite monomer, possibly stabilizing dimer formation in mammals.

The linker between dI and the additional helix TM1 in dII contacts dIII in the 'face-down' orientation (upper left panel of Figure 3.13). C-terminal part of the dI-dII linker (residues 410-424) forms a loosely defined α helix in the face-down state, while the N-terminal region (400-410) is a flexible loop and only displays very weak density. V412 and T411 from the linker are forming van der Waals interactions with the N974-P977 stretch of the loop D (residues 974-986) on dIII. This explains why K410-T411 bond is protected from trypsinolysis in the apo and $NADP^+$-bound state (Yamaguchi, Wakabayashi and Hatefi, 1990). In the dIII-

detached state, the density for the C-terminal part of the linker disappears, which could indicate partial unfolding of the helix to accommodate an increased distance between dI and dII. The role for the dI-dII linker is not clear and the sequence is not well conserved but connection from the dI domain via TM1 to TM5 of the opposing monomer could help in the coordination of the catalytic cycle events between the monomers as discussed below.

3.2.5. Structure in lipid nanodiscs

Structures of NNT in MSP1E3D1 or MSP1D1ΔH5 nanodiscs in the presence of $NADP^+$ were refined to between 10 and 13 Å. Overall the protein exhibited the same global conformation as in LMNG (Figure 3.14). This confirms that detergent solubilisation does not radically change the conformation of NNT. In NNT-1E3D1 we could classify a major class (54%) and a minor class (46%). The minor class had a slightly lower resolution and appeared to have more dissociated dIII. Major class could correspond to the double face-down and the minor to the single-face down, although this is hard to establish at this resolution and also the major class had dIIIb density weaker than dIIIa. In NNT-ΔH5 we could only classify a major class. The nanodisc density was noticeably smaller, but only by about 10 Å. Elution at 10.3 mL still suggests an apparent molecular weight of 440 kDa. This again underscores the finding that the global conformation of NNT is not dependent on the choice of surfactant. The fact that higher resolution could not be achieved could be due to the microscope hardware used for this project, because of the heterogeneity of the protein conformations or because of the large unstructured density corresponding to the lipidic nanodisc interfering with alignment.

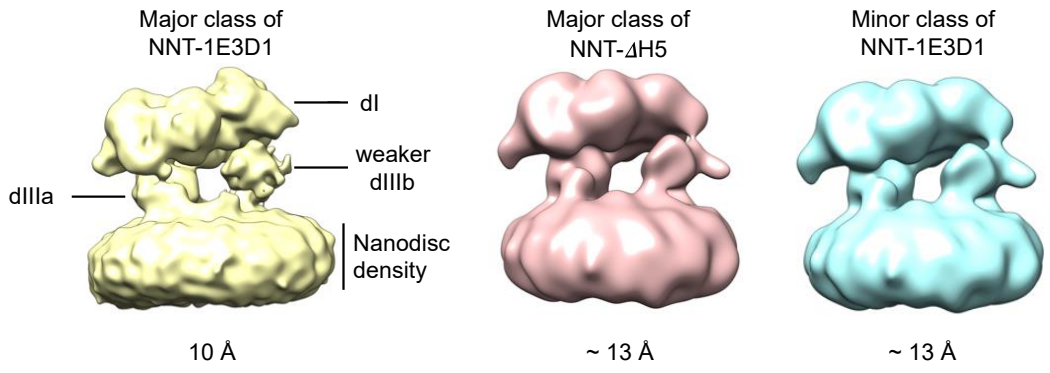


Figure 3.14. Overall conformation of NNT-NADP+ in lipidic nanodiscs.

In all of the structures, dIIIb has a weaker density than dIIIa, suggesting that these structures are mixtures of double- and single-face down conformations as observed in NNT-LMNG. In NNT-ΔH5 nanodiscs the nanodisc diameter is about 10 Å smaller than in NNT-1E3D1 nanodiscs.

3.3.Detailed comparison of NNT conformational states

3.3.1. Structure of the dII and the proton translocating pathway

The ovine dII forms a dimer of 14 transmembrane helices each (Figure 3.15a). The twelve conserved transmembrane helices adopt the same fold as demonstrated in the isolated *T. thermophilus* dII domain (Leung *et al.*, 2015; Padayatti *et al.*, 2017). Two supernumerary helices form the outer part of the dimer interface and contribute to a formation of a semi-enclosed lipid binding pocket in which two molecules of phosphatidyl choline were observed (Figure 3.15c). The IMS-facing side of the dII has several exposed glutamate residues which contribute to a strong negative charge, relevant for proton binding (Figure 3.15b).

The likely proton translocation pathway is formed by a hexagram of TM3-4, TM9-10, and TM13-14, as proposed earlier based on the mutagenesis studies in *E. coli,* structures of *T. thermophilus* dII and molecular dynamics (MD) simulations (Glavas, Hou and Bragg, 1995; Leung *et al.*, 2015; Padayatti *et al.*, 2017). These are the most conserved helices of the entire structure that contain several polar residues. Furthermore, the architecture around TMH13 and TMH14 is highly unusual and completely conserved between *O. aries* and *T.* thermophilus. TMH13 and 14 are significantly shorter than the rest and connected with a very short loop on the IMS side and by a highly conserved salt bridge on the matrix side (R893-D787). TMH14 is followed by a short horizontal amphipathic alpha helix CL8 (Figure 3.15d).

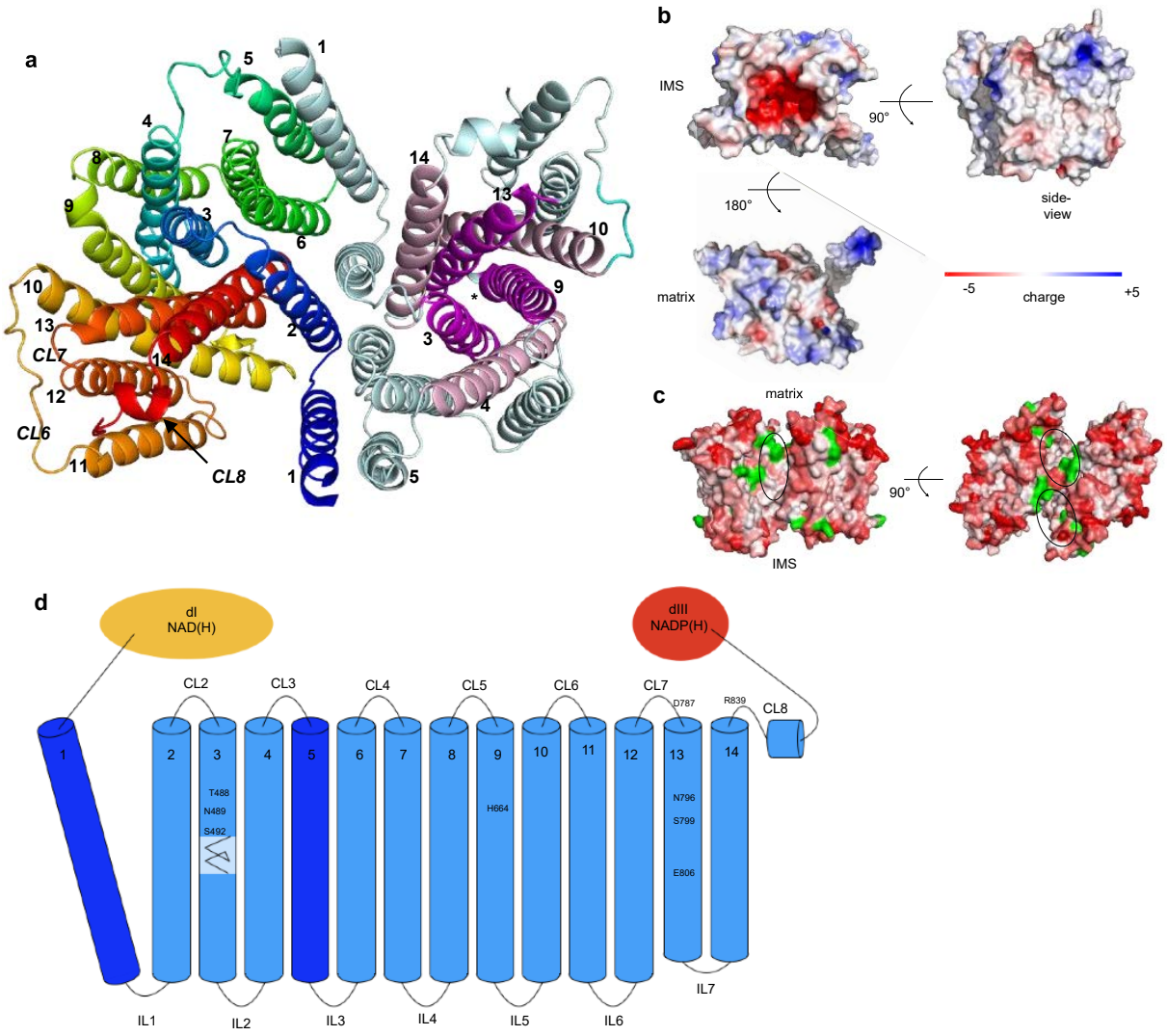


Figure 3.15. Architecture of the transmembrane domain dII.

***a**, Organisation of transmembrane helices in the NNT dimer. The monomer on the left is coloured blue-to- red from N- to C-terminus and the monomer on the right has the inner proton channel-lining TM helices highlighted in magenta and the outer scaffolding helices in pink. The dimer interface consisting of eight helices is significantly larger than the four-helix interface of T. thermophilus. **b,** Electrostatic surface potential of the proton entry cavities on the matrix (bottom) and IMS sides (top) as well as that of the membrane-facing side (right). **c,** Hydrophobicity of residues on the surface of dII_2 (coloured white to red from hydrophobic to hydrophilic). Surface-exposed tyrosine and tryptophan residues, which often delineate lipid membrane surface, are highlighted in green and the lipid-binding pocket is circled. **d**, The architecture of mammalian dII. Supernumerary helices TM1 and TM5 are coloured in a darker shade of blue. Residues of the proton transfer pathway are shown in circles. 3_{10}-helix stretch within TM3 is depicted as a triangular helix.*

Adapted from Kampjut & Sazanov, 2019.

The proton translocation pathway was proposed to have two protonatable residues: one on either side of the membrane, connected by a dry channel. In ovine NNT, N-side (matrix) part of the proton channel consists of a single protonatable histidine and two asparagine residues (H664 on TMH9, N489 on TMH3 and N796 on TMH13), while the P-side (IMS) part

was proposed to consist of a single glutamate (E806 on TMH13) (Figure 3.16). The dry connecting channel is composed of helices 3, 9 and 13 and strikingly, this is where a clear and significant distortion of helix TM3 occurs: two turns of 3_{10} helix, the only such distortion in dII. A similar distortion in a similar position is also present in the TM3 of *T. thermophilus* dII structure (PDB 5UNI and 4O93), exactly where the key α_2H42 is located. Such disruptions of secondary structure in TM helices happen almost exclusively in functionally important sites, where flexibility of otherwise rigid TM helices is required (Cooley, Arp and Karplus, 2010; Efremov and Sazanov, 2011). Additionally, secondary structure is disrupted also in TM9 around His664, with backbone oxygen atoms of residues 661, 662 and 667 facing away from the helix-forming hydrogen bond network. Therefore, it is likely that these key TM3 and TM9 residues undergo conformational changes during the catalytic cycle, promoting proton translocation.

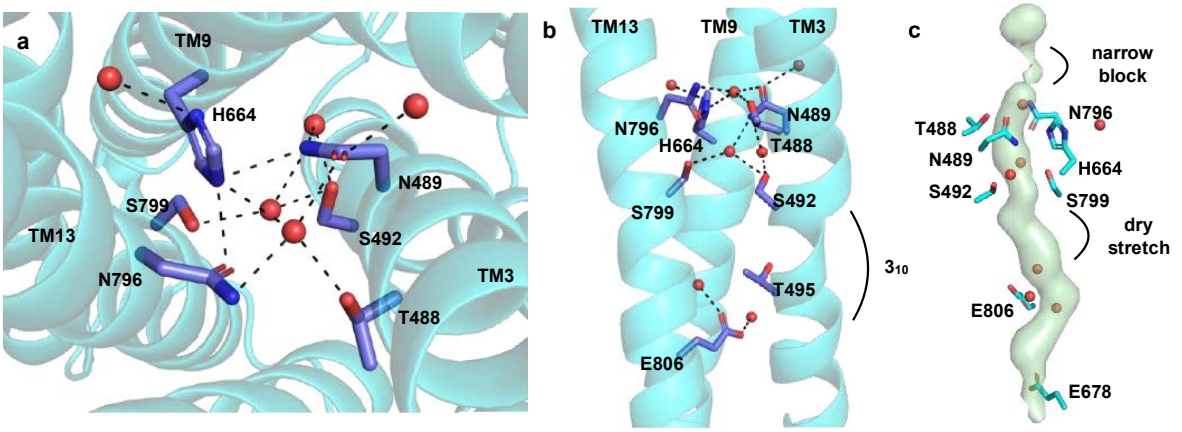


Figure 3.16 The transmembrane dII domain and the proton translocation pathway.

***a**, Conserved residues of the proton transfer pathway. Dowser-predicted water molecules are shown as red spheres. The distorted α-helical area is indicated as 3_{10} helix. **b**, Proton translocating pore features a narrow block close to the matrix (N) side and a dry stretch in the middle of the membrane. Four of the predicted water molecules lie within the central pore. **c**, Detailed view of the matrix proton entry site with conserved residues and hydrogen bonding network shown.*
Figure adapted from Kampjut & Sazanov, 2019.

Modelling of water molecules by Dowser revealed that both the N-side and P-side residues coordinate two water molecules each and that the channel between them is separated by a dry stretch. MD simulations of *T. thermophilus* dII (Padayatti *et al.*, 2017) showed that protonation of the histidine induces a transient wetting of the transmembrane proton channel between α_2H42 and βE221 (ovine H664 and E806) accompanied by a flip of βT214. The role

of the threonine could be played by the ovine homologue S799 or by S492, a homologue of *T. thermophilus* α_2H42, or by a non-conserved T495 (Figure 3.16b) (Padayatti *et al.*, 2017).

3.3.2. Conformation of the proton translocation channel depends on the dIII attachment

dIII domain forms a large interface with dII and inserts the helix 4 (residues 929-937) and the $NADP^+$-coordinating loop (residues 920-927) deeply between the key TM helices 3, 9 and 13, pushing them apart (Figure 3.17a). This induces a significant shift of the loops CL6/CL7, helix CL8 and the matrix-facing sides of TM9, 10, 11, 13 and 14 away the centre of the dII, with up to 3.5 Å displacements (Figure 3.17a).

Dramatically, when dIII attaches H664 flips "up" (towards N-side) and S799 "down" (towards P-side), in line with the MD predictions (Padayatti *et al.*, 2017) (Figure 3.17b). This affects the dimensions and chemical properties of the proton translocation channel. In the dIII-attached dII the channel starting from the N-side (i.e. facing bacterial cytoplasm or mitochondrial matrix) consists of a narrow block followed by a group of conserved polar residues (His664, Asn796 and Asn489) and a dry hydrophobic stretch. There is also a cluster of negatively charged residues at the exit, appropriate for proton "gathering" from the P-side in the forward reaction. Analysis of the structure with dIII attached failed to reveal any water-permeable channels leading towards the N-side. The nicotinamide-binding loop, $NADP^+$ and M485 very effectively plug the N-side opening of the proton channel. We consider this quite a definitive feature of the structure. dIII binding to dII may also allow temporary wetting of the dry patch in the middle of the channel (Figure 3.16c) and thus His664 protonation from the P-side.

With dIII detached, the proton channel adopts a more closed conformation from the P-side. dII from apoNNT appears similar to the dII from NNT-$NADP^+$, suggesting that binding of dIII is enough to open the channel, and that nucleotide binding to dIII is not important. Additional restriction points appear between H664 and the P-side (Figure 3.17e,f). However, from the N-side (matrix) the channel is more open in this case. Overall, in the dIII-detached conformation, H664 appears to be more accessible for protonation from the matrix side, partly due to the re-arrangements in the channel and partly due to the removal of the dIII "plug", exposing the protein surface close to the channel to the solvent.

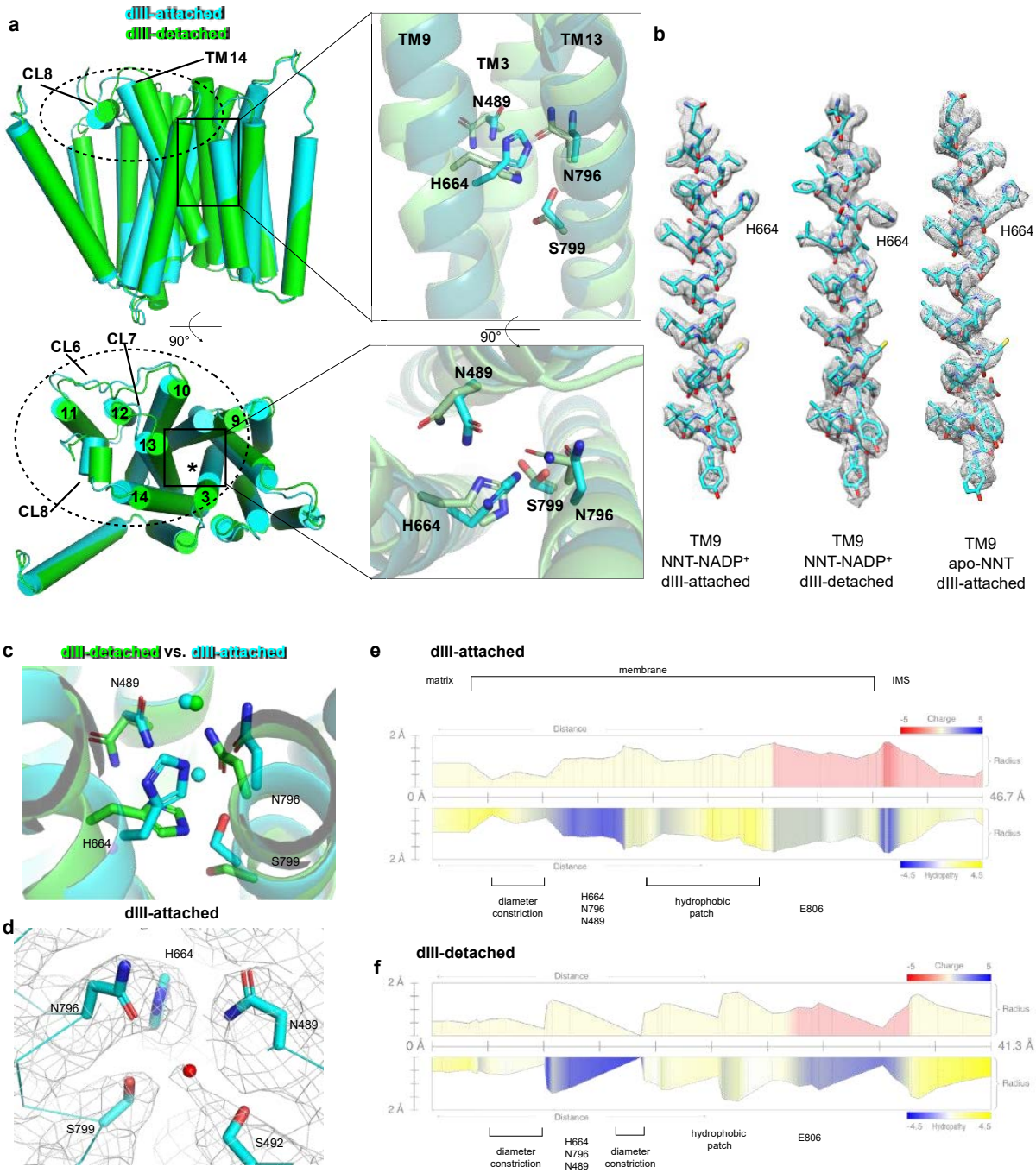


Figure 3.17. Comparison of proton translocation channel in dIII-attached and dIII-detached conformations.

***a**, Comparison of dII in dIII-associated (cyan) and dissociated (green) states from "single face-down" structure. dIII dissociation results in the tilting of TM helices towards the central axis. Expanded views show rearrangement of the N-side proton entry side in dII. Notably, H664 and S799 flip and the surrounding N489 and N796 change conformations significantly, opening the channel to the N-side when dIII is detached. **b**, Comparison of the TM9 density in the "double face-down", "single face-down" (dIII-detached monomer) and apo dII clearly displaying a H664 flip in the dIII-detached dII. **c**, Comparison of Dowser-predicted waters in the dIII-attached (cyan) and dIII-detached (green) dII. The dIII-attached structure has two water molecules, one below and one above the N-side proton gate while the dIII-detached structure only has one water molecule above the gate, consistent with the proposal that channel is open to P-side when dIII is attached. **d**, Density for a water molecule coordinated between H664 and S492, consistent with Dowser-predicted water is beginning to show in our cryo-EM density. **e**, Proton pathway profile (calculated in Mole 2.5) in dIII-attached dII reveals a diameter constriction between N-*

side and H664, a hydrophobic stretch between H664 and E806 and a negatively charged P-side proton entry site. ***f****, Proton pathway profile in the dIII-detached dII. Additional constriction appearing between H664 and P-side upon dIII detachment is indicated.*
Figure adapted from Kampjut & Sazanov, 2019.

Modelling of water molecules with Dowser (Zhang and Hermans, 1996) is consistent with this conclusion. In the dIII-attached dII one water molecule is bound above and one below the H664-N489-N796 ring while in the dIII-detached state, the P-side water molecule disappears and the S799 flips "downwards" (Figure 3.17c). Cryo-EM density of the highly coordinated P-side water molecule is visible in our 2.9 Å map (Figure 3.17d). Overall this shows that the part of the channel between H664 and the P-side is less hydrated when dIII is detached.

The conclusion is that the channel is open to N-side in the dIII-detached state and open to the P-side in the dIII-attached state. When dIII is detached, H664 is more accessible for protonation from the matrix side, both due to the re-arrangements in the channel and, importantly, the removal of the dIII-NADP(H) "plug", exposing the dII surface to the solvent.

3.3.3. Comparison of proton translocation machinery with *T. thermophilus* NNT

We noted that *T. thermophilus* dII resembles the dIII-detached ovine dII much more than the dIII-attached dII (Figure 3.18a,b). However, a direct comparison of the fine details of the proton channel between these two species is difficult because the protonatable histidines sit on different TM helices and at a different depth within the proton translocation pathway in the two species. This is because while all transhydrogenases have a single key protonatable histidine at the N-side of dII, the histidine sits either on TM3 (α_2H42 in *T. thermophilus*) or on TM9 in *E. coli* and mammals (βH91/H664) (Figure 3.18c).

The differences between the ovine and *T. thermophilus* dII are also in the P-side (intermembrane space) cavity, gated by a conserved residue E806 which in ovine dII is in an extended conformation in contrast to the *T. thermophilus* structures, with T674 and E678 allowing access to the P-side. Residues at the P-side of the proton transfer path are generally much less conserved between species than at the N-side. The P-side entry cavity in mammals also has a stronger negative charge, appropriate for the attraction of protons in the forward reaction.

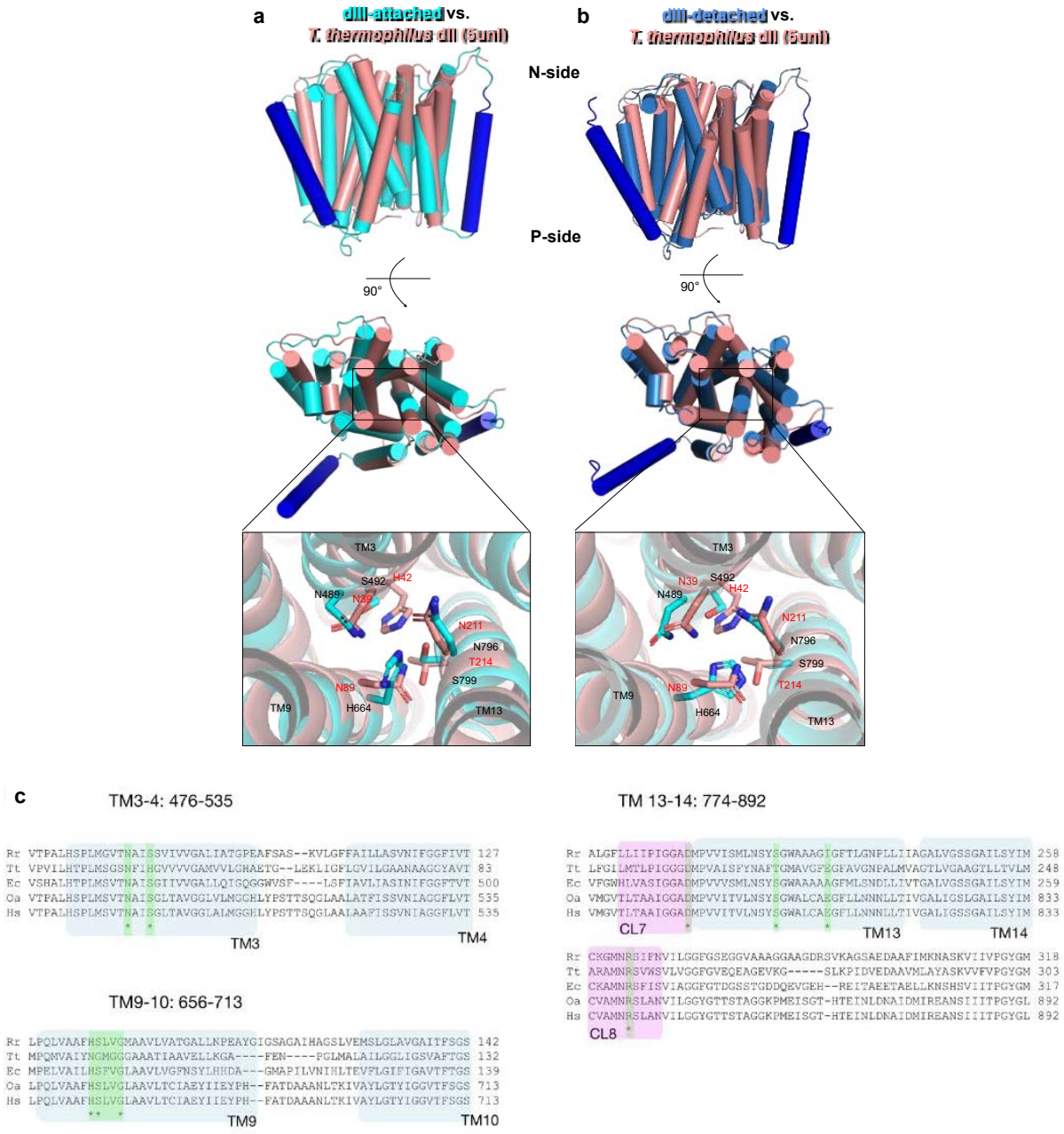


Figure 3.18. Comparison of ovine and T. thermophilus dII.

***a**, Comparison of T. thermophilus dII (salmon, PDB 5UNI) and "double face-down" (dIII attached) ovine dII (cyan, supernumerary TM1 and TM5 in blue). Residues in the N-side cavity display markedly different conformations. **b**, Comparison of T. thermophilus dII (salmon, PDB 5UNI) and "single face-down" (dIII detached monomer) ovine dII (blue). Residues in the N-side cavity match more closely as both of these dII structures are detached from dIII. **c**, Alignment of conserved residues on TM3, TM9, TM13, TM14 and CL7 and CL8 important for proton translocation and reaction coupling in T. thermophilus (Tt), R. rubrum (Rr), E. coli (Ec), O. aries (Oa) and H. sapiens (Hs). T. thermophilus and a handful of other species with an NGXGG motif on TM9 have the protonatable histidine on TM3 (2H42 in T. thermophilus) while most other species that share an HSXXG motif on TM9 have the protonatable histidine on TM9 (H91 in E. coli and H664 in O. aries).*
Adapted from Kampjut & Sazanov, 2019.

3.3.4. dII-dIII interface formation opens dIII nucleotide binding pocket

The overall conformation of dIII in our structures revealed important differences to the crystal structures of isolated dIII at the interface-forming sites and in the nucleotide binding-pocket (Prasad *et al.*, 1999; Mather *et al.*, 2004). These changes can be explained by the fact that in our structures, dIII was bound to dII, while in previous crystal structures dIII was solved either in isolation or in complex with dI. This gives us an unexpected insight into the nucleotide exchange process. dII-dIII interface formation in effect opens the nucleotide binding pocket, structurally explaining how the dIII can transition from "occluded" into "open" state.

In ovine NNT, helix 4 and loop D, which are conserved transhydrogenase-specific elements flanking the central Rossmann fold, formed extensive contacts with the cytoplasmic loops of the dII (Figure 3.19a). They also form contacts with dI domain in the face-up orientation as evident from the dI_2dIII heterotrimeric structures (Figure 3.13) (Cotton *et al.*, 2001; Leung *et al.*, 2015).

Loop E (residues 1002-1010) is another transhydrogenase-specific element that closes over the nucleotide to impede its exchange with solution. In all the previous dIII structures (isolated or in complex with dI), loop E was closed, indicating that dIII was in the "occluded" state (Prasad *et al.*, 1999; White *et al.*, 2000; Sundaresan *et al.*, 2003). In ovine dIII bound to dII, however, loop E was displaced by the dII CL7 and had only very weak density in our "double face-down" $NADP^+$ structure, which allowed NADP+ to insert deeply into the dII (Figure 3.19c). In the apo-NNT structure, loop E was completely disordered allowing free exchange of nucleotide (Figure 3.19d). Some opening was also observed around the ribose diphosphate moiety of $NADP^+$, namely the breaking of the R925-Y890 stacking interactions and the R925 coordination bonds with the diphosphate which have so far not been observed in any known dIII or dI_2dIII crystal structures and must be induced by dII binding (Figure 3.19b) (Prasad *et al.*, 1999; White *et al.*, 2000; Sundaresan *et al.*, 2003). In apo-NNT, R925 was rotated outwards into the solvent.

These changes strongly suggest that dIII bound to dII should be considered in the "open" state. In the absence of nucleotide dIII appeared even more open and disordered than in the presence of $NADP^+$, because $NADP^+$ was partially occluded by the dII loops in the face-down binding site, but far less so than in the detached dIII. This arrangement and the disorder of the loop E strongly suggest that dIII-dII interaction is important for opening of the dIII and nucleotide exchange.

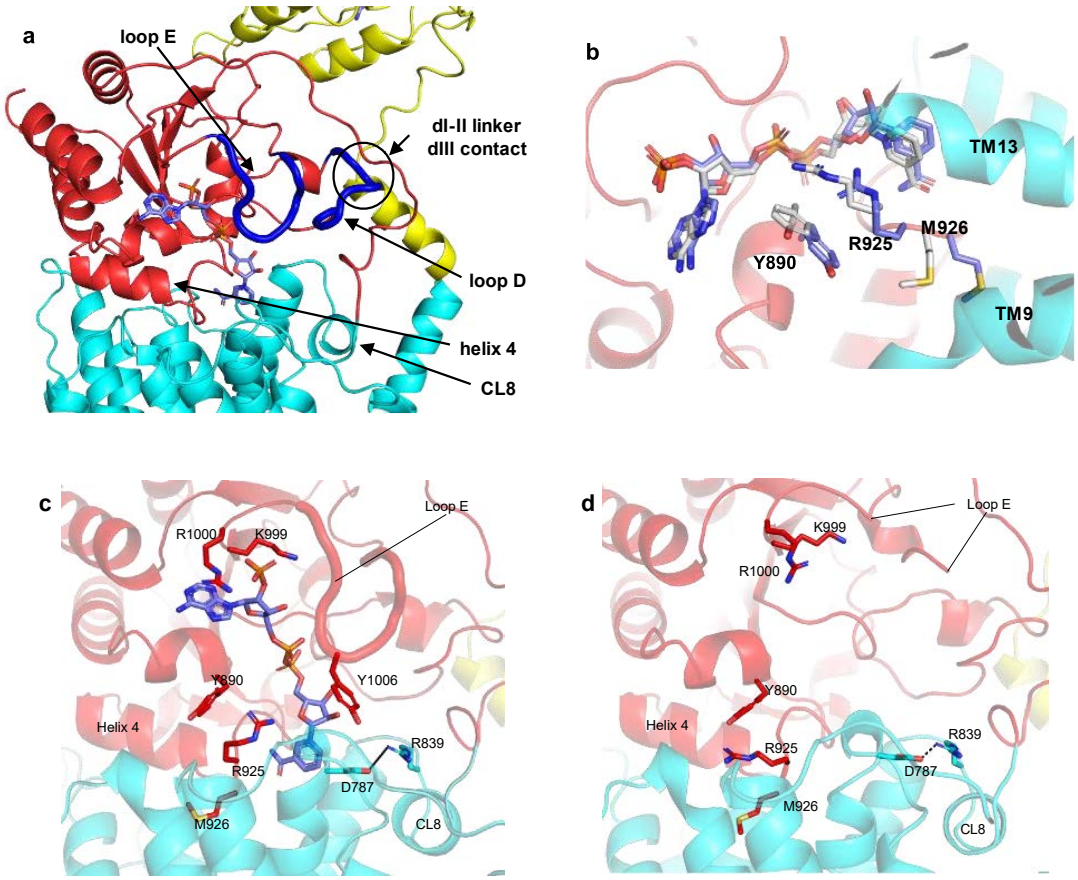


Figure 3.19. dIII attachment to dII opens the NADP(H) binding pocket.

***a**, The interface between dIII and dII. Helix 4 and loop D, which are conserved transhydrogenase-specific elements flanking the central Rossmann fold, form extensive contacts with the cytoplasmic loops of the dII. Loop E is displaced by dII binding. **b**, NADP(H) binding pocket comparison between ovine dIII-dII and human dIII-NADP+ (PDB 1djl, in grey) reveals breaking of R925 interactions with diphosphate and Y890. **c**, NADP(H) binding site in "double face-down"-NNT. **d**, NADP(H) binding pocket in apo-NNT. Loop E, K999 and R1000 are disordered and R925 flips into an outward-facing orientation, opening the site to the solvent*
Figure adapted from Kampjut & Sazanov, 2019.

3.3.5. Electrostatic interaction between H664 and dinucleotide controls dIII attachment

To ensure the entire catalytic cycle to take place, dIII must sequentially associate and dissociate from dII and this process needs to be meaningfully coupled to proton transfer to ensure coupling between the redox and proton pumping reactions. Protonatable H664 and the nicotinamide ring of the dinucleotide bound in the dIII come to within 13 Å when dIII is bound to dII (Figure 3.20). This is much closer than 19 Å predicted previously (Leung *et al.*, 2015). The short distance between H664 and nicotinamide immediately suggests that the electrostatic interaction between them directly influences the strength of the dII-dIII interaction.

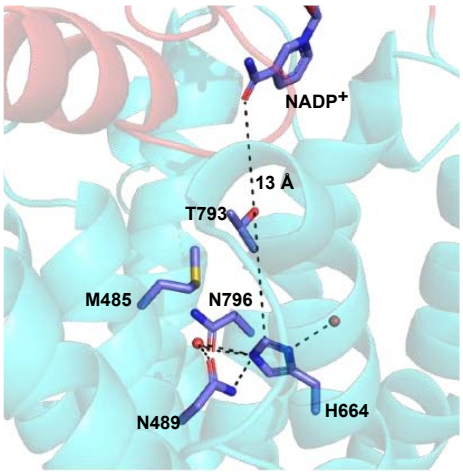


Figure 3.20. Nicotinamide ring is bound 13 Å away from H664.

Their electrostatic interaction influences the strength of the dII-dIII interface.
Adapted from Kampjut & Sazanov, 2019.

The effect of a Coulombic interaction of a unit charge at a distance r can be estimated as $\Delta G = 332\ \varepsilon^{-1}\ r^{-1}$ kcal mol^{-1}. Dielectric constant for the protein is $\varepsilon = 4$. Unit charge (protonated histidine) at 13 Å distance from nicotinamide thus contributes ~6 kcal/mol, which is a whole order of magnitude more than thermal fluctuations, which are on the order of a few k_BTs (1 k_BT at T = 310 K is ~0.616 kcal/mol). This proposal is therefore thermodynamically feasible, despite the fact that the dIII-dII interface is relatively large and strong as shown above.

This proposal is also experimentally supported by the phenotypes of mutants of H664 and R544. Mutations of the conserved H664 to arginine/lysine, thereby introducing a permanent positive charge within dII, lead to the dissociated conformation mimicking our NADPH structures. They also exhibit an elevated cyclic reaction rate (suggesting dIII spends more time dissociated from dII and being in the occluded state, preventing nucleotide exchange). When dII-dIII interface is destabilised by mutating R544, disrupting its salt bridge with D942, we observe the same phenotype. Change of a single charge at the dII-dIII interface drastically influences its stability, giving strong support for the hypothesis that dIII attachment and detachment is regulated by electrostatic interactions between H664 and nicotinamide.

The electrostatic interactions can also be paraphrased in terms of changes of pK_a. It has been shown that mutations can change pK_a of residues which are 10 Å away by up to two units (Tynan-Connolly and Nielsen, 2006). dIII-NADP(H) attachment modifies the pK_a of H664. Bound NADPH promotes its H664 protonation and $NADP^+$ promotes H664 de-protonation.

pKa values of protein residues are notoriously difficult to predict. The only reference to the H664 pKa in the literature was predicted to be 5.5-6 based on pH profile of cyclic reaction (Sazanov and Jackson, 1995). This is consistent with ProPka calculations on our structures (suggesting pK_a ~5). It would also allow H664 to exist in a protonated state in sufficiently large proportion of molecules to allow reverse reaction. However, dII-dIII interface itself may be sensitive to pH changes especially far away from the optimum pH values, regardless of histidine protonation state. Since our datasets were collected at pH 7.4, when H664 is mostly de-protonated, we observed mostly attached dIII-$NADP^+$ and mostly detached dIII-NADPH, consistent with this proposal. dIII detachment from dII in the presence of $NADP^+$ and $H664^+$ is also validated by low pH-stimulation of $NADP^+$-dependent cyclic reaction (hydride transfer from NADH to NAD^+ via occluded $NADP^+$) (Hutton *et al.*, 1994; Sazanov and Jackson, 1995).

This is further validated by our data where some dIII-$NADP^+$ were dissociated compared to virtually none of apo-dIII being dissociated. In the apo state, dIII can bind to dII regardless of the H664 protonation status, while dIII-$NADP^+$ cannot bind to dII if H664 is protonated.

Based on the structural data presented here, supported by previous biochemical work we conclude that the H664 and nicotinamide interact electrostatically. Protonated H664 interacts unfavourably with $NADP^+$, while neutral H664 interacts unfavourably with NADPH. dIII-$NADP^+$ is detached from protonated $H664^+$ but is attached to the neutral H664. Vice versa, dIII-NADPH is detached from H664 and is attached to $H664^+$. Thus, both NADP(H) binding and histidine protonation/deprotonation (depending on the direction of the reaction) are necessary to induce dIII detachment from dII.

3.4.Mechanism of NNT

3.4.1. Principles

On the basis of the structural observations and previous functional and mutagenesis results we proposed a revised catalytic mechanism of transhydrogenase, involving dIII flipping as proposed earlier (Jackson *et al.*, 2015). For the sake of clarity I will outline the principles of NNT reaction that have been proposed earlier and those that have been proposed or confirmed based on the structural work described in this thesis.

Previous work established the following:

a) dI remains open to nucleotide exchange throughout the reaction cycle. This is confirmed by high rates of binding of NAD(H) to dI during the catalytic cycle (Yamaguchi and Hatefi, 1997; Venning *et al.*, 2000).

b) The dI-dIII interface is transient, hydride transfer is fast and does not limit the reaction (Venning *et al.*, 1998).

c) When dIII is detached, bound NADP(H) is occluded, as shown by a decreased K_d for NADP(H) in isolated dIII (~10^{-9} M) compared to the intact enzyme (~10^{-6} M), copurification of bound nucleotide with isolated dIII and well-ordered and closed loop E in dIII structures (Fjellström, Johansson and Rydström, 1997; Prasad *et al.*, 1999; Sundaresan *et al.*, 2003; Bizouarn *et al.*, 2005). H664 and R544 mutations which lead to dIII detachment also lead to co-purification with NADP(H) (Glavas and Bragg, 1995; Glavas, Hou and Bragg, 1995).

The following principle was speculatively proposed earlier (Jackson *et al.*, 2015) and is now validated by our structures:

d) The formation of the dII-dIII interface is necessary for nucleotide exchange, as shown by our structures demonstrating loop E opening. It is further validated by R544 mutants which disrupt the dII-dIII interface (Glavas and Bragg, 1995).

Crucial new insights from the structural data presented here are:

e) Apo-dIII is attached to dII (regardless of the protonation state of His664) and can detach only when bound to NADP(H).

f) The proton channel is open only to the N-side when dIII is detached and only to the P-side when dIII is attached.

g) The dIII-NADP(H) detachment from dII depends on the protonation state of the key histidine in the proton channel (H664 in ovine).

On the basis of these principles we propose a robust and elegant mechanism which explains the full reversibility of NNT, pH dependencies of reverse and cyclic reactions, effects of $NADP^+$ and NADPH on the NNT conformation and is validated by the phenotypes of various mutants, particularly within the proton translocation pathway, E loop and dII-dIII interface. Proton translocation and nucleotide exchange are tightly coupled by direct interactions between H664 and NADP(H): attachment/detachment of NADP(H)-bound dIII is strictly dependent on the protonation state of H664 and opens the proton channel to the opposite sides of the membrane. "Slipping", or uncoupling, is prevented because proton translocation cannot happen without nucleotide exchange.

3.4.2. Forward transhydrogenation reaction

The forward reaction occurs under conditions with high pmf and excess of $NADP^+$. With dIII detached, H664 is unprotonated as it is exposed to the matrix with its relatively high pH (step 1, Figure 3.21). After hydride transfer, dIII-NADPH swivels down and attaches to dII-H664, opening it to the P-side (step 2, Figure 3.21). This allows H664 protonation from the P-side, where pH is lower (and His pK_a is shifted higher due to interaction with NADPH). Nucleotide exchange follows, due to $NADP^+$ excess and opened dIII (step 3, Figure 3.21). dIII-$NADP^+$ detaches from dII-$H664^+$, which opens dII to the N-side where $H664^+$ is de-protonated (step 4, Figure 3.21). dIII-$NADP^+$ associates with dI, allowing for hydride transfer (back to step 1). As described below, all the steps can be easily reversed in the appropriate conditions (high NADPH and low pmf), giving rise to the reverse reaction.

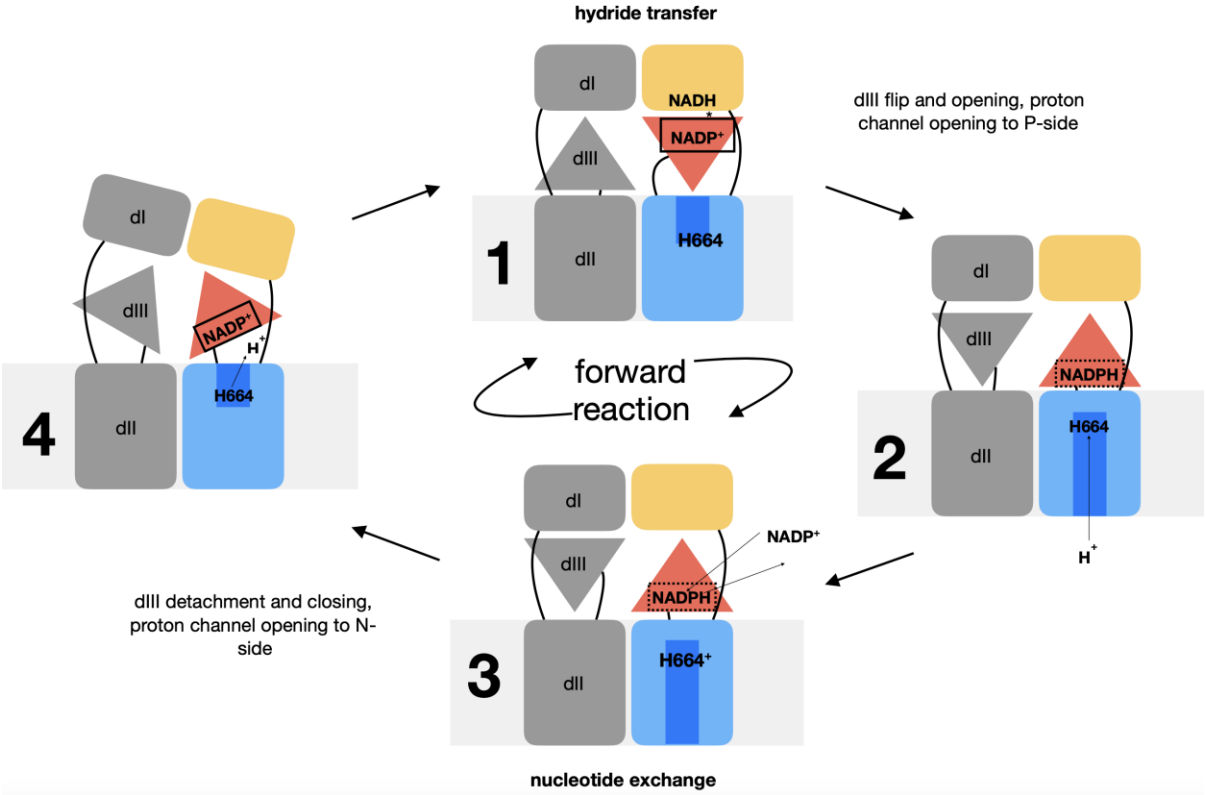


Figure 3.21. The catalytic mechanism of transhydrogenase

*NNT always works as a dimer. Key steps in the mechanism for the monomer in colour are described in the main text. The other monomer (grey) works in the anti-phase fashion. Key check-point in the mechanism is that proton translocation across the membrane cannot proceed to completion without nucleotide exchange in step 3. The dIII-NADP(H) and dII-H664 interactions, ensuring this coupled reaction, can be summarised as follows, with (*) meaning strong interaction and (<->) weak interaction or repulsion: $NADP^+$ * H; $NADP^+$ <-> H^+; NADPH * H^+; NADPH <-> H. Reverse transhydrogenation reaction consumes NADPH and NAD+ and results in proton pumping, supporting proton motive force (pmf). The driving forces for this reaction are the nucleotide ratios and pmf which promotes protonation of H664 from the matrix side.*
Figure adapted from Kampjut & Sazanov, 2019.

3.4.3. Comparison to the Jackson mechanism

This mechanism is quite different, and in many ways simpler, to the previously proposed Jackson mechanism (Figure 1.8). Firstly, the Jackson model assumed that dIII exists face-down in either occluded or open state depending on H664 protonation and nucleotide bound to dIII, whereas we show that essentially any dIII in face-down conformation is open. Secondly, the Jackson model assumes that dII can exist in no less than four different states: open to N-side, open to P-side and occluded with or without the proton bound, whereas we observe only two states of dII, depending on the dIII attachment, which satisfactorily explain the mechanism. Thirdly, the Jackson model postulated that $NADP^+$- or NADPH-bound dIII opens the channel specifically either to the P- or N-side, respectively. This is not consistent

with our data and mechanism, which does not require such specificity and operates on simpler principles. Finally, Jackson model does not explain how exactly nucleotide exchange happens only in step 1 and not between steps 2 and 3 or 5 and 6.

The Jackson mechanism also did not explain why NADPH and $NADP^+$ have different effects on the enzyme (as demonstrated by trypsinolysis and C893 labelling experiments). According to the Jackson mechanism, NADPH and $NADP^+$ should both stabilize the same, purely asymmetric state. This difference is now fully accounted for by the new proposal as the overall conformation of NNT depends on the protonation state of H664 and the nucleotide bound in dIII.

3.4.4. Reversibility of the proposed mechanism

In the reverse reaction, when pmf is low and there is an excess of NADPH, all the steps are naturally reversed, resulting in proton pumping into the P-side (Figure 3.22). After hydride transfer, dIII-$NADP^+$ is dislodged by an incoming dIII from the opposing monomer and is attached to dII with protonated $H664^+$ (as pH is now lower in the matrix) (steps 1-2). This opens the channel to the P-side and H664 loses its proton there, as its pK_a is reduced due to interaction with $NADP^+$ (step 2). Nucleotide exchange follows due to the excess of NADPH (step 3) and dIII-NADPH swivels "up", due to its low affinity to neutral H664, back to step 1. In both forward and reverse reactions, NAD^+/NADH formed in dI in step 1 would be exchanged for NADH/NAD^+ before step 4 completes, so that the cycle can resume, since dI remains open throughout.

In the reverse direction, due to the excess of NAPDH, dIII is mostly detached from dII and the channel is open towards the N-side. Once a hydride transfer event takes place, dIII-$NADP^+$ binds transiently to the dII and promotes NADP+ dissociation, followed by stronger binding of apo-dIII, H664 de-protonation into the P-side, which is facilitated by a cluster of negatively charged residues around the P-side opening. dII-dIII interface is broken by binding of NADPH, which causes dissociation of dIII and restarting of the cycle. Slipping in this direction is prevented by the fact that nucleotide exchange will only be stimulated when H664 is protonated.

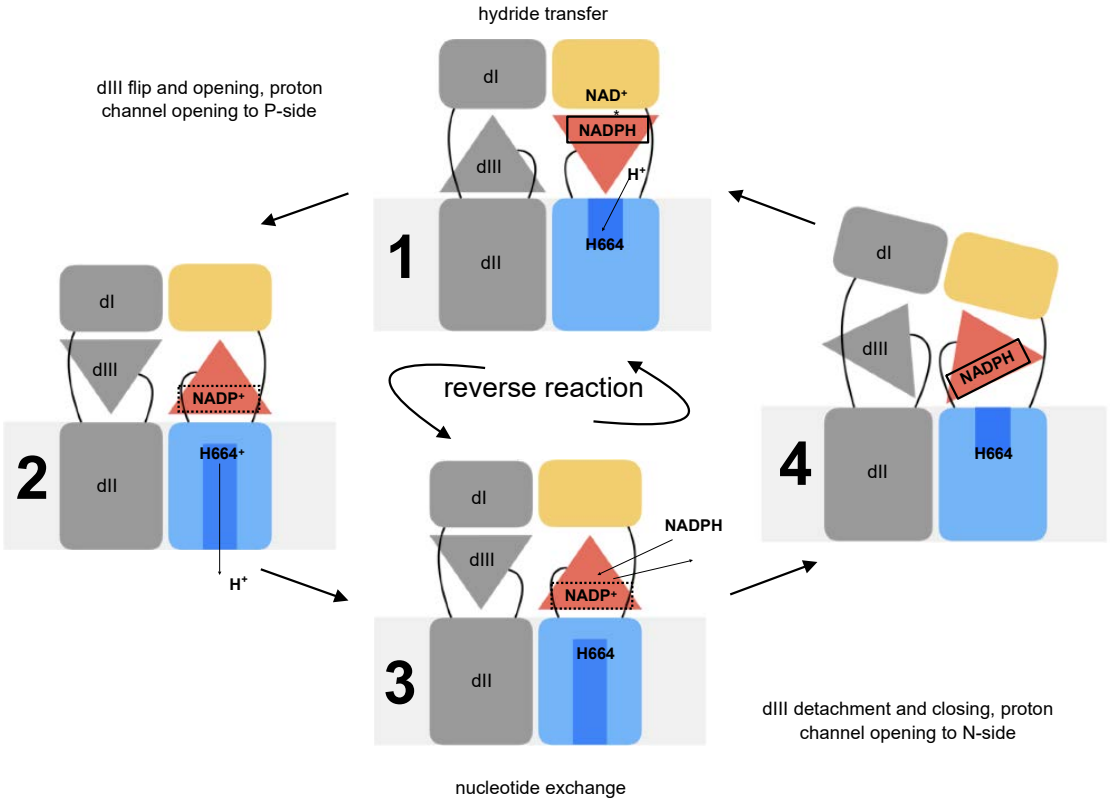


Figure 3.22. Reverse transhydrogenation.

Adapted from Kampjut & Sazanov, 2019.

3.4.5. Comparison of reverse and forward transhydrogenation

In the forward direction, nucleotide exchange is stimulated by displacement of loop E from dIII-NADPH. In the reverse direction, nucleotide exchange is stimulated by unfavourable interaction between protonated H664 and NADP+ bound to dIII. This ensures that hydride transfer is always coupled to proton pumping, because nucleotide exchange will not happen unless H664 is protonated from N-side first.

dIII detachment in the forward direction is stimulated by unfavourable interaction between protonated H664 (from P-side under the influence of PMF) and NADP+ bound to dIII. In the reverse direction, dIII detachment is stimulated by binding of NADPH to dIII (as supported by our structural evidence).

These notions explain why the reverse and forward reactions are coupled to proton translocation in different directions. In the reverse direction, H664 is protonated from N-side and it is relatively close to the surface of the membrane. Deprotonation then involves the loss of the proton into the P-side across a relatively large depth of the membrane. Locally, this is stimulated by negatively charged residues at the P-side entry. In the forward direction,

protonation of H664 would have to happen from the P-side and cross a large proportion of the membrane depth. This is clearly only favourable when there is an already pre-existing pmf across the membrane.

In the forward direction, dIII-$NADP^+$ is mostly associated with dII and the channel is open towards the P-side (Figure 3.21, step 1). In the reverse direction, dIII-NADPH is mostly dissociated from dII and the channel is open towards the N-side. Since NADPH promotes dIII detachment and $NADP^+$ promotes dIII attachment to dII, the relative concentration of $NADP^+$/NADPH determines which side of the membrane the proton channel will spend most time open to. This is of course only valid at neutral pH when H664 is mainly deprotonated.

Under typical conditions, forward reaction proceeds in the presence of high $NADP^+$ and high pmf, which both ensure H664 protonation from the P-side. Conversely, reverse reaction happens at high NADPH and low pmf, when H664 protonation is more likely from the N-side. Excess of NADPH at zero starting pmf will drive build-up of pmf until the equilibrium is reached.

3.4.6. Preventing slipping

Ensuring a correct stoichiometry between hydride transfer and proton transfer requires a tight coupling between hydride transfer, nucleotide exchange and proton translocation. "Slipping", or uncoupling, at any step is prevented because proton translocation cannot happen without nucleotide exchange. For example, step 3 (Figure 3.21) cannot progress to step 4 and switch proton channel access unless there is nucleotide exchange – only the dIII-$NADP^+$ and $H664^+$ combination will allow dIII dissociation and dII opening to the N-side. If nucleotide exchange does not happen, bound dIII-NADPH will interact strongly with $H664^+$ and prevent dIII detachment and dII opening to the N-side. In other words, two half-paths of proton across the membrane (up to H664 and after) in both forward and reverse reactions are connected only by nucleotide exchange, excluding any uncoupling.

The proposed mechanism ensures tight coupling also close to the equilibrium conditions because our mechanism does not assume certain protonation state of H664 in order for the reaction to proceed. The figures just illustrate the most abundant species under typical conditions of forward and reverse reactions. However, we can simply consider all the possibilities. After hydride transfer in step 1, dIII associates with dII, allowing nucleotide exchange. Subsequent to nucleotide exchange possible scenarios will be as follows:

1. If it is dIII-NADP+ and $H664^+$ (weak dIII-dII interaction), then dIII would be flipped up first, then dII is closed from periplasm and opened to cytoplasm, so that proton is released into the cytoplasm – this is forward reaction, when there is excess of $NADP^+$.

2. If it is dIII-$NADP^+$ and H664 (strong dIII-dII interaction), dIII will remain attached to dII until H664 is protonated from the periplasm. This prevents slipping of forward reaction. After H664 protonation dIII will be flipped up and proton released into the cytoplasm, in the forward reaction as above.

3. If it is dIII-NADPH and H664 (weak dIII-dII interaction), dIII will flip up first, then dII will close to the periplasm and open to cytoplasm, allowing H664 protonation from the cytoplasm - this is reverse reaction, when there is excess of NADPH.

4. If it is dIII-NADPH and $H664^+$ (strong dIII-dII interaction), dIII will remain attached to dII until H664 is de-protonated into the periplasm. This prevents slipping of reverse reaction. After H664 de-protonation, dIII can flip up and dII will close to the periplasm and open to cytoplasm, allowing H664 protonation from the cytoplasm, in the reverse reaction as above.

5. In the apo state, dIII remains attached to dII, preventing proton leak.

Thus, we achieve tight coupling between hydride transfer and proton translocation, without any slipping, at all conditions – either favourable for forward or reverse reaction, or close to equilibrium. Close to equilibrium the overall net direction of reaction will depend on the pmf value and nucleotide ratios. It is known that high pmf can drive ($NADPH/NADP^+$) / ($NADH/NAD^+$) ratio up to 500 from about 0.8 in the absence of pmf.

3.4.7. Driving force for dIII dissociation from dI

After a full cycle and hydride transfer back at step 1, subsequent dIII dissociation from dI in step 2 is unlikely to be initiated by hydride transfer itself as it is a fast step and dI_2dIII structures were solved with many combinations of bound nucleotides (Cotton *et al.*, 2001; Mather *et al.*, 2004). However, NNT always operates as a dimer, in which the two dIII domains cannot simultaneously attach to the dI dimer, as noted above. Therefore, as the second monomer (in grey in Figure 3.21) goes through a similar cycle as the first monomer (in colour), the second dIII, swivelling up after detachment from dII, will clash with the already bound dIII from the first monomer and dislodge it, initiating its flip “down” in step 2, resuming the cycle. As shown in the literature, monomeric dI cannot catalyse reverse transhydrogenation reaction. Presumably, detachment of dIII-NADP+ from dII causes conformational changes in the other

monomer which will prevent dIII-NADP+ to reassociate with dII after proton has been lost into the N-side.

3.5.Validating biochemical data

3.5.1. Trypsinolysis at different pH shows opposing effects of NADP and NADPH

It has been demonstrated before that NADPH and $NADP^+$ have different effects on the trypsinolysis rate of dI-dII linker (Yamaguchi, Wakabayashi and Hatefi, 1990). This can be explained by our structural data which show that trypsinolysis of the dI-dII linker is possible only in the dIII-detached conformation. In the dIII-attached conformation dI-dII linker contacts dIII and is likely inaccessible to trypsin active site (Figure 3.13). To validate the proposal that dIII detachment depends not only on the NADP(H) binding but also on the H664 protonation state, we analysed the full pH dependence of trypsinolysis in the presence of all substrates (Figure 3.23). These data show that NADPH stimulates trypsinolysis with increasing pH, while $NADP^+$ stimulates it with decreasing pH. This validates the predicted H664-NADP(H) interactions.

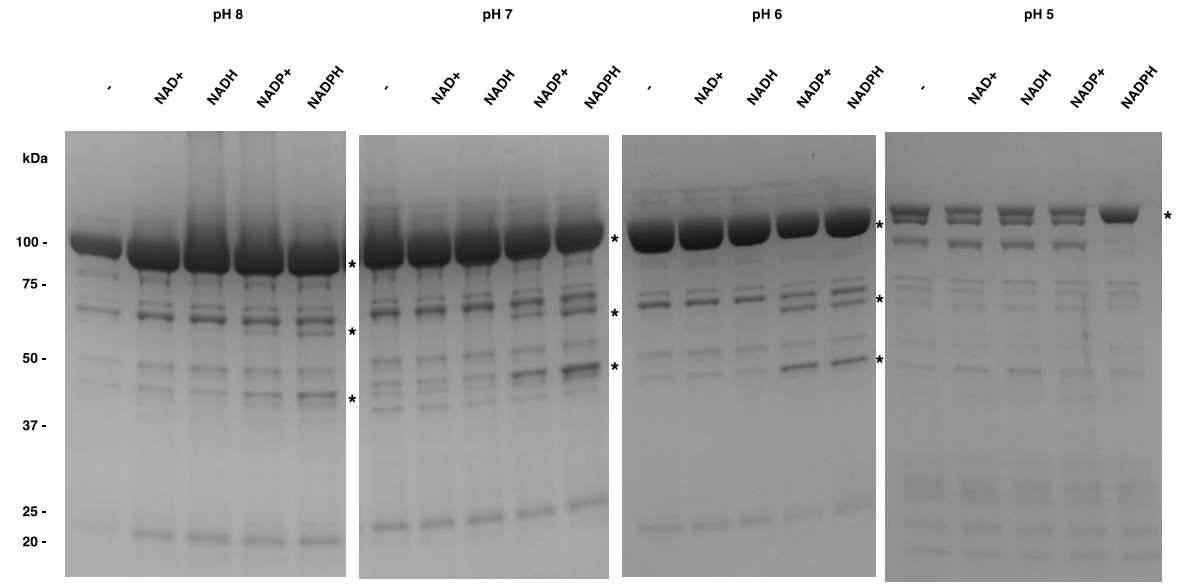


Figure 3.23. Trypsinolysis of NNT at different pH in the presence of substrates.

Trypsinolysis of ovine NNT confirms the interdependence between H664 protonation state, NADP(H) bound in dIII and dIII attachment to dII. 110 kDa full NNT and previously identified 66 and 43 kDa fragments are labelled with asterisks. As pH decreases from 8 to 6, NADPH- induced proteolysis diminishes relative to that induced by NADP+. At pH 5 trypsin produces different fragments, but stabilisation of the intact NNT by NADPH is evident. Trypsinolysis was performed independently three times with similar results. NNT purification was repeated independently ten times with similar results as shown in panels a-b.
Figure adapted from Kampjut & Sazanov, 2019.

3.5.2. Cyclic reaction

It has also been known for a long time that the pH dependence of the cyclic reaction is shifted towards the acidic range compared to the reverse reaction. Cyclic reaction has a pH optimum at around 5.5-6, very different from the pH optimum for reverse transhydrogenation, which is around pH 7 (Bizouarn and Jackson, 1993; Sazanov and Jackson, 1995). In the light of our structural data, it is easy to see that cyclic transhydrogenation will proceed the fastest when dIII spends as much time as possible dissociated from dII and thus in the occluded state. This will prevent loss of nucleotide from its binding pocket and reduce the time the enzyme spends in the "waiting" face-down conformation. The dII-dIII interface will be the weakest when pH is near pK_a of H664, as that's when dIII with *either* $NADP^+$ or NADPH will be the most dissociated.

Cyclic reaction therefore cannot inform us on specific $NADP^+$ vs NADPH effects as the main species in the NADPH-initiated cycle is still the occluded dIII-$NADP^+$ (Hutton *et al.*, 1994). The authors of the study above concluded that the release of either $NADP^+$ or NADPH decreases the pK_a of the key group in the enzyme based on their measurements of NADP(H)-linked cyclic transhydrogenation. However, since the effects of either $NADP^+$ or NADPH on the cyclic reaction would thus inevitably be the same, this conclusion does not contradict the proposed mechanism of NNT described here. Disentangling specific effects of NADPH and $NADP^+$ on the protein conformation can be done using trypsinolysis as described above.

3.5.3. Labelling of C893

It has been established previously that NADPH, but not $NADP^+$, increases the susceptibility of C893 to chemical labelling (Yamaguchi and Hatefi, 1989). Since C893 is buried in the dII-dIII interface, its accessibility to chemical modifiers is an independent measure of dIII attached/detached ratios. In chapter 3.1 I show that ovine NNT is inhibited by oxidised glutathione, the mechanism of which could be glutathionylation of the conserved C893. Because the inhibition depends strongly on the pH and redox potential it has so far not been possible to establish whether pH and different substrates (NAD(P)H or $NAD(P)^+$) modulate glutathione inhibition appreciably.

3.5.4. Residue conservation

An additional piece of evidence for the proposed mechanism where the interfaces and the proton pumping pathway play essential roles is the conservation of residues across the

organisms. When conservation scores are mapped on a model of transhydrogenase it becomes clear that the proton translocation pathway and the domain interfaces are the most conserved across the species, suggesting that they are indeed functionally significant (Figure 3.24).

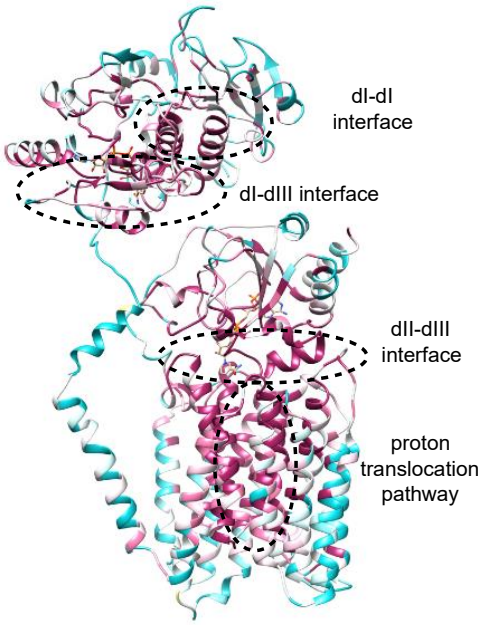


Figure 3.24. Conservation of transhydrogenase residues.

Residue conservation scores (calculated in Consurf, coloured cyan to magenta from low to high conservation) mapped on the structure of a single monomer of transhydrogenase. The most highly conserved regions are the proton translocation pathway, nucleotide binding sites and dI-dIII, dII-dIII and dI-dI interfaces. Adapted from Kampjut & Sazanov, 2019.

3.5.5. Review of available mutagenesis data

There have been over 300 different point mutants of *E. coli, R. rubrum* and human NNT described in the literature, which have been collected and summarized elsewhere (Kampjut and Sazanov, 2019). Here I am going to discuss a few of the key mutants which support the proposed mechanism (Figure 3.25).

The fact that dIII detachment depends on the electrostatic interaction between H664 and nicotinamide has been demonstrated by mutants of H664 and other nearby residues involved in proton translocation (N796, N489, S799). The key protonatable H664 has been mutated to 16 different amino acids. The effects of this mutation on the reverse, cyclic and proton-pumping activity have been measured and replicated in up to four different laboratories. The proposal that dIII-NADP+ can stably bind to dII with H664 in neutral state and protonation of H664 causes dIII-NADP+ detachment has been validated by H664K/R mutants in *E. coli* (Glavas, Hou and Bragg, 1995). Positively charged K/R in place of histidine mimic a

protonated histidine and as predicted by our mechanism, binding of dIII to dII should be severely diminished, resulting in detached dIII with tightly bound nucleotide. NNT was trapped in the NADP(H)-induced conformation with nucleotide bound to dIII, i.e. with both dIII detached from dII. The mutant had inhibited reverse activity but increased cyclic activity. In H664N/S/T mutants, however, dIII is always attached because N/S/T are not protonatable. In this case both the cyclic and reverse activities were inhibited and the enzyme could not adopt the NADP(H)-induced conformation, i.e. was trapped with dIII attached to dII (Glavas, Hou and Bragg, 1995). This pattern has been replicated with other residues in the vicinity of H664: N796 and T709. Mutating these residues to positively charged residues locked NNT in the dIII-detached state (Bragg and Hou, 2001; Yamaguchi, Stout and Hatefi, 2002; Yamaguchi and Stout, 2003). Additional complication is that sometimes mutations to large and neutral residues such as phenylalanine or tryptophan also stabilised the dIII-detached state, such as in the case of G826, but this can be attributed to the fact that insertions of large residues in the finely-tuned dII environment results in large scale conformational changes of the whole domain and its interaction surface with dIII.

The fact that dII-dIII interface formation is important for nucleotide exchange has been elegantly validated by mutations that affect the dII-dIII interface but are not involved in proton translocation (R544, R839, D787). The effects of these mutations replicated the above-described effects of H664K/R mutations and stabilised the enzyme in the dIII-detached conformation, as demonstrated by co-purification with NADP(H) and increased cyclic activity (Glavas, Hou and Bragg, 1995; Yamaguchi and Hatefi, 1995; Althage, Bizouarn and Rydström, 2001).

The fact that dII-dIII interface formation results in nucleotide exchange in dIII subunit and that this is connected to proton pumping has also been validated by mutations of loop E residues. In full transhydrogenase, loop E mutations result in decoupled phenotype, as predicted by the mechanism described here (Olausson *et al.*, 1993; Bergkvist *et al.*, 2000). In dI-dIII systems loop E mutations, however, increase the rates of reaction, as NADP(H) release from dIII becomes less limiting for the overall reaction. This validates the proposal that dIII is occluded when shuttling between dI and dII subunits, however, this has been known also from purification and structures of isolated dIII from several species. In all of them, NADP(H) is tightly bound and loop E is closed over the nucleotide (Prasad *et al.*, 1999).

Finally, mutations that disrupt the dimer by interfering with the dI-dI interface lead to a full inhibition of all the activities (Diggle *et al.*, 1996; Bizouarn *et al.*, 1997; Obiozo *et al.*,

2007). This confirms the fact that transhydrogenase must exist as a dimer, either for stability, but more likely for the catalytic reasons.

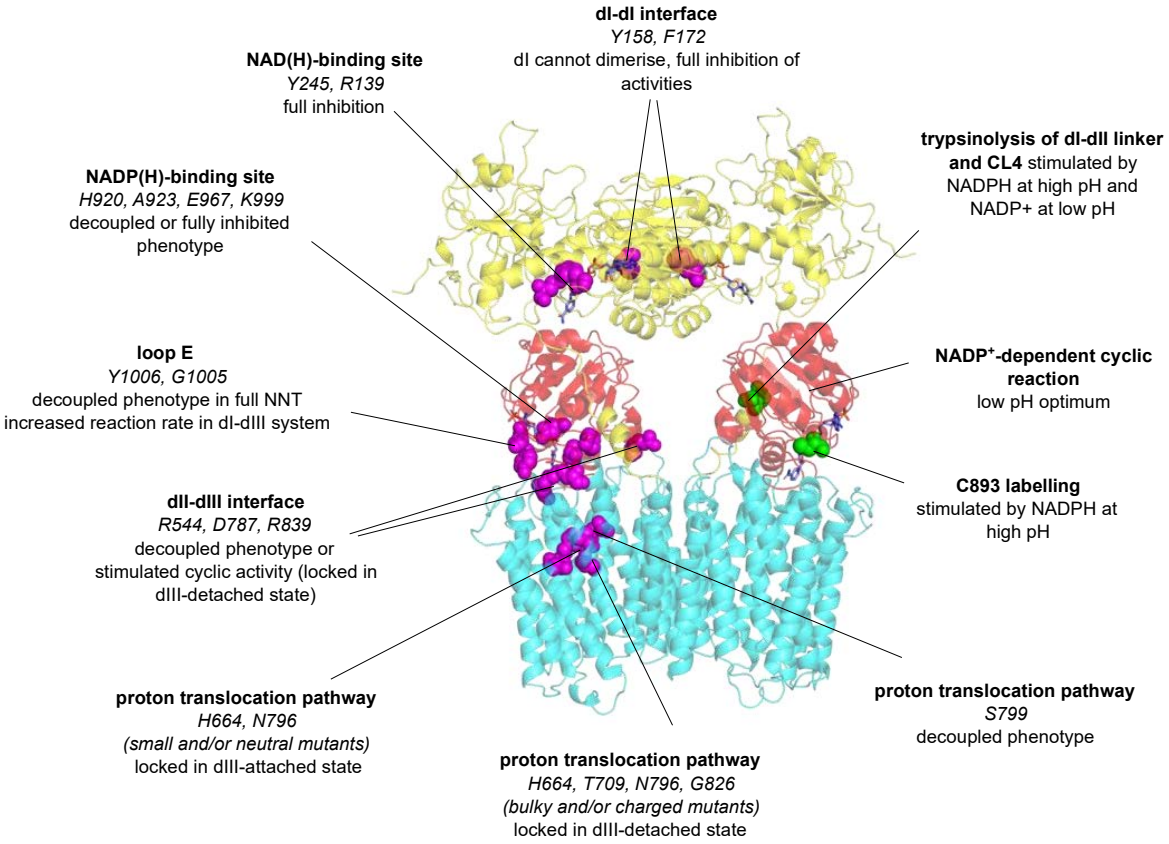


Figure 3.25. Summary of previously published biochemical evidence and mutagenesis data validating the proposed mechanism.

Adapted from Kampjut & Sazanov, 2019.

3.6.Discussion

3.6.1. Role of the dimer

Our revised mechanism implies anti-phase action of two monomers, or "division of labour" as suggested previously (Jackson *et al.*, 2015). Such a mechanism is absolutely necessary for bacterial enzymes without the dI-dII linker, where one dIII has to be attached to dI_2 at all times to prevent the loss of dI_2. Based on our mechanism and structures, it would seem that a strict out-of-phase operation of NNT is not obligatory since all the catalytic steps apart from the hydride transfer could in principle proceed simultaneously on both sides of the dimer. Empirically, however, mammalian NNT also demonstrates half-of-sites inhibition,

catalytically inactive monomeric dI and negative cooperativity of NAD(H) and NADP(H) sites between monomers, strongly suggesting an anti-phase mechanism (Phelps and Hatefi, 1984; Yamaguchi and Hatefi, 1993; Obiozo *et al.*, 2007). Even though the presence of dI-dII linker could make a strict anti-phase action less of a requirement for the mammalian enzyme, there are probably other advantages to operating an antiphase mechanism. Firstly, the coordinated anti-phase action of the two monomers would avoid steric clashes between independently functioning dIIIs. Secondly, swivelling "up" of one dIII is the best explanation for the initiation of the detachment, after hydride transfer, of the other dIII from dI_2 (Step 2, Figure 3.21). Thirdly, operating as a dimer in an antiphase mode could potentially also prevent slipping (decoupling of proton pumping from hydride transfer) as hydride and proton transfer are not only coupled within one protomer but also between the two protomers. All of these reasons probably allow the dimer to secure a kinetic advantage over the monomer (Obiozo *et al.*, 2007; Jackson *et al.*, 2015). Nevertheless, further data will be required to determine the significance of the anti-phase action in NNT as well as to test all the implications of the proposed mechanism.

3.6.2. Communication between the two protomers

Our structures suggest different ways in which the two monomers can communicate with each other to structurally explain the basis of anti-phase dimer action. As described above, rotation of one dIII into the face-up position could help dislodge dIII from the opposite protomer from dI.

Rotation of one dIII into the face-up position would bring it very close to the other dIII in the face-down conformation (Figure 3.13). They could form many hydrogen bonds without clashing and so interact when swivelling.

dI dimer is stabilised by a large buried surface extended by an elongated hairpin (ovine residues 172-184) reaching out from each monomer into the neighbour (Prasad *et al.*, 2002; Johansson *et al.*, 2005; Sundaresan *et al.*, 2005). The hairpin from one dI monomer interacts only weakly or not at all with the "face-down" dIII in another monomer (as seen in our structures), but strongly with the "face-up" dIII (as in dI_2dIII crystal structures). This hairpin originates directly from the Rossmann fold coordinating the NAD(H) in the first monomer, which may allow for communication of the redox state from one dI monomer to the other.

Additionally, the mammalian enzyme has another possibility for protomer communication. The dI-dII linker contacts the dIII from one monomer and is continued by the TM1(unique to metazoans and a few other species), which forms extensive interactions with

the TM5 (unique to metazoans) and the TM4-5 loop from another monomer, which in turn directly contact dIII from that monomer. This feature, in addition to the TM1/TM5 dimer stabilisation, may allow for more efficient dIII cross-talk. Together, all these possible routes of communication may allow for the two monomers to work more effectively in an anti-phase fashion.

3.6.3. Comparison with the bacterial NNT

In our datasets we did not observed a face-up dIII conformation since it is likely that the dI-dIII interface formation requires binding of the corresponding nucleotide to dI and once formed, has a very short life span as the hydride transfer step is several orders of magnitude faster than nucleotide exchange (Hutton *et al.*, 1994; Venning *et al.*, 1998, 2000). To explain the existence of a pure asymmetric dimer in *T. thermophilus* we need to note that a) this was observed in a crystal structure which is inherently more static than cryo-EM structures solved in solution and b) the lack of the dI-dII linker in *T. thermophilus.* We propose that the dI-dIII interface in bacteria is strengthened to prevent a simultaneous dissociation of both dIIIs which would lead to the loss of the dI_2 in the absence of the dI-dII linker. This is supported by studies of the hydride transfer activity which is two orders of magnitude faster in the isolated mixtures of *R. rubrum* and *E. coli* dI/dIII than in the mixtures of isolated bovine dI/dIII (Yamaguchi and Hatefi, 1995, 1997; Fjellstrom *et al.*, 1999). This notion explains why $NADP^+$ affects the structures of mammalian and bacterial NNT differently. It also explains the evolutionary reason for acquisition of the dI-dII linker, as it contributes towards higher efficiency of the mammalian enzyme, as the dI-dIII interface can be weaker and allow for faster turnover. dI-dII linker also allows NNT to exist in an "auto-inhibited" double face-down conformation, which could be important for additional regulation of mammalian NNT.

3.6.4. Potential role of kinetic restriction

The mechanism and ensuring the directionality could also be helped by a kinetic mechanism. Specifically, dIII-$NADP^+$ re-association with dII before hydride transfer takes place (step 4, Figure 3.21) could be prevented by the relative rates of H664 deprotonation into the N-side (slow) and hydride transfer (fast). Furthermore, dI has had time to exchange the nucleotide since hydride transfer of the previous cycle, while deprotonation of dII into the matrix can only take place after dIII has dissociated in step 4. Presumably, the kinetic parameters of NNT are tuned such that the reaction is driven forward. It will be interesting to

investigate the role of kinetics for the catalytic mechanism of NNT, but these parameters cannot be inferred from cryo-EM data.

3.7.Conclusions

The structures presented here are the first high-resolution structures of an intact transhydrogenase from any species. We describe for the first time in atomic detail the dII-dI linker, the face-down conformation of dIII, the structure of dIII without bound nucleotide, the dII-dIII domain interface and the structural changes induced in the proton channel and the NADP(H) binding site by dIII binding to dII. These observations, together with a wealth of published biochemical evidence on NNT, allow us to revise and streamline the catalytic mechanism. The revised mechanism elegantly explains the full reversibility of NNT, pH dependencies of reverse and cyclic reactions, different effects of $NADP^+$ and NADPH on the overall conformation of the NNT as well as complex phenotypes of mutants, promoting decoupling or cyclic reactions. Further work on NNT will need to explain how exactly the reaction is coupled in the two monomers. Our structures combined with a better understanding of the mechanism will facilitate further understanding of the functional role of NNT and could be used for structure-based drug design.

4. Complex I

4.1. Biochemical characterization of ovine complex I

4.1.1. Effects of detergents on NADH:DQ activity

As has been described previously, complex I activity depends highly on the type of detergent in which the protein is purified and assayed (Sharpley and Hirst, 2006; Letts *et al.*, 2016). Firstly, because the integrity of complex I depends strongly on the type of detergent, and many detergents interfere with the quinone or lipid binding pockets. Secondly, because the NADH:DQ assay employs a hydrophobic substrate (DQ, Figure 4.1) and adequate mixing and accessibility to the protein is required to achieve maximum activity that is not diffusion-limited. The assay conditions have been thoroughly optimized before (Letts *et al.*, 2016), and included the addition of CHAPS to the assay buffer, which greatly improves accessibility of quinone. Here we mainly compared the different preparations of complex I used for cryo-EM grid screening.

The highest activity (~6 U/mg) was observed in CxI preparations in very mild detergents such as FOM or LMNG (Figure 4.1). Surprisingly, activity in nanodiscs (MSP2N2) or amphipols, which should theoretically preserve the near-native lipid environment around complex I much better than detergents resulted in strong inhibition of activity (~1 U/mg). Nevertheless, the remaining activity was fully sensitive to rotenone, suggesting that the particles maintained structural integrity and remained well-coupled. The reduced activity was thus very likely due to the decreased rate of quinone exchange between the buffer and the nanodisc/amphipol particle. Activity in all of the detergents was higher than in Brij-35 which was used previously to determine the ovine complex I structure (Fiedorczuk *et al.*, 2016).

Since the FOM preparation activity was so high, the activity in a buffer containing only FOM and no CHAPS was also measured, but this resulted in lower final activity measurements, likely because DQ solubility in a FOM buffer is lower (Figure 4.2). To assess the effects of the temperature on the activity, the preparations with the highest activities were also assayed at 4ºC, to determine the rate of complex I reaction at the temperatures at which cryo-EM grids are typically prepared (Figure 4.2). Activity at 4ºC was reduced by about 90%. These measurements were used for predicting the consumption of substrates in the CxI sampled during turnover as described later.

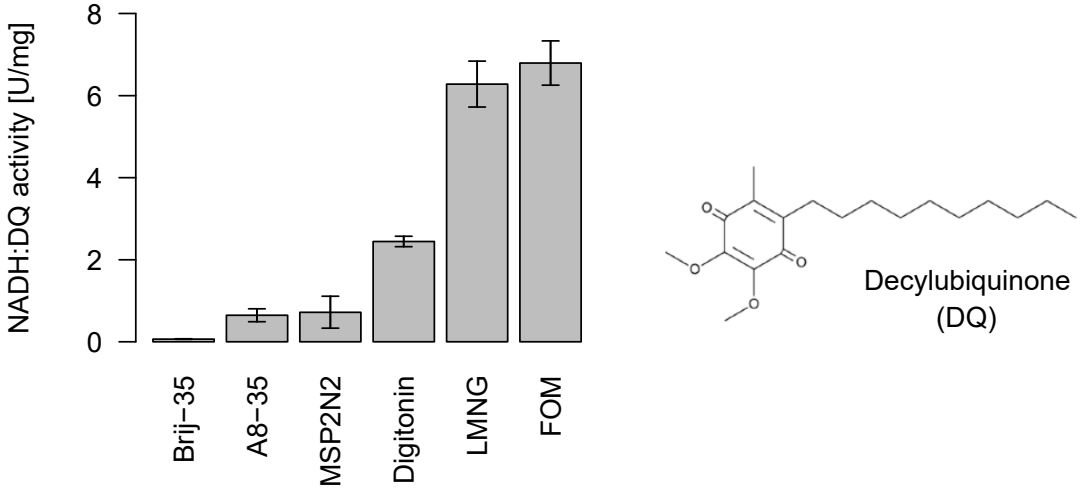


Figure 4.1. Complex I activity depends on the detergent.

Complex I has the highest activity in mild detergents such as LMNG and FOM but low activity in nanodiscs and amphipols.

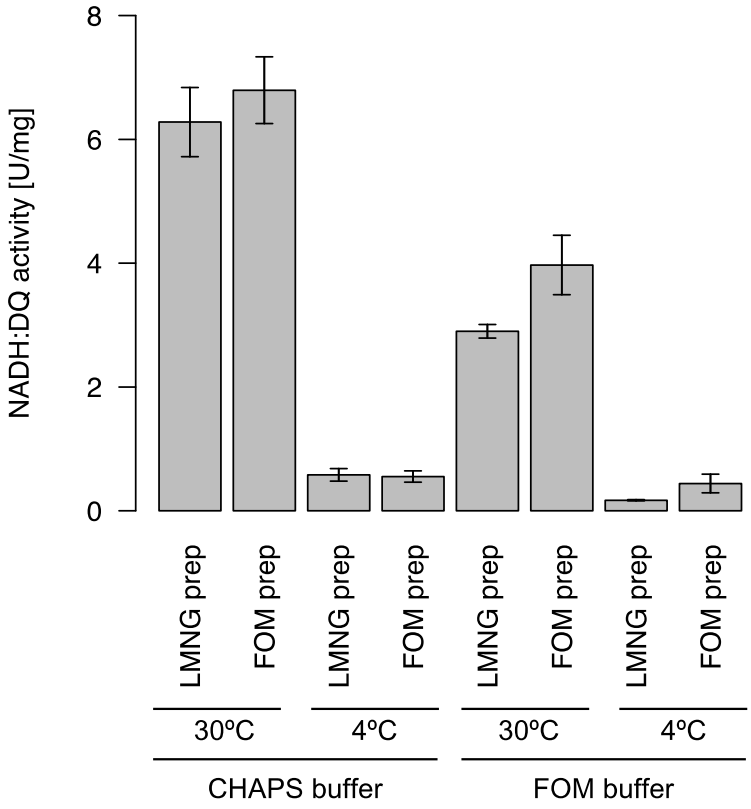


Figure 4.2. Complex I activity of preparations in FOM and LMNG in two different buffers at 30ºC and 4ºC.

4.1.2. Inhibition by rotenone and imiquimod

Many inhibitors of complex I are known and continue to be described (Murai and Miyoshi, 2016). Rotenone is one of the strongest known inhibitors of complex I (Figure 4.2). Ovine complex I preparation was fully sensitive to rotenone as described previously (Grivennikova *et al.*, 1997). Rotenone inhibition was also structurally characterized in this work for the first time.

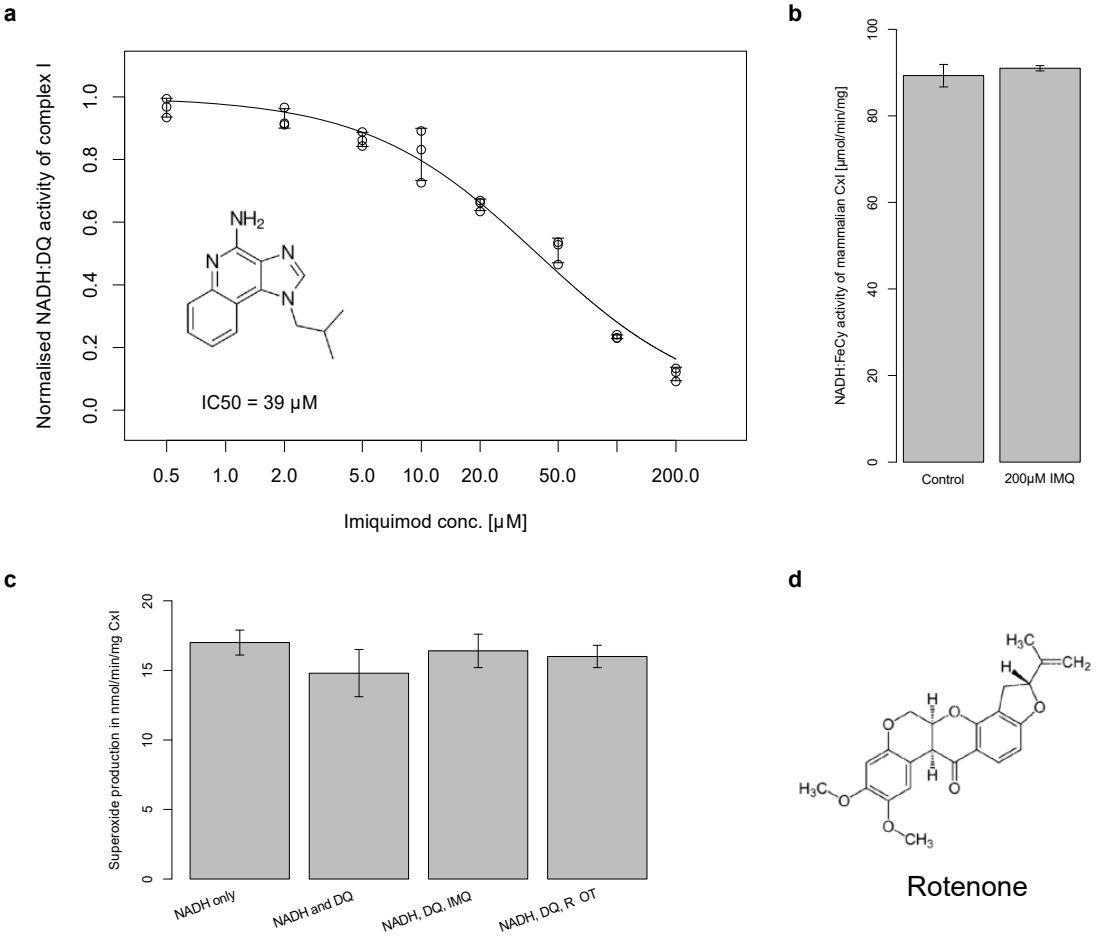


Figure 4.3. Inhibitors of complex I.

a. *Imiquimod is an inhibitor of ovine complex I with IC50 of 39 µM.* ***b.*** *Imiquimod does not inhibit NADH:Ferricyanide activity of complex, indicating that imiquimod binds within the quinone cavity and not at the FMN site.* ***c.*** *Neither rotenone nor imiquimod significantly affect ROS generation by complex I.* ***d.*** *Structure of rotenone.*

In addition to rotenone, imiquimod has recently been identified as a novel complex I inhibitor in vivo with potential clinical implications (Groß *et al.*, 2016). Imiquimod indeed inhibits the NADH:DQ reaction of the purified ovine complex I, albeit with a relatively high IC_{50} of around 39 mM (Figure 4.3a). Imiquimod does not inhibit NADH:ferricyanide oxidoreduction activity of complex I, indicating that imiquimod likely binds in the quinone

binding pocket (Figure 4.3b). Molecular similarity to quinone and rotenone as well as hydrophobicity are consistent with this conclusion.

Since an increase of ROS production in the presence of imiquimod was observed in cells, we tested effects of imiquimod on ROS production by purified complex I. There was no difference in the initial rates of ROS production by ovine complex I in the presence of NADH and either absence or presence of DQ, rotenone or imiquimod (Figure 4.3c). We measured superoxide production, since mammalian complex I produces ROS as ~95% superoxide/~5%H_2O_2 (Esterhazy et al., 2008). Thus, ROS production by ovine enzyme is similar across all conditions, as reported previously for complex I from *Yarrowia lipolytica* (Dröse, Galkin and Brandt, 2009). In vivo it is likely that the ROS increase comes from inhibition of complex I increasing NADH concentrations significantly and thus leading to increased ROS production at the FMN site, which is highly sensitive to NADH/NAD+ ratio.

4.2. Preliminary structure of ovine complex I in nanodiscs

Initially, reconstitution and structural studying of complex I in nanodiscs was initiated as nanodiscs provide a near-native lipid environment. As outlined in chapter 2, a preliminary cryo-EM dataset of complex I reconstituted into MSP2N2 nanodiscs and in the presence of DQ was collected. It showed that MSP2N2 can wrap around the elongated membrane arm of complex I, despite having a predicted nanodisc diameter of about 140 nm.

We could not observe bound quinone in this structure, either because too low concentration was used (20 μM) or because of the low resolution achieved (up to 4 Å in the peripheral arm and 4.4 Å in membrane arm). Lipid environment of this preparation preserved the structural integrity of loosely bound membrane subunit B14.7. Due to the small size of the dataset, the focus-reverse-classify procedure could not be implemented. Ultimately, due to the much higher activity of complex I purified in LMNG, the latter preparation was used for all subsequent high-resolution datasets.

4.3. High-resolution features of complex I model

The sample of complex I in LMNG proved to be the best candidate for high resolution structural analysis and it allowed us to build the most complete model of mammalian complex I so far (96.9% residues modelled, Table 4.1). Resolutions up to 2.3 Å for peripheral arm (PA) and up to 2.5 Å for membrane domain (MD) resulted in a better-defined geometry in many ambiguous regions (Figure 4.3). This is a significant improvement over the previous resolution of 3.3 Å (Agip *et al.*, 2018) and resulted in several novel features described below.

4.3.1. Post-translational modifications of complex I

Complex I is post-translationally modified at several residues, either during polypeptide synthesis and assembly or as part of regulation (Carroll *et al.*, 2005; Carroll, Fearnley and Walker, 2006). While many of the regulatory post translational modifications occur in intrinsically disordered N- or C-terminal parts of polypeptides, we observed density for several post-translational modifications connected with stability and assembly of complex I. These included N-terminal acylation of three of the supernumerary subunits, B18, B14.5a and B14.7 (Figure 4.4e). Almost all of the supernumerary subunits in complex I undergo some sort of N-terminal modifications upon import into the mitochondria but in many supernumerary subunits N-termini were disordered (Carroll *et al.*, 2005).

Furthermore, mitochondrially-encoded subunits of complex I (ND1-ND6) all have an N-formyl methionine, as a bacterial-like remnant of mitochondrial gene expression synthesis and the density for N-formyl methionine was clearly observed in three of these subunits (Figure 4.4e). Symmetric dimethylation of the conserved Arg85 of 49 kDa has been observed before and occurs in the vicinity of the N2 cluster (Agip *et al.*, 2018).

Finally, phosphorylation of the Ser36 of 42 kDa subunit has been described before (Schilling *et al.*, 2005) and it was suggested that the phosphorylated Ser36 would clash with nucleotide binding by 42 kDa subunit (Agip et al., 2018). 42 kDa is a nucleoside kinase homologue but despite possessing all the active site residues, no nucleoside kinase activity could be measured in the sample of purified ovine complex I. Phosphorylated S36 indeed clashed with ADP (or ATP) binding in the 42 kDa subunit, as suggested for the mouse complex I (Agip *et al.*, 2018), but there remained a smaller density in the active site which could be modelled as a molecule of AMP (Figure 4.4f).

Table 4.1. Model overview and comparison with the previous ovine complex I model (5LNK)

	Subunit name	Chain	Range built / (out of total) residues	Un-modelled residues	% atomic model NEW	% atomic model in 5LNK	Cofactors	Notes
CORE	51 kDa	1	9-438 / 445	1-8, 439-445	95.6	97.1	FMN N3(4Fe) NADH	
	24 kDa	2	5-217 / 217	1-4	98.2	98.6	N1a (2Fe)	
	75 kDa	3	6-693 / 704	1-5, 694-704	97.7	97.7	N1b(2Fe) N4 (4Fe) N5 (4Fe)	
	49 kDa	4	1-430 / 430	0	100	90.0	DQ, also rotenone	several conformations; dimethylated Arg85
	30 kDa	5	7-214 / 228	1-6, 215-228	91.2	91.2		
	PSST	6	24-179 / 179	1-23	87.2	86.6	N2 (4Fe)	several conformations
	TYKY	9	1-176 / 176	0	100	100	N6a (4Fe) N6b (4Fe)	
	ND3	A	1-115 / 115	0	100	100		15-50 re-modelled
	ND1	H	1-318 / 318	0	100	100	DQ, also rotenone	several conformations, parts re-modelled
	ND6	J	1-175 / 175	0	100	73.7		
	ND4L	K	1-98 / 98	0	100	87.8		N-formyl Met
	ND2	N	1-347 / 347	0	100	100		
	ND4	M	1-459 / 459	0	100	100	rotenone	N-formyl Met
	ND5	L	1-606 / 606	0	100	84.0		N-formyl Met
SUPERNUMERARY	10 kDa	a	32-75 / 75	1-31	58.7	54.7		
	13 kDa	b	1-95 / 96	96	99.0	99.0	Zn^{2+}	
	18 kDa	c	8-133 / 133	1-7	94.7	92.5		
	39 kDa	d	1-340 / 345	341-345	98.6	87.0	NADPH	previously missing C-terminus and two internal loops
	B8	e	13-98 / 98	1-12	87.8	85.7		
	B13	f	3-115 / 115	1-2	98.3	97.4		
	B14	g	14-127 / 127	1-13	89.8	89.8		
	B14.5a	h	1-112 / 112	0	100	84.8		N-acetyl Ala
	B17.2	i	1-145 / 145	0	100	99.3		
	SDAP-α	j	5-84 / 88	1-4, 85-88	90.9	96.6	Phospho-pantetheine	
	42 kDa	k	1-320 / 320	0	100	70.9		
	15 kDa	l	1-105 / 105	0	100	90.5		
	B9	m	4-83 / 83	1-3	96.4	96.4		
	B12	n	11-89 / 97	1-10, 90-97	81.4	73.2		
	B14.5b	o	1-120 / 120	0	100	100		
	B15	p	1-128 / 128	0	100	62.5		
	B16.6	q	5-143 / 143	1-4	97.2	97.9		
	B17	r	1-38, 66-126 / 127	37-65, 127	76.4	66.1		
	B18	s	1-122 / 136	123-136	89.7	78.7		N-myristolyl Gly
	B22	t	1-177 /178	178	99.4	93.3		
	AGGG	u	5-69 / 72	1-4, 70-72	90.3	91.7		
	ASHI	v	4-158 / 158	1-3	98.1	27.2		
	ESSS	w	6-123 / 125	1-5, 124-125	94.4	68.8		
	KFYI	x	1-49 / 49	0	100	98.0		
	MNLL	y	8-57 / 57	1-7	87.7	93.0		
	MWFE	z	1-70 / 70	0	100	98.6		
	B14.7	V	1-140 / 140	0	100	0		previously polyAla; N-acetyl Ala
	SGDH	W	5-143 / 143	1-4	97.2	97.2		
	SDAP-β	X	2-88 / 88	1	98.9	100	Phospho-pantetheine	
	PGIV	Y	1-171 / 171	0	100	100		
	PDSW	Z	3-173 / 175	1-2, 174-175	97.7	97.7		
	TOTAL		**8256 / 8516**	**260**	**96.9 %**	**88.3%**		

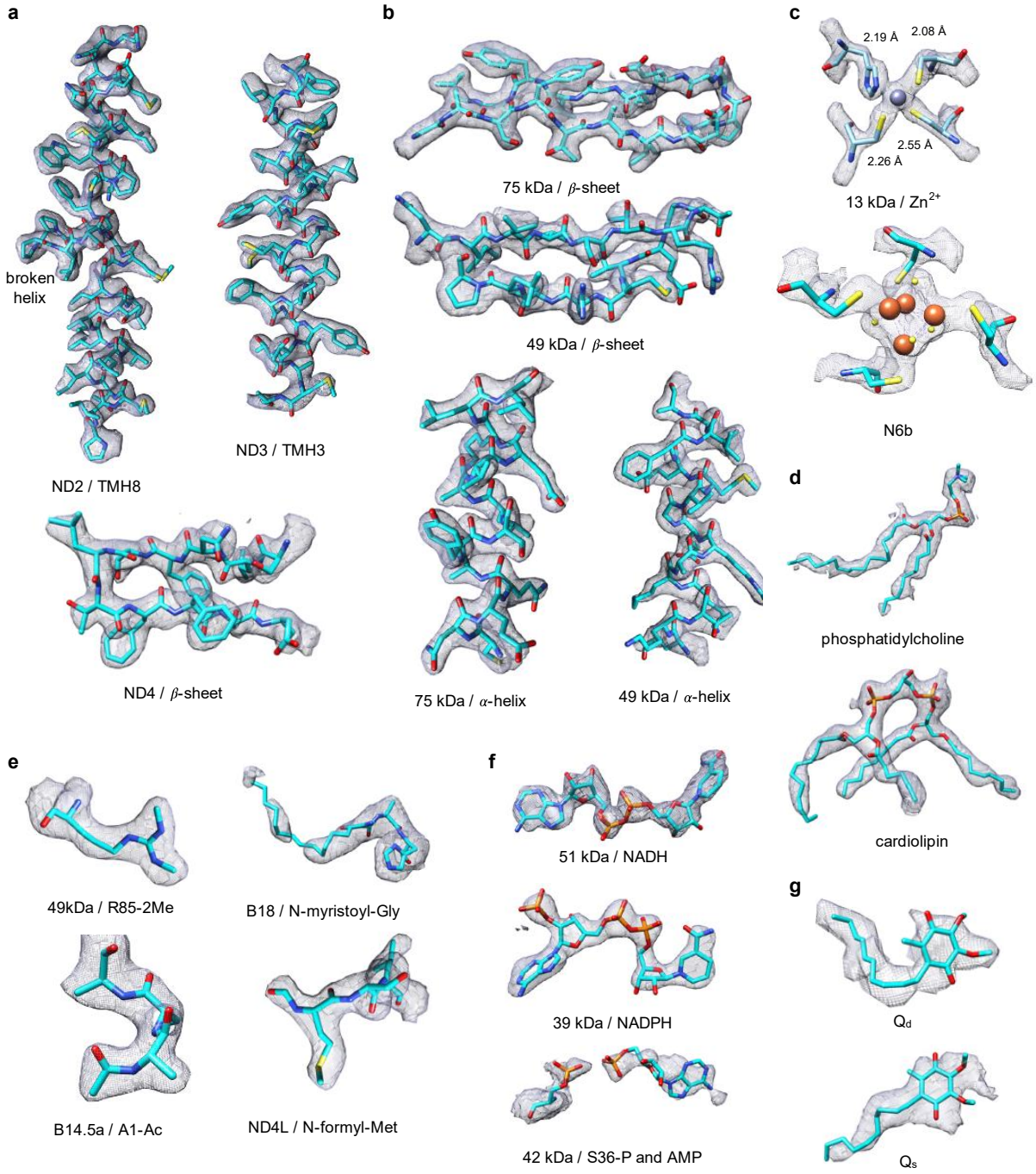


Figure 4.4. Cryo-EM density examples.

a. *Density in various regions of the MD focus-refined map from the turnover dataset.* ***b.*** *Density in various regions of the PA focus-refined map from the turnover dataset.* ***c.*** *Densities for metal ions and iron-sulfur clusters.* ***d.*** *Density for lipid molecules.* ***e.*** *Densities for various post-translational modifications.* ***f.*** *Densities for bound nucleotides.* ***g.*** *Densities for quinones.*
Adapted from Kampjut & Sazanov, 2020.

4.3.2. Lipid binding sites

We modelled more and better-defined lipid molecules than previously, including cardiolipins and several lipid molecules bound above the quinone entry site, indicating the local distortion of the membrane plane around the quinone entry site (Figure 4.5). This puts the quinone entry site roughly between the two lipid leaflets and was proposed earlier to facilitate quinone flipping and diffusion (Parey *et al.*, 2019). Cardiolipin, known to be essential for respiratory enzymes activity, was found in 8 locations, some of them consistent with locations in other species (Agip et al., 2018; Blaza, Vinothkumar and Hirst, 2018) along with 5 new ones. Supernumerary subunit B14.7 is almost fully separated from the MD by a lipid cushion, hence this subunit could have been acquired solely for the interaction with complex III in the respiratory supercomplexes.

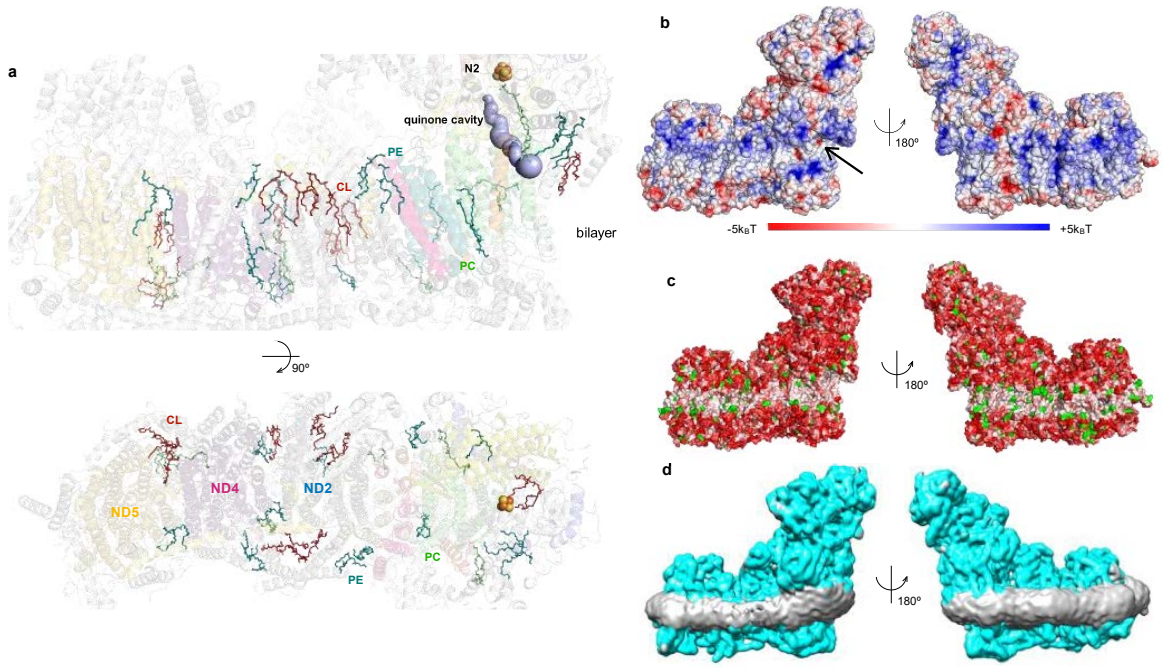


Figure 4.5. Interactions of complex I with lipids.

a. *Lipids modelled into the complex I structures shown as sticks. Phosphatidylethanolamine (blue), cardiolipin (red) and phosphatidylcholine (green).* ***b.*** *Surface electrostatic potential of complex I reveals many lipid binding sites, negatively charged quinone entry site and membrane curvature around quinone entry point.* ***c.*** *Hydrophobicity plot of complex I with white regions showing hydrophobic and red regions showing hydrophilic residues reveals the transmembrane region. Tryptophan and tyrosine residues, which often denote membrane edge are shown in green.* ***d.*** *Open2-turnover structure filtered to 8 Å reveals curvature of the detergent belt around the quinone entry site.*
Adapted from Kampjut & Sazanov, 2020.

4.3.3. Ion and water densities

Finally, we observed a highly coordinated (four backbone oxygens, a glutamine and one water molecule) cation bound about 7 Å away from FeS clusters N4 and N5. We tentatively assigned the cation as K^+, based on its density (much stronger than water) and distances (~ 2.7 Å) to coordinating oxygen atoms (Figure 4.6a) (Zheng *et al.*, 2008). Presence of the cation in the vicinity of N5, which unusually does not show EPR signal, could be necessary to fine-tune its potential or that of the nearby N4 cluster (Bridges, Bill and Hirst, 2012). The cation probably also stabilizes a very tight turn of the 75 kDa loop between the N4 and N5 clusters.

We were also able to model water molecules, with important implications for the proton translocation mechanism which will be discussed in detail later. Densities for water molecules were especially well interpretable in local resolution filtered oversampled maps of focus-refined MD and PA of complex I during turnover (Figure 4.6b,c).

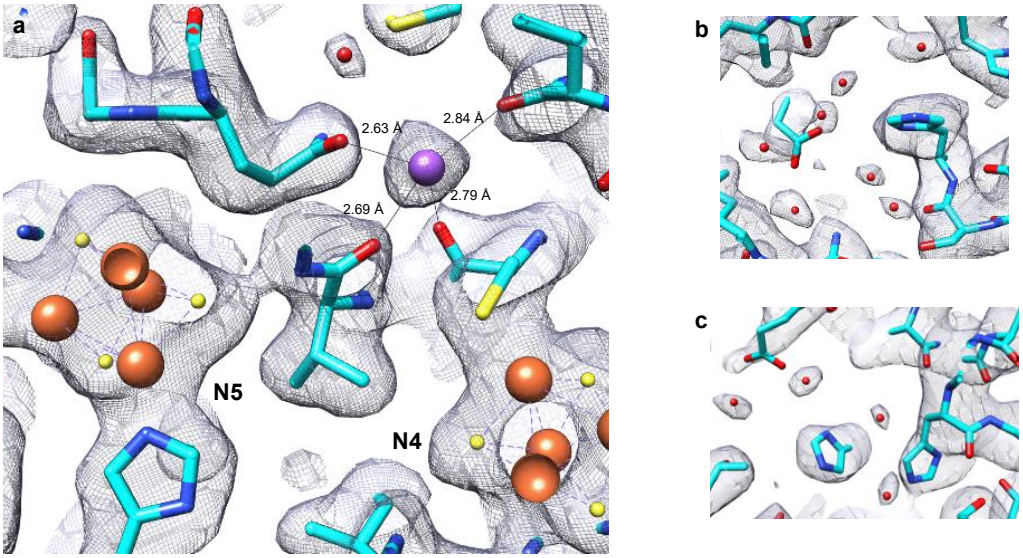


Figure 4.6. Further Cryo-EM density examples.

a. *Density and surroundings for the putative K+ ion bound next to the N5 and N4 cluster.* ***b.*** *Examples of water densities near ND4/E378 in the open MD class.* ***c.*** *Example of water density around ND4/E378 in closed MD structure.*
Adapted from Kampjut & Sazanov, 2020.

4.4. Comparison of open and closed complex I

4.4.1. Overview

We collected datasets of ovine complex I in LMNG in five different conditions: native, with added NADH, with added NADH and DQ, with added rotenone and NADH, and in the deactive state. All of the datasets contained a mixture of open and closed conformations of complex I arising from the combined side- and frontward tilting of the PA by up to 7º as described previously (Vinothkumar, Zhu and Hirst, 2014; Fiedorczuk *et al.*, 2016; Zhu, Jiapeng, Vinothkumar and Hirst, 2016; Letts *et al.*, 2019) (Table 4.2, Figure 4.5). The relative ratios between open and closed conformations were roughly the same in most of our datasets with the exception of the deactive complex I, which completely lacked a closed conformation and complex I inhibited with rotenone, which had a smaller proportion of the closed state, presumably as a result of the presence of rotenone (or DMSO in which it was dissolved) (Figure 4.7). In each dataset, several open classes could be classified. The most open classes (largest PA-MD angle) of the native, deactive and rotenone datasets showed partial disordering of the ND6 TMH4, 49 kDa N-terminus, anchor helix of ND5 and B14.7, which could represent a partially disassembled or deactivated enzyme. Open classes from NADH dataset on the whole appeared slightly more closed than in the native dataset, possibly due to the localized differences of the PA-MD interface as described below.

Otherwise, the open classes did not differ from each other, particularly at the catalytically significant sites. The different degrees of openness in open classes thus likely reflect the loose PA-MD interface rather than functionally different states of complex I. The transitions between open classes should be viewed as continuous and we therefore treat open complex I here as a single, albeit heterogeneous, state of the complex.

Table 4.2. Proportion of open and closed conformations of complex I in the collected LMNG datasets and resolutions achieved

Condition	Closed (%)	res. (Å)	Open (%)	res. (Å)	PA res. (Å)	MD res. (Å)
native	12	3.8	88	3.2	2.8	3.2
NADH	10	3.4	90	2.8	2.5	2.8
rotenone	5	3.7	95	2.8	2.4	2.8
turnover	11	3.1	89	2.6	2.3	2.5
deactive	0	N/A	100	2.8	2.5	2.8

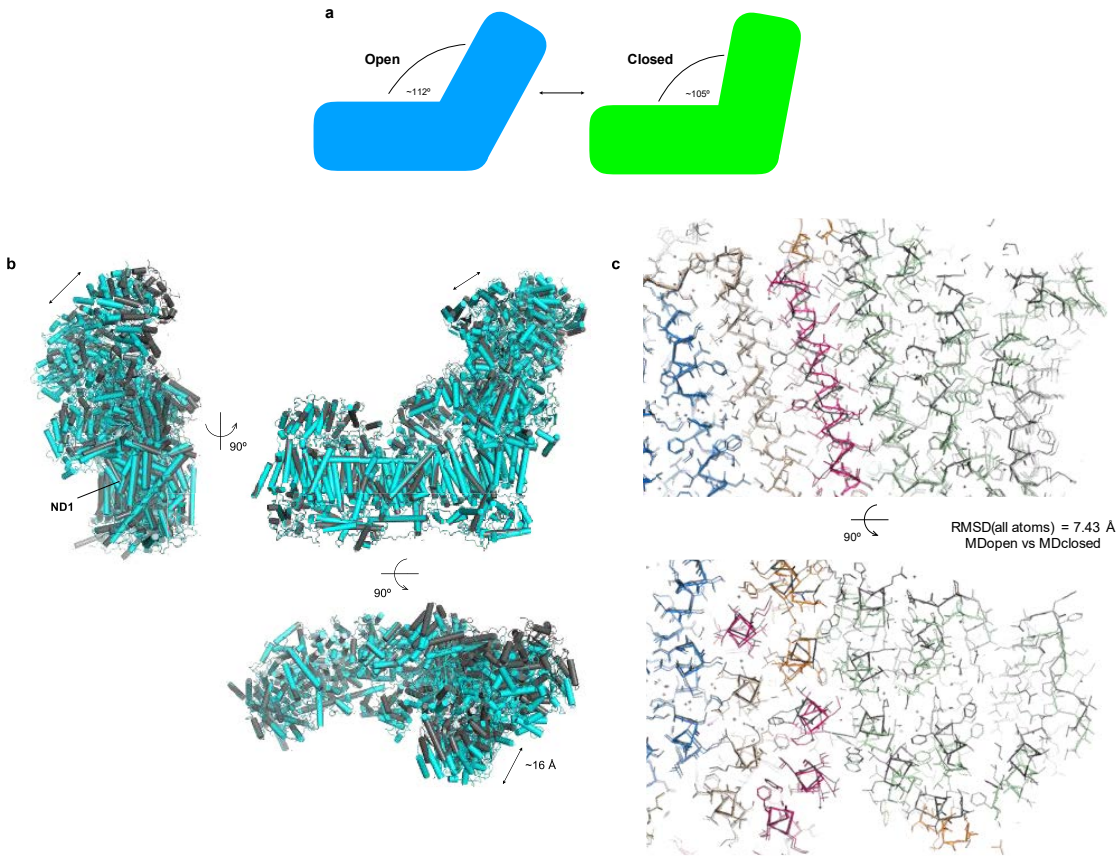


Figure 4.7. Closed and open conformations of complex I.

a. *Ovine complex I exists in open and closed conformations, the ratios of which differ between different conditions investigated (Table 4.2).* ***b.*** *Overall comparison of closed (cyan) and open (grey) complex I reveals tilting of the entire PA and large changes in ND1 subunit.* ***c.*** *E- channel in open and closed MD differs radically in tilting and displacements of several helices and different side chain conformations.*
Adapted from Kampjut & Sazanov, 2020.

The opening of complex I coincides with unfolding of several loops, which can roughly be divided into the quinone cavity forming loops and the interface forming loops (Figure 4.6). The former consists of the 49 kDa β1-β2 loop (yellow) and the PSST loop (grey) that form the deep part of the quinone binding pocket and the ND1 TMH5-6 loop (green) that forms the shallow part of the pocket. The 49 kDa loop binds the quinone head group and is connected to the MD via the rest of the β-sheet. The interface-forming loops consist of the ND3 loop (orange) and the TMH3-4 loop of ND6 (blue). Unfolding of these loops on closed-to-open transition is associated with a notable rotation of the C-terminal half of TMH3 of ND6 by about 120º and the appearance of the π-bulge in the middle of the helix (Agip *et al.*, 2018; Letts *et al.*, 2019).

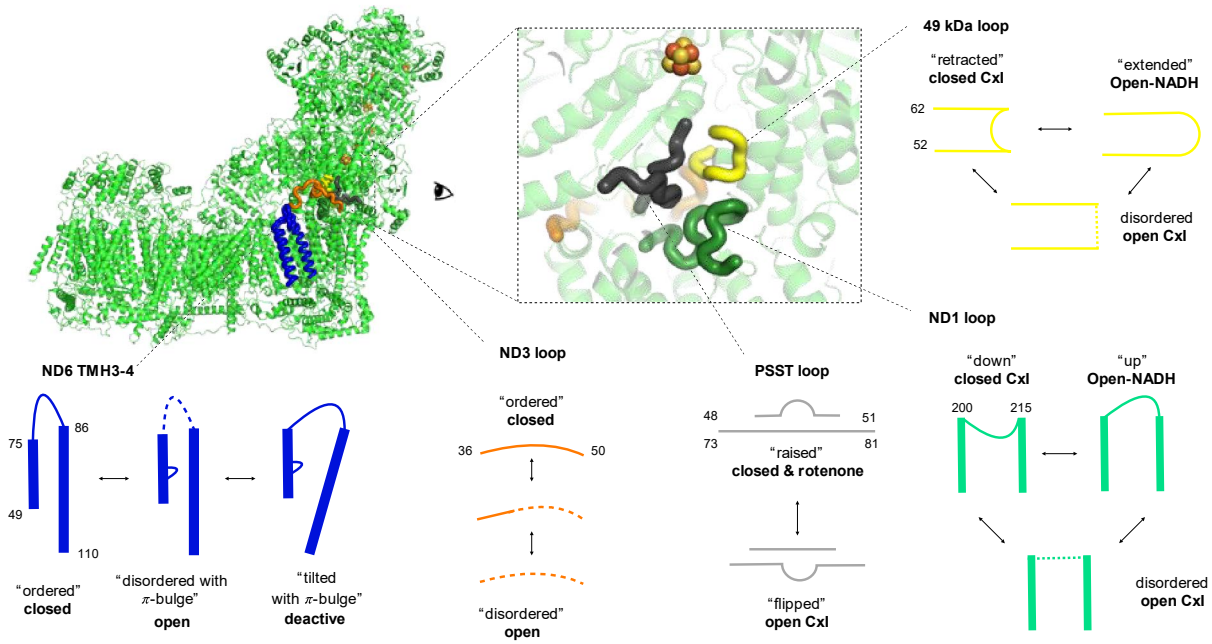


Figure 4.8. Conformational dynamics of the quinone binding site loops.

Loops at the interface between the peripheral and membrane arms, ND3 (orange) and ND6 (blue) adopt different states shown schematically. Quinone binding cavity loops include the 49 kDa loop (yellow), PSST loop (grey) and ND1 loop (green), which also undergo radical changes, as indicated.
Adapted from Kampjut & Sazanov, 2020

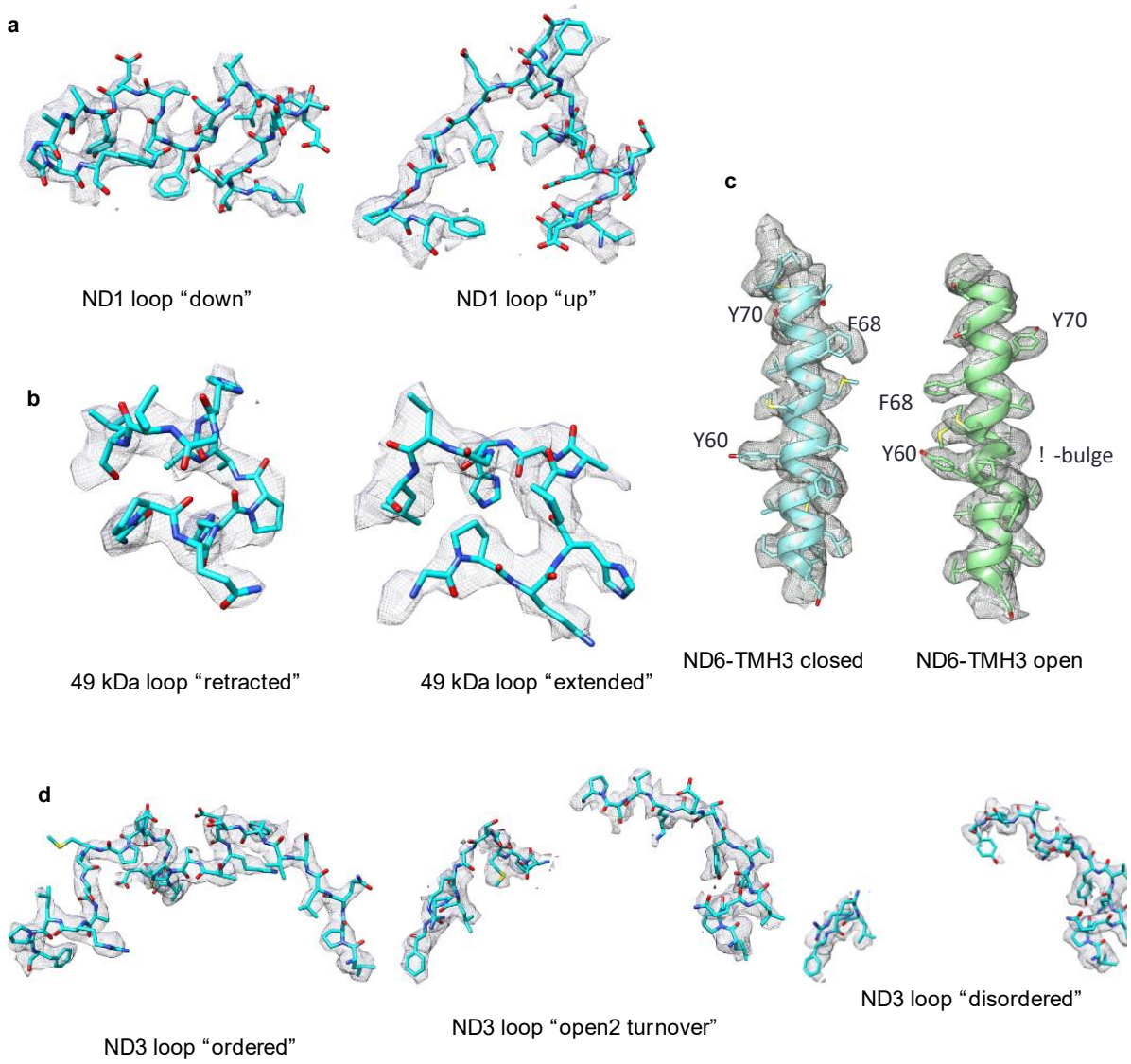


Figure 4.9. Densities for the mobile loops in different conformations.

a. *Density for the ND1 loop in the “down” conformation (Turnover closed class, left-hand side) and the “up” conformation (NADH open2 class, right-hand side).* ***b.*** *Density for the 49 kDa loop in the retracted (Turnover closed class, left-hand side) and in the extended conformation (NADH open2 class, right-hand side).* ***c.*** *Density of the ND6 TMH3 in closed state (without* ***π****-bulge) and in open state (with* ***π****-bulge). N.B. the rotation of the upper part of the helix.* ***d.*** *Density for the ND3 loop in the fully ordered (turnover closed class, top), partially disordered (open2 turnover class, middle) and disordered conformation (NADH open2 class, bottom).*
Adapted from Kampjut & Sazanov, 2020.

4.4.2. Opening and reduction result in vast rearrangements of the mobile loops

In all of the open conformations (from all datasets) we observed disordering of the central part of the ND3 loop (residues 38-42 involving also small changes in the adjacent termini of TMH2 and TMH3), the ND6 loop and the amphipathic helix (residues 259-272) of the supernumerary 39 kDa subunit. While the N-terminal part of the ND3 loop was ordered in all datasets (residues 27-37), the C-terminal part of ND3 loop (residues 43-48) was disordered in NADH dataset, but ordered in most others.

Dramatic changes were observed also in the whole of ND1 subunit, with shifts of most helices increasing from the IMS towards the matrix side and reaching up to 5 Å displacements (Figure 4.10a). These movements in turn lead to the observed tilting of PA and probably also help to accommodate the widening of quinone cavity as described below. Furthermore, we observed formation of the π-bulge in TMH3 of ND6, resulting in a remarkable 120º rotation of the C-terminal half of the helix above the π-bulge upon closed/open transition (Figure 4.10b).

In contrast to the interface forming loops, the conformations of the quinone binding site loops (49 kDa, PSST and ND1), depended not only on the openness, but also on the ligands bound and the reduction state of the whole complex.

The 49 kDa loop was ordered in the retracted (out of quinone cavity) conformation in closed classes and disordered in the open classes of the native, turnover and deactive datasets. When rotenone was bound in the deep binding site (Q_d) in open complex I, 49 kDa loop was also ordered in the retracted conformation. In open classes of NADH-reduced enzyme (but not closed classes), however, 49 kDa loop was ordered in the extended-into-the-quinone-cavity conformation, shortening the cavity considerably (Figure 4.10d). These results suggest that binding of a ligand in the Q_d site can order the 49 kDa loop. Since the extended conformation was not observed in the rotenone and turnover datasets, which also contained NADH, quinone-like ligand probably must be fully ejected from the cavity before the loop can extend.

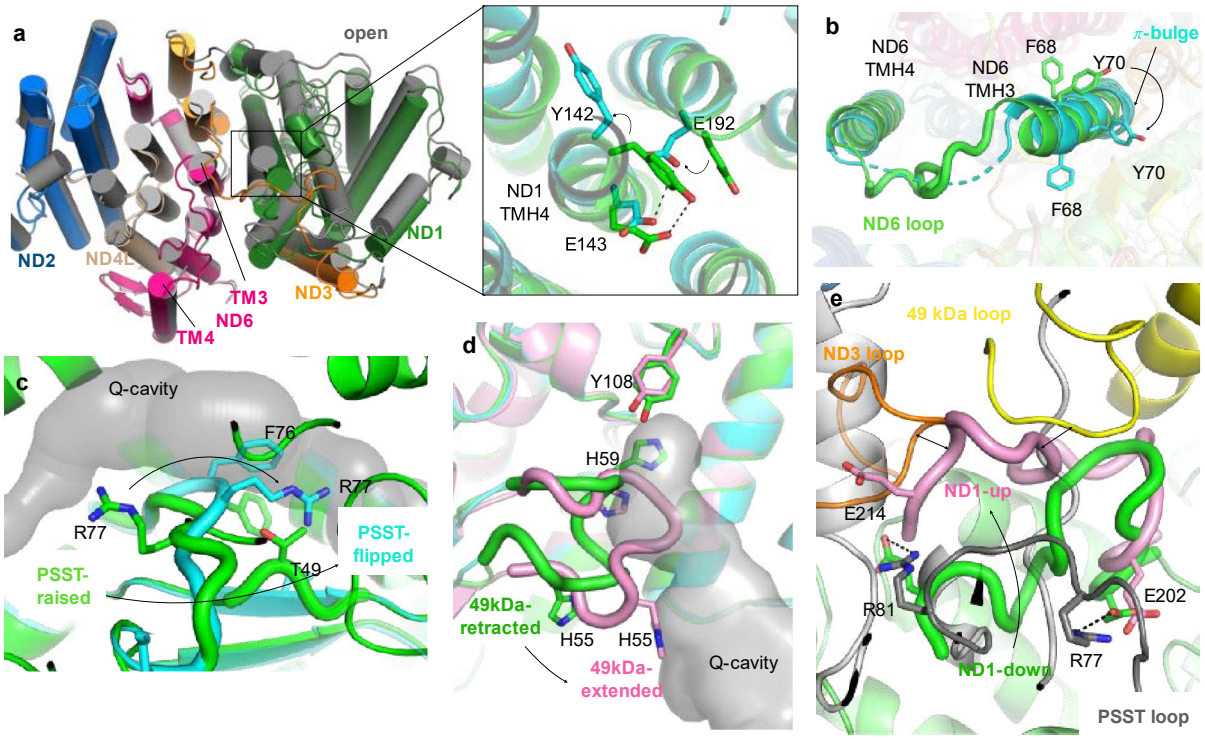


Figure 4.10. Opening and reduction of the complex induce long-range conformational changes.

a. Reogranization of the ND1 subunit upon opening with a detailed view of the rearrangements of the conserved glutamates which are part of the E-channel. On the right open state is gray and closed state is colored by subunit, while on the right open state is cyan and closed is green. b. TMH3 of ND6 rotates by almost 180o in the open state (cyan), acquires a π-bulge near Y60 and the TMH3-4 loop gets disordered. c. PSST loops and R77 undergo large conformational changes between open (cyan) and closed (green) states. d. 49 kDa loop is ordered in the closed state (green), gets disordered upon opening (cyan) and reorders in the extended position the reduced open state (pink). e. ND1 loop which is in the "down" conformation in the closed state (green), gets disordered upon opening and switched to the "up" conformation in the reduced open state (pink). In the "up" position ND1 clashes with the ND3 (orange) and 49 kDa (yellow) loops which are also reordered. Reorganization of the ND1 loop results in breaking of two salt bridges with PSST/R77 and R81 (grey).
Adapted from Kampjut & Sazanov, 2020.

ND1 loop was similarly ordered in the "down" conformation in the closed classes and disordered in the open classes of the native, turnover, rotenone and deactive datasets. In NADH open classes (but not closed), however, ND1 loop rearranged into an "up" position, which displaced the ND3 loop as well as the other two ND1 matrix side loops (Figure 4.10e). Disordering of ND1 loop breaks several salt bridges, among which are highly conserved E214-PSST/R81 and E202-PSST/R77. Mutagenesis studies showed inhibitory effects of R77 and R81 mutations and MD simulations suggested that quinone dynamics in the cavity is correlated with the state of the salt bridge PSST/R81-ND1/E214; i.e. quinone can only diffuse within the cavity when the salt bridge is broken and the ND1 loop disordered (Galemou Yoga *et al.*, 2019).

In the “up” conformation, E214 formed an alternative salt bridge with ND1/R62 and a hydrogen bond with PSST/Q106, while E202 was exposed to the quinone cavity lumen (Figure 4.10e). ND1 loop did not switch into the “up" conformation in rotenone and turnover datasets, which also contained NADH, suggesting that binding of the ligand in the quinone cavity reverses the NADH-induced changes and could be an important step during the catalytic cycle.

In contrast to 49 kDa and PSST loops, ND1 loop not only reforms the cavity upon the transition into the “up” conformation but it also appears to push ND1 TMH3 and TMH4 about 5 Å towards MD. Strikingly, as TMH4 bends near Y142 upon such transition, this tyrosine swings out in the open state by almost 180º, freeing the space for E192 to move within hydrogen bonding distance to E143 (Figure 4.10a). Changes in density appearance and the predicted pK_a also indicate that at least one of the glutamates (E192, pK_a increases from 7.52 to 9.15 as calculated by ProPKa software (Olsson *et al.*, 2011)) becomes protonated in the open state. These extensive conformational changes propagate in the membrane domain from the Q cavity all the way towards the ND6/TMH3 π-bulge and the E-channel (Figure 4.7), likely representing a major part of the coupling mechanism.

PSST loop itself also changed the conformation in the native, turnover, NADH and deactive datasets. In the closed classes it adopted the “raised” conformation, in which residues 48-51 bulge out into the quinone cavity, while in the open classes they form part of the β-sheet. Furthermore, residues 75-80 flip in the open classes, resulting in a 180º rotation of R77 (Figure 4.10c). In the rotenone dataset, however, PSST loop always adopted the raised conformation and R77 contributed to rotenone binding as discussed later.

4.4.3. Rearrangements of the quinone binding cavity

Opening of complex I and change in the conformations of the above loops completely reshape the negatively charged quinone binding cavity (Figure 4.11a). Closed cavity is very tight, clearly excluding the possibility of quinone protonation by the bulk solvent. In the open conformation, quinone cavity is much wider, in a large part due to PSST re-arrangements. It becomes possibly even accessible to bulk solvent near the ND3 loop (although that may be closed by a disordered loop) and also through extra cavity passing by ND1/R134. This widening will facilitate quinone movements in and out of cavity before and after reduction. In the NADH open conformation cavity becomes much shorter, mostly due to 49 kDa loop insertion. This will facilitate ejection of quinol upon quinone reduction and protonation.

Our open NADH structures do in fact show a density near the quinone binding cavity entry (Q_s site), which could be modelled as a quinone, presumably co-purified with the enzyme (Figure 4.11b). The density appeared stronger in the open NADH classes than in the closed class. There was also a weaker density found in a similar position in the native complex I (Figure 4.11b). We cannot unambiguously assign these densities as they could come from lipid, detergent or native quinone (in the case of native complex) or quinol (in the case of NADH-complex I). Density in a similar position has also been reported in *Y. lipolytica* and has been hypothesized to correspond to a negatively charged lipid head group (Parey *et al.*, 2018) and more recently to a native quinone (Parey *et al.*, 2019). Based on the appearance, connectedness and strength of these densities we modelled quinol into the NADH-open complex I but not in the other classes.

Finally, our results do not support the quinone shuttling mechanism (Brandt, 2011; Wikström and Hummer, 2012; Wikström, Sharma, Ville R I Kaila, *et al.*, 2015; Haapanen and Sharma, 2018). UQ_{10} bound in the Q_d site would occlude with its long tail the entire cavity along with the shallow binding site (Baradaran *et al.*, 2013; Warnau *et al.*, 2018).

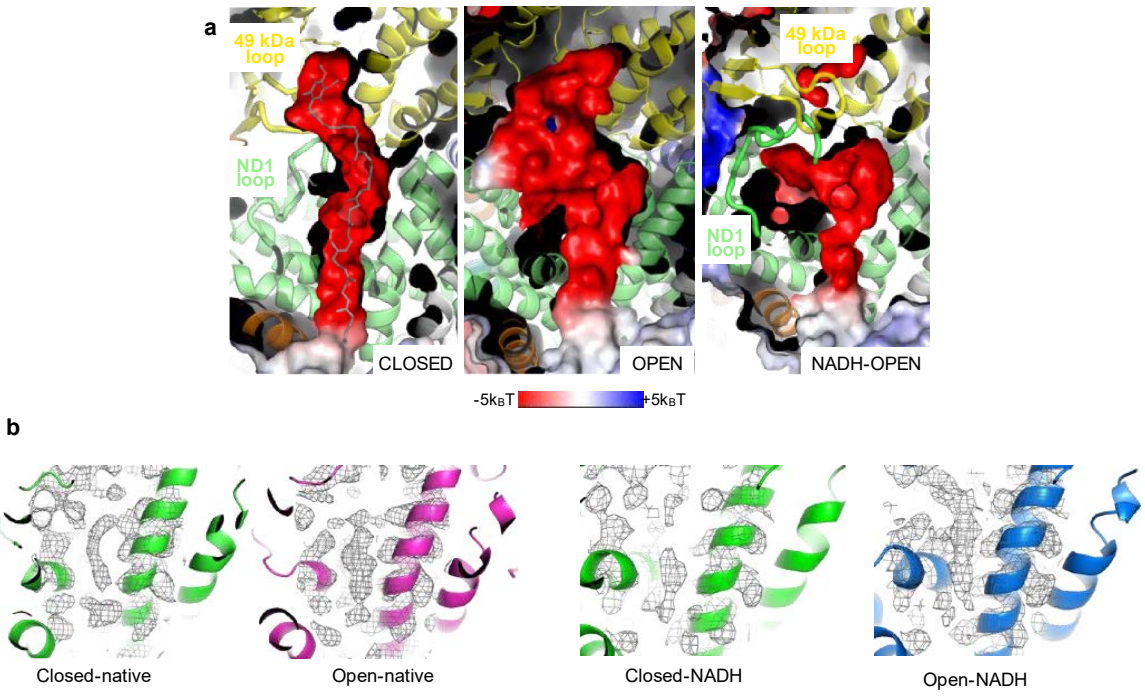


Figure 4.11. Changes in the quinone binding cavity.

a. *Cross sections through the quinone binding cavity in closed, open and NADH-open states colored by electrostatic potential reveal that opening and reduction result in large changes of the quinone binding cavity shape. In closed state UQ-10 is modelled (gray), based on decyl-ubiquinone binding and illustrating tight fit of the native substrate.* ***b.*** *Comparison of the densities near the entry into the quinone cavity in closed native complex*

(green) vs. open native complex (magenta) and closed NADH complex (green) vs. open NADH complex (blue). NADH open state has the strongest density, which was modelled as quinone.
Adapted from Kampjut & Sazanov, 2020.

4.5.Rearrangements of the NADH binding site

In all datasets containing NADH (NADH, turnover and rotenone) a clear NADH density was resolved in the same position near FMN as in *T. thermophilus* (Figure 4.12a) (Berrisford and Sazanov, 2009). As the NADH binding site is rather shallow and open to the solvent, no further changes were observed upon NADH binding. It was reported based on the NuoEF subcomplex crystal structures from *Aquifex aeolicus* that the peptide bond 51kDa/E99-G100 flips away from FMN when the subcomplex is reduced by NADH, as a mechanism to prevent ROS production (Schulte *et al.*, 2019). However, C=O bond is always facing away from FMN in ovine enzyme in all the datasets of sufficient resolution, whether the enzyme is reduced or oxidized (Figure 4.12b,c). This casts doubts over the proposed mechanism (Schulte *et al.*, 2019) and the difference could be because of substitution of serine for glycine at position 100 in ovine complex I or the effect of working with a subcomplex in *A. aeolicus*.

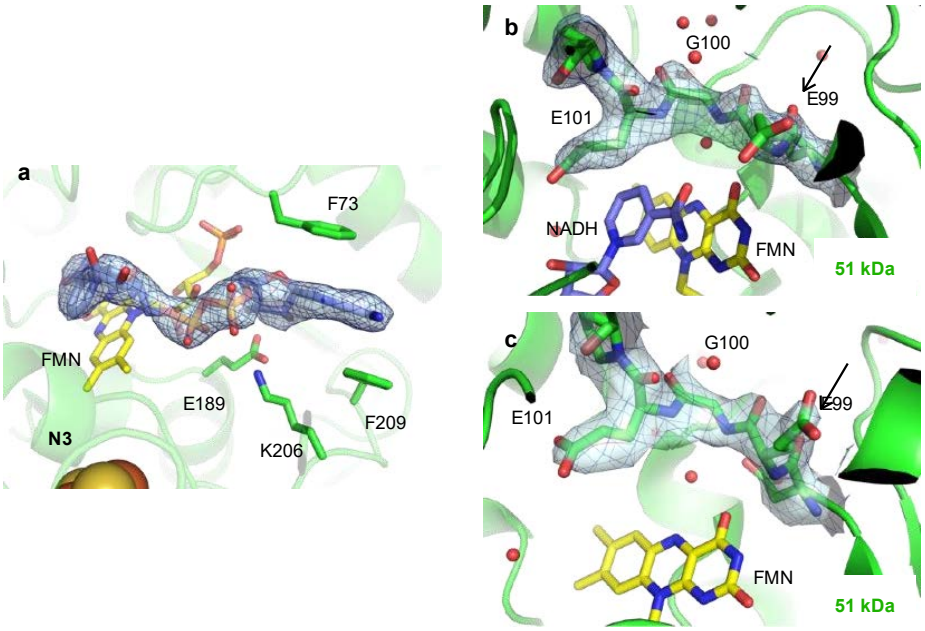


Figure 4.12. NADH-induced changes in the NADH-binding pocket.

a. *NADH density in the 51 kDa subunit. NADH binds by forming -stacking interactions between the nicotinamide ring and the FMN as well as between the adenine and F73 and F206.* ***b-c.*** *The conformation of the backbone between residues 98 and 101 remains unchanged in all the models and backbone carboxyl of G99 (indicated) faces away from the FMN in all cases. Turnover PA focus-refined map (**b**) and deactive PA focus-refined map (**c**) are shown for comparison.*
Adapted from Kampjut & Sazanov, 2020.

4.6. Rotenone binding sites differ between open and closed complex I and include a unique site in the antiporter subunit.

Structurally, inhibition of complex I is not well described and the traditional three classes of complex I inhibitors, A (competitive), B (non-competitive), and C (amilorides, biguanides, capsaicin), were predicted to have distinct binding sites (Degli Esposti, 1998; Murai and Miyoshi, 2016). In the *Y. lipolytica* complex I structure, a class A inhibitor DQA density was reported in the vicinity of the 49kDa/H59 homologue. Rotenone, one of the strongest known specific inhibitors of mammalian complex I (IC_{50} ~ 1nM) and an atypical class B inhibitor, was predicted to bind next to the ND1 loops by photoaffinity labelling studies (Earley and Ragan, 1984; Earley *et al.*, 1987). Because rotenone is much bulkier than typical Q-like inhibitors (such as piericidin A), it was hypothesized that it cannot penetrate into the Q cavity of complex I and might instead block the cavity by binding outside (Haapanen and Sharma, 2018). Rotenone was also reported to have different K_i values for the active (1 nM) and deactive (80 nM) complex I and its tight binding to the active form partially prevents and reverses enzyme deactivation (Kotlyar and Vinogradov, 1990; Grivennikova *et al.*, 1997).

We observed rotenone density in the Q_d site in all the conformational classes of the ovine complex I, coordinated by the key Y108 and H59, mimicking DQ and piericidin A binding in *T. thermophilus* (ROT1, Figure 4.13a,c) (Baradaran *et al.*, 2013). Additionally, in the open classes rotenone was also found in the Q_s site close to ND1/R25, exactly where a stable quinone binding site #4 has been proposed based on MD simulations (Haapanen, Djurabekova and Sharma, 2019) (ROT2, Figure 4.13a,d). Finally, in the open2 and 3 classes (84% of all particles) rotenone density was unexpectedly observed far away from the quinone cavity in the ND4 subunit. This binding site is created by tilting out of the TMH6 and is deep between TMH5-7, directly contacting the key K206 in the proton channel (ROT3, Figure 4.13a,e,f).

We cannot unambiguously assign the K_i values to the three binding sites. The ND4 binding site could correspond to the lower affinity (80 nM) binding site reported for the deactive conformation (Grivennikova *et al.*, 1997), which would also be consistent with changes observed in the deactive conformation that expose this rotenone binding pocket (Figure 4.13f). Binding in the ND4 site could also explain why rotenone has been observed to inhibit Na/H^+ antiporter activity of deactive complex I (Roberts and Hirst, 2012). ND4-bound rotenone, curiously, did not influence the conformations of the proton pumping residues of the

ND4 subunit, despite being bound directly by the connecting lysine ND4/K206. It would be interesting to see if any other inhibitors bind in this region, particularly biguanides which have been shown to bind more strongly to the deactive form of complex I (Matsuzaki and Humphries, 2015). Existence of inhibitors specific for the antiporter-like subunits of complex I could also be a useful tool for the detailed studies of proton pumping.

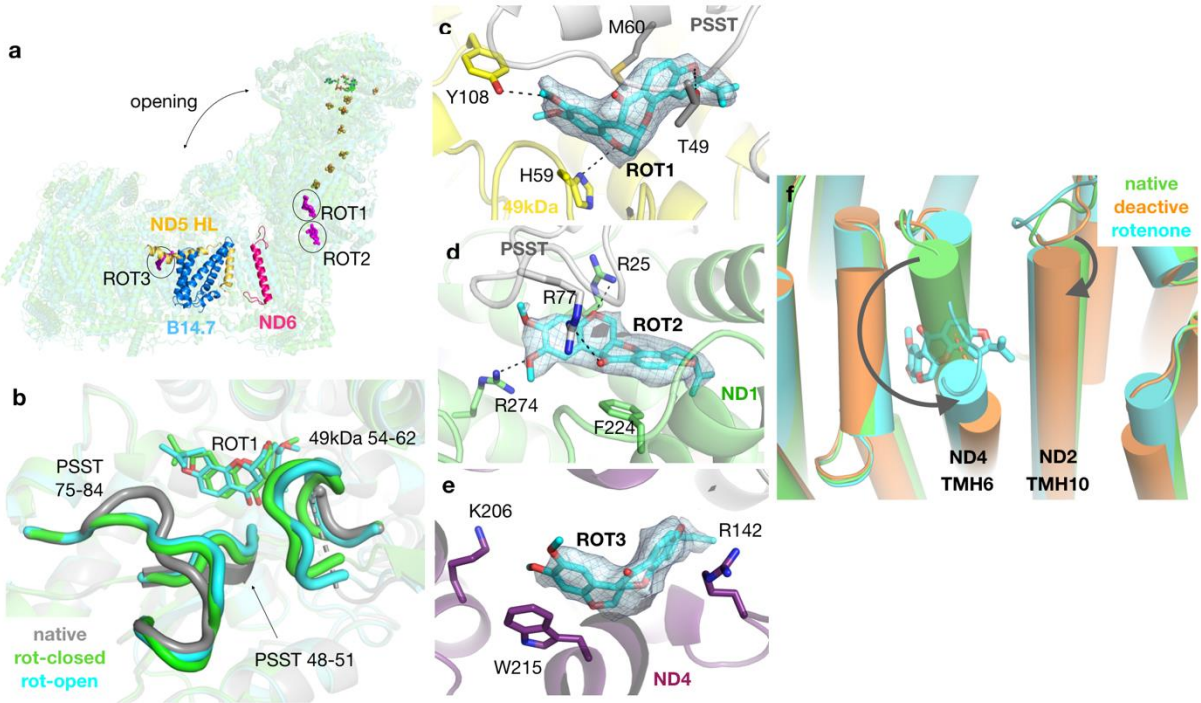


Figure 4.13. Rotenone inhibition of complex I.

a. *The most open classes of rotenone-inhibited and deactive complex I exhibit disorder in the ND5 horizontal and anchor helix as well as in the B14.7 subunit (colored). TMH4 of ND6, which undergoes rearrangements in deactive state is labelled in pink. The three rotenone binding sites are indicated as well; ROT2 binds to open classes only and ROT3 binds to open2 and open3 classes.* ***b.*** *PSST and 49 kDa loops do not change between open (cyan) and closed (green) rotenone states. Native open conformation is shown in grey.* ***c.*** *The deep rotenone binding site: rotenone is within hydrogen binding distance to H59 and Y108 and mimics the conformation of bound quinone during reduction from the N2 cluster.* ***d.*** *The shallow rotenone binding site: rotenone is stabilized by polar interactions with conserved arginines and stacking interactions with F224.* ***e.*** *ND4 rotenone binding site: rotenone is stabilized by positively charged residues, including proton-pumping K206, R142 and by stacking interactions with W216.* ***f.*** *To open the ND4 subunit binding site, the outer TMH6 needs to undergo a slight tilting and rearrangement observed in rotenone open2 and open3 classes (cyan), but also in the most open class of the deactive dataset (orange). Native structure is overlaid in green.*
Adapted from Kampjut & Sazanov, 2020.

Rotenone binding did not change the diameter of the quinone cavity entry or induce massive rearrangements of the quinone cavity as proposed earlier (Haapanen and Sharma, 2018). Bulky rotenone can thus fit, perhaps with local re-arrangements, through the narrow quinone cavity entry site. This casts doubt on the existence of an alternative access pathway to the N2 cluster proposed earlier, as none of the bulky synthetic quinone derivatives, apart from

the largest one, investigated in these studies were significantly larger than rotenone and should therefore be able to fit through the quinone cavity opening (Uno *et al.*, 2019, 2020). As noted above, in the open complex I binding of two rotenone molecules in the quinone cavity ordered the 49 kDa and PSST loops (Figure 4.13b) but not ND1 or ND3 loops, which suggests that ligand binding to the open complex I directly affects the conformations of the PSST and 49 kDa loops, while the ordering of the ND3 and ND1 loops can only happen upon reduction or closing of the complex. Overall, rotenone appears to be unique among complex I inhibitors by having multiple binding sites, which may explain its extremely high affinity for the enzyme.

4.7. Deactivation leads to tilting of the ND6 TMH4 helix, arresting the complex

To better understand the structural rearrangements leading to complex I deactivation, we prepared deactive ovine complex I by incubation without substrates at an increased temperature as described previously (Figure 4.14) (Kotlyar and Vinogradov, 1990) and collected a cryo-EM dataset. The deactive ovine complex I sample did not contain any closed conformation but we could classify four different open classes, which were overall more open than the native open classes. They also had a large portion of the B14.7 subunit disordered and classes 3 and 4 also showed some disorder around the horizontal helix HL and its TM anchor, the N-terminus of the 49 kDa subunit and the TMH5 of ND4 (Figure 4.14a).

The most striking feature of all our deactive classes, however, was the almost complete relocation (by tilting) of the ND6 TMH4 helix, accompanied by the insertion of the TMH3-4 loop between the matrix and membrane arms, as well as unfolding of the 112-118 ND6 β-sheet (Figure 4.14a). TMH3-4 loop gets disordered in the open state and TMH3 acquires a π-bulge, which is likely the "waiting" state of the catalytic cycle when quinone cavity loops get partially and transiently disordered. During prolonged absence of turnover, unfolding proceeds to encompass wider regions of the complex further away from the quinone cavity and results in TMH4 tilt and a stable insertion of ND6 loop between the two arms, preventing any transition into closed conformation and thus deactivating the enzyme. Reversal of this process is consequently slow and requires multiple turnovers. The stable decoupling of matrix and membrane arms of complex I and perhaps also the rearrangements within the ND4 subunit are consistent with the observation that deactive bovine complex I is a sodium proton antiporter (Roberts and Hirst, 2012).

Opening of complex I during deactivation was confirmed using the NEM assay, which is the standard biochemical assay for the deactive complex I and depends on testing the liability

of the ND3/C39 to alkylation (Galkin et al., 2008). However, while ND3 accessibility is a good proxy for the opening of complex I as it assays disorder of the ND3 loop, this assay cannot assess the conformational states of the other loops and we use it here strictly to assay "openness" of complex I. On the other hand, the other known signature A/D feature, the delay in reaching maximal activity upon NADH addition to pre-equilibrated lipid-protein-quinone mixture is indeed a clear indicator of the A/D state of the enzyme. While the deactivated enzyme has consistently shown a significant lag, our standard preparations did not show any delay (Figure 4.14b), confirming that in all our datasets (apart from deactive) the protein represents the active enzyme.

According to the NEM assay the ND3 loop was disordered in virtually all the deactive particles and in ~80% of the native particles, consistent with the structural data (Figure 4.14c). Interestingly, activating isolated complex I with 5 μM NADH (and 200 μM DQ) did not result in a higher proportion of ND3 loop being ordered, also consistent with our structural data from the turnover sample. Since the reaction was allowed to proceed for 20 s before freezing the turnover sample, if activation or catalytic activity were involved in "closing" of the complex, significant difference from the native/NADH closed proportion would be observed. This again confirms that openness of complex I is not correlated with the deactive state, but is an intrinsic property of the sample, with open/closed ratio reflecting the thermodynamically more favorable open state co-existing with the less favorable, but accessible, closed state.

Interestingly, when activating submitochondrial particles, we achieved 40% closed complex I upon activation. This indicates that the lipid membrane is probably necessary for closing of complex I (ordering of the 39 kDa amphipathic helix, lateral tension of the lipid bilayer on the ND1 subunit). If the proton pumping into the IMS happens in the closed step as predicted by our mechanism, closed state should be more populated after the establishment of a high proton motive force.

Our observation of a very distinct deactive state structure is a definite proof that open state of complex I is a part of catalytic cycle and not a deactive form of the enzyme as proposed earlier (Agip *et al.*, 2018; Blaza, Vinothkumar and Hirst, 2018).

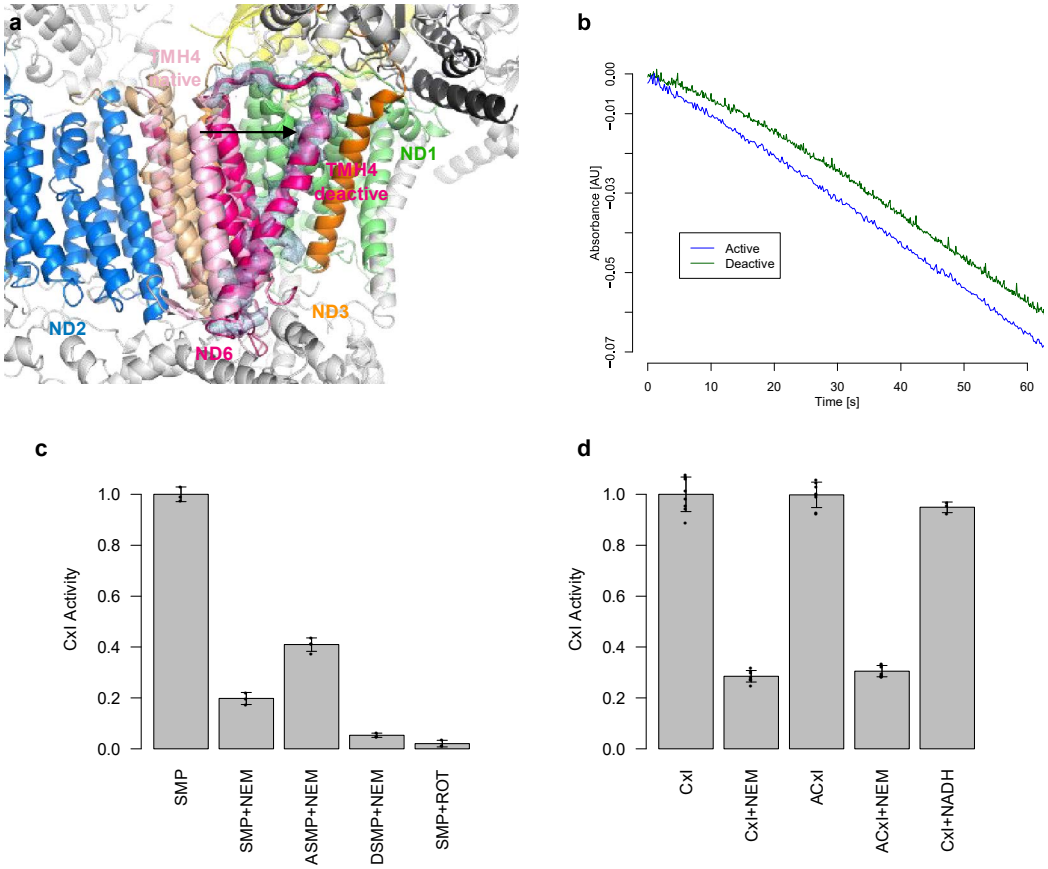


Figure 4.14. Deactivation of complex I.

a. *In deactive complex I ND6-TMH4 gets tilted towards the matrix-membrane arm interface and the TMH3-4 loop inserts between the two arms, stabilizing them in the open conformation.* ***b.*** *Deactive complex I shows a lag in reaching maximum activity. Reaction was started by the addition of NADH to complex I preincubated in lipid and DQ containing buffer.* ***c.*** *Complex I within SMPs is ~20% closed, which increases to ~40% closed upon activation with 5 μM NADH and 200 μM DQ (ASMP), while deactivation of SMPs (DSMP) decreases it to ~0%.* ***d.*** *Activation of purified complex I (ACxI) does not change closed/open ratio. 30 min incubation with 5 mM NADH as in the NADH dataset does not inhibit complex I (CxI+NADH).*
Adapted from Kampjut & Sazanov, 2020.

4.8. Under turnover conditions quinone binds both in the deep and shallow sites

Finally, we collected a dataset of complex I incubated with NADH and decylubiquinone and flash-frozen before the substrates were exhausted. Under such active turnover conditions, one may expect to identify quinone binding modes and perhaps any conformational changes associated with the catalytic cycle, including those in the proton channels, provided that any relevant conformations can be separated by classification and that the enzyme explores the entire conformational space in the absence of membrane and proton-motive force (*pmf*). Such an approach has previously been used to determine the *Y. lipolytica* complex I turnover

structure, albeit at a medium resolution (4.5 Å) which prevented mechanistic insights (Parey *et al.*, 2018).

This approach depends on the assumption that substrate availability during blotting remains the same and that the sample is well mixed and that there is no partitioning of substrates. To circumvent these problems, we included relatively high concentrations of surfactants and substrates, such that not more than 10% of them would be consumed in the time between setting up the reaction and grid freezing. Furthermore, it assumes that there is no preferential binding of some intermediates to carbon or air-water interface and that it is possible to precisely classify rare conformations from the dataset. The fact that rare conformations can be classified has been demonstrated before and in the datasets presented here.

The ratios between closed and open classes remained unchanged in the turnover dataset. Both open and closed classes of this dataset contained a clear density for bound NADH, showing that this is the most populated state of the NADH-binding site during turnover, even though Fe-S clusters and FMN are likely to be mostly reduced since quinol release is a limiting step during complex I turnover (Vinogradov and Grivennikova, 2001; Wikström, Sharma, Ville R I Kaila, *et al.*, 2015).

Closed complex I contained two bound quinones. Quinone in the deep binding pocket (Q_d) was bound between H59 and Y108, exactly where rotenone was observed to bind (Figure 4.15a). We also observed a strong density near the entry to the quinone binding cavity, overlapping with the second rotenone binding site and site #4 predicted by MD, which we modelled as another molecule of decylubiquinone (Figure 4.15a,c). Residues 49 kDa/D160 and H59, which coordinate quinone in the Q_d site had a continuous density between them, indicating the existence of a salt bridge, probably as a result of proton donation to quinone (Figure 4.16a). This density was not present in any of the other conditions (Figure 4.16b,c) and D160 density was weak as is usual for unbound carboxylates in cryo-EM (Wang, 2017). Therefore, it is likely that H59 first donates a proton to reduce quinone and then accepts a proton from D160, so that negative charge resides on D160 and the salt bridge is formed. This is the first experimental demonstration that H59/D160 pair act as a proton donor for the quinone. It also indicates that the species bound in the Q_d site in the turnover dataset is quinol rather than quinone.

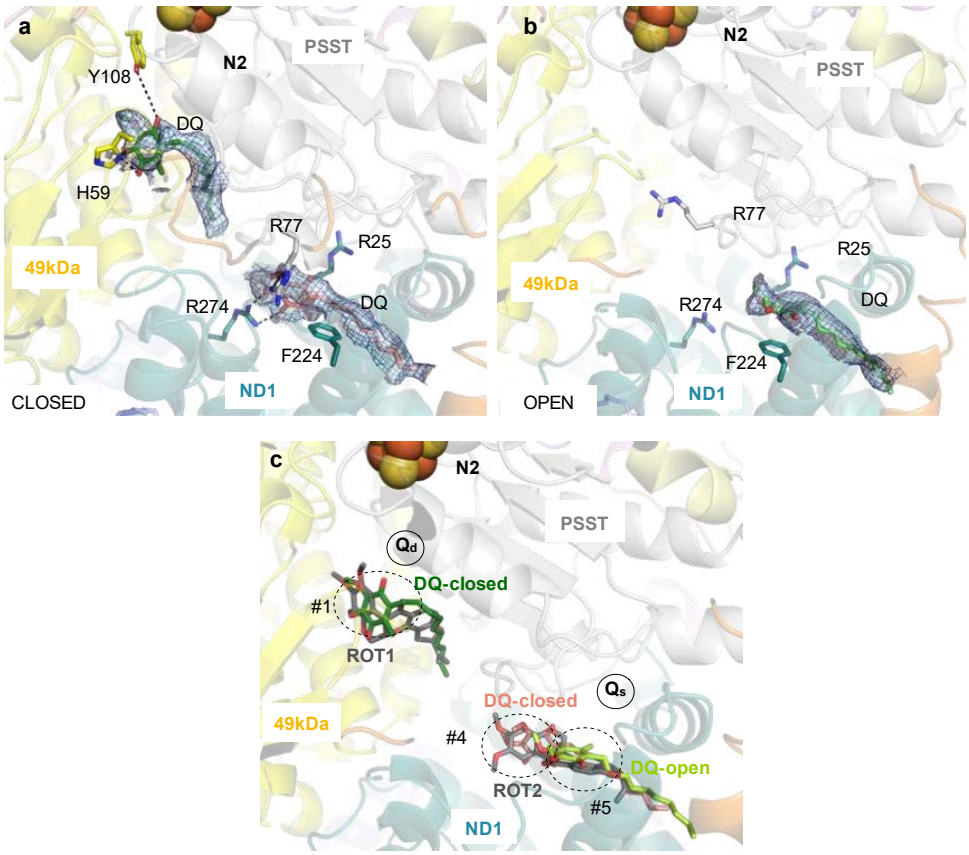


Figure 4.15. Quinone density is observed at different binding sites in open and closed complex I during turnover.

a. *In closed complex I, one quinone molecule is bound in the deep pocket between Y108 and H59 and another in the shallow pocket close to ND1/R25.* ***b.*** *In open complex I, a single quinone is bound at the entry to the quinone binding cavity close to F224.* ***c.*** *Comparison of binding positions of quinone and rotenone in different classes. Adapted from Kampjut & Sazanov, 2020.*

In the open turnover classes a quinone density was present only in the shallow pocket (Q_s) but it was located further outwards compared to the closed class and was overlapping with the site #5 predicted by MD (Figure 4.15b,c). The mobile loops of the quinone binding cavity were disordered to approximately the same extent as in the native open classes and we did not see any density that could correspond to quinone in the deep binding pocket. This suggests that opening of the complex I affects quinone binding and dynamics in the cavity and could facilitate quinol ejection or quinone uptake by the complex. Presence of two molecules of quinone in the closed class is probably an artefact of using a soluble decylubiquinone analogue of ubiquinone with a short carbon tail. If native ubiquinone-10 were bound in Q_d, its isoprenoid tail would clearly prevent quinone binding in Q_s (Figure 4.11). Nevertheless, it gives the indication of the quinone path in the cavity, as the Q_s position subtly differs between open and

closed classes (Figure 4.15c). None of the classes contained a quinone bound in the position #2 as observed in *Y. lipolytica*, or in position #3, as predicted by MD, presumably because these states are much more shortly lived than the others or are species-specific (Parey *et al.*, 2018; Haapanen, Djurabekova and Sharma, 2019).

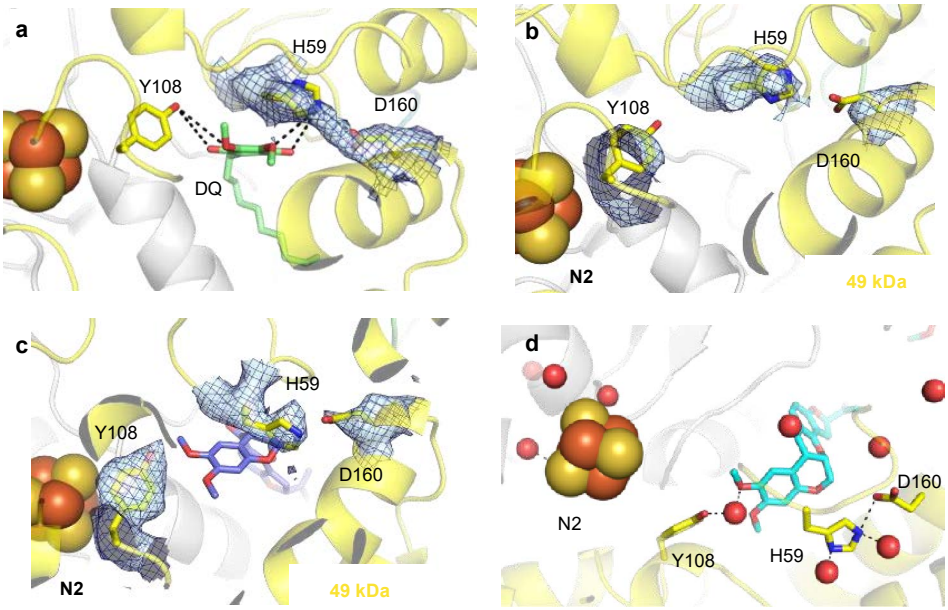


Figure 4.16. Protons for quinone protonation are donated from the H59-D160 pair.

a. *Detailed view on the binding of quinone next to the N2 cluster and the density showing the salt bridge between H59 and E160.* ***b-c.*** *D160 displays weak density and no connection to H59 in all of the classes apart from closed turnover class. NADH closed class (**b**) and rotenone open2 class (**c**) densities are shown for comparison.* ***d.*** *Several water molecules are bound near the deep quinone binding site in the open state.*
Adapted from Kampjut & Sazanov, 2020.

4.9. The role of water molecules in electron transfer and quinone protonation

Abundant water molecules that we observed in the cryoEM density of the peripheral arm of complex I were mostly consistent across the different conditions; hence they are likely to represent tightly bound stationary molecules (Figure 4.17a). Several of them are close to the iron-sulfur clusters, but none lie directly between them, hence it is not likely that they participate in electron tunneling as proposed before from simulations (Hayashi and Stuchebrukhov, 2011). Uniquely from other FeS clusters, N2 contains two water molecules at

hydrogen bonding distance, which may help to elevate its redox potential (Figure 4.17b). We also observed several water molecules near the deep quinone binding site and the Y108-H59-D160 triad in the rotenone open structure, which suggests that in the open state waters can enter this area (Figure 4.16d).

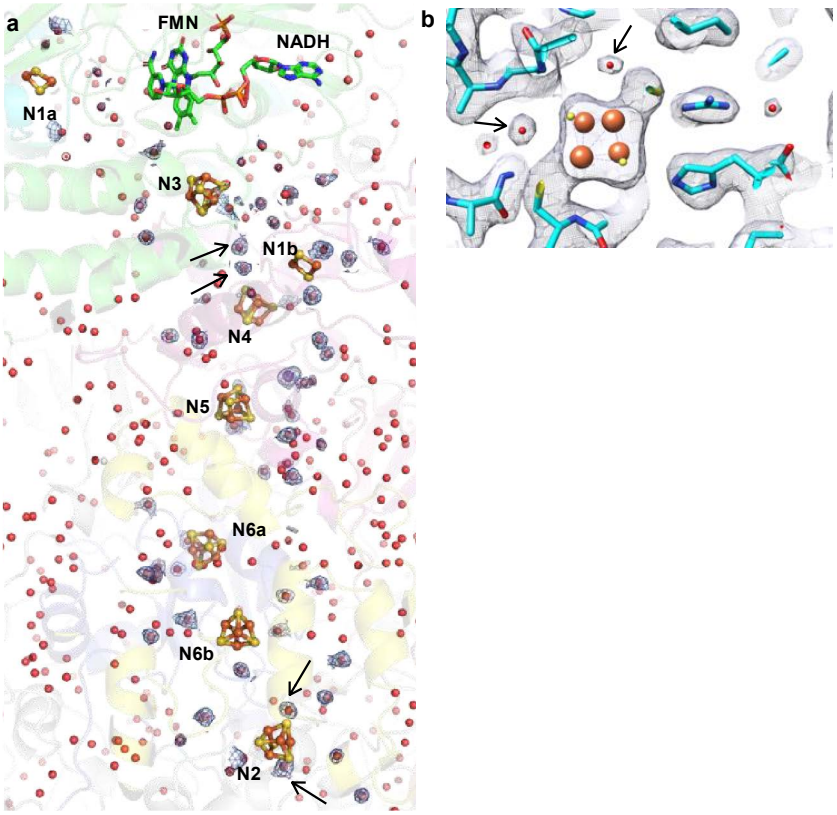


Figure 4.17. Water molecules in the peripheral arm.

a. *Peripheral arm, with water molecules within 10 Å of FeS clusters shown as red sphere with corresponding density, with other identified waters in the background. Two water molecules within hydrogen bonding distance to N2 and two molecules between N3 and N4 clusters (i.e. not in a direct N3-N1b-N4 path) are indicated with arrows.* ***b.*** *Detailed view of waters bound next to N2 cluster.*
Adapted from Kampjut & Sazanov, 2020.

4.10. The role of water molecules in proton pumping

The membrane domain overall was less hydrated than the peripheral arm, with a clear pattern of hydrated matrix and IMS surfaces and abundantly hydrated central hydrophilic axis with relatively dry regions in between (Figure 4.18a). The structures and hydration patterns of the membrane arm of complex I allowed us to experimentally define the proton transfer pathways within complex I for the first time (Figure 4.18a). The total number of identified water molecules was higher in the open state due to higher overall resolution, but the pattern

of distribution of waters did not change in the closed state with one notable exception. All the key protonatable residues along the entire hydrophilic axis from the ND5 tip to the Q cavity are clearly interconnected via many water molecules, with the exception of a long break in the E channel between ND4L/E34 and ND3/D66 in the open state (Figure 4.18a, green arrow). This is caused by the ND6 TMH3 π-bulge with bulky hydrophobic residues such as F68 and M64 blocking the path. Strikingly, when TMH3 rotates and the π-bulge disappears in the closed state, these residues are rotated away and are replaced by the conserved glycines 62-63, creating a water-filled cavity between ND3/D66 and ND4L/E34. This allows for water chain to extend from ND1 all the way to ND2 only in the closed state (Figure 4.18b,c). TMH3 of ND6 is the most conserved TM helix in complex I membrane domain, and even conservative mutations in the π-bulge area are highly detrimental to activity (Efremov and Sazanov, 2011), indicating that identified here breaking and reforming of the water wire is probably essential for the mechanism. Furthermore, three critical glutamates of the E channel, ND1/E143, ND4L/E34 and E70 change their conformations and become charged in the closed complex I as suggested by their density appearance and pattern of hydrogen bond interactions (Figure 4.20). In particular, ND4L/E34 rotates away from ND4L/E70 as they became charged and surrounded by waters. Apart from these most critical changes, many more side-chain and backbone re-arrangements occur in this area upon transition from open to closed state, as shifts of ND1 helices propagate to the E channel.

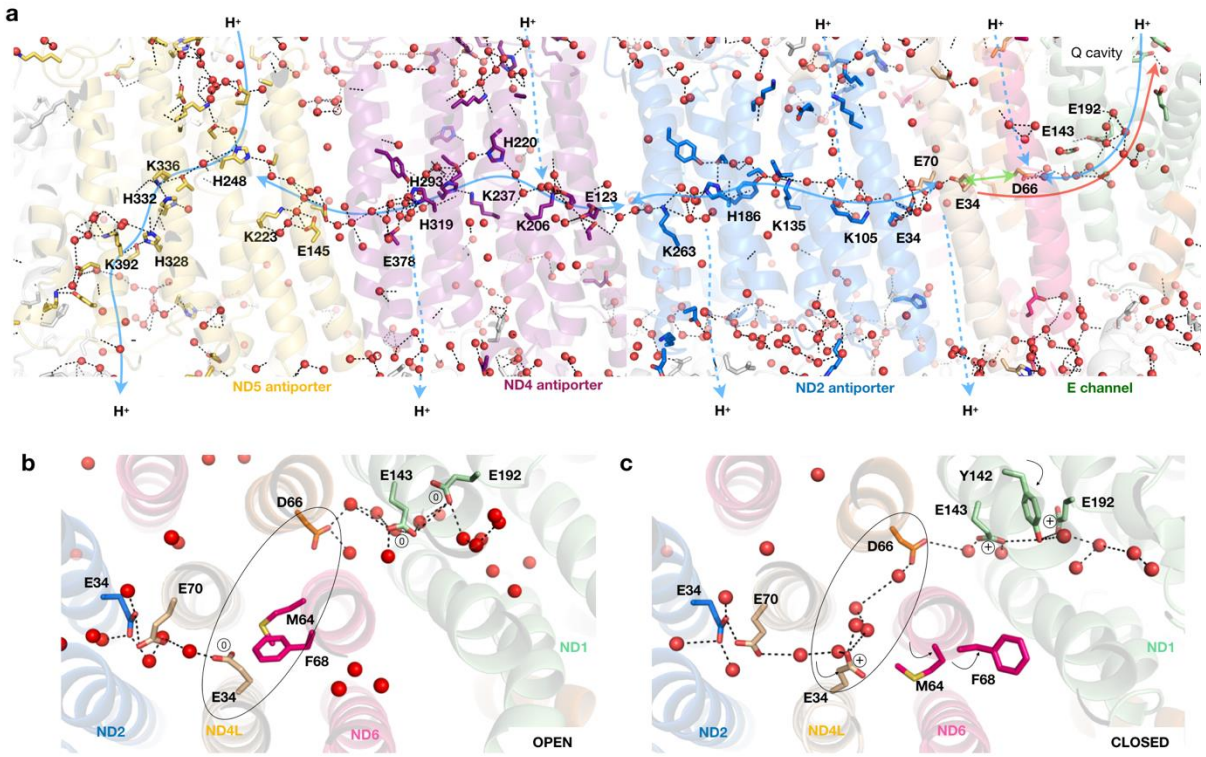


Figure 4.18. Water molecules delineate proton pathways in the membrane arm of complex I.

a. *Central hydrophilic axis connects the three antiporters in the open state of complex I. Both N- and C- terminal Lys/Glu are abundantly hydrated. Water wires towards IMS and matrix side are visible only in ND5. E-channel has a dry region between D66 and E34 and no connection to IMS or matrix side. It is connected to the quinone binding cavity only in the closed state (green arrow), allowing for protons from E34 and E70 to be delivered to Q-site (red arrow).* ***b.*** *Top view of the E-channel in open state of complex I shows that the E-channel dry region is caused by the ND6 π-bulge formation and E34 rotation.* ***c.*** *Continuous water wire in the E-channel of the closed state of complex I.*
Adapted from Kampjut & Sazanov, 2020.

In contrast to the E-channel, antiporters did not change significantly the conformation between closed and open states or between open structures of the rotenone, NADH, turnover and deactive datasets. Analysis of the proton transfer pathways connecting all the protonatable residues (Lys, Glu, His, Asp, Tyr, Ser, Thr) and waters within hydrogen bonding distances (proton hopping via Grotthuss mechanism, with bond slack up to ~4 Å) suggested that E-channel and ND2 and ND4 antiporters lack any connection to the IMS site (Figure 4.18). These parts of the structure are particularly dry and hydrophobic and it is hard to imagine conformational changes which might bring about connections there (some water molecules visible in these areas in Figure 4.18 are actually far away from the key residues). Any clear connections are also lacking on the matrix site, but in principle relatively minor conformational changes there may bring in extra waters allowing connections from this side (dashed H^+ input lines in Figure 4.18).

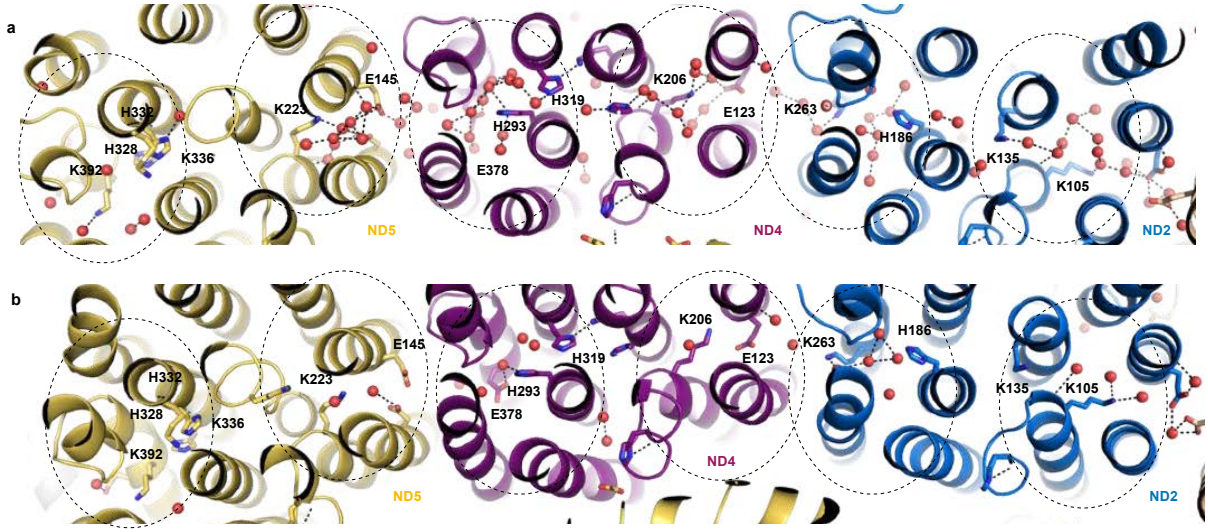


Figure 4.19. Comparison of antiporters in open and closed complex I during turnover.

***a**. Central hydrophilic axis in the open MD class. **b.** Central hydrophilic axis in the closed MD class. Due to lower resolution fewer waters could be modelled but the overall hydration pattern remains the same. Adapted from Kampjut & Sazanov, 2020.*

Notably, the distal antiporter ND5 is radically different – K392 from the C-terminal half-channel is clearly connected to the IMS via the conserved D393 and many waters around. Connection to the matrix side via TM8 H248 also appears to exist with only small barriers, which may be important to prevent back-leak of protons. The only other clear link to the matrix side is via a water wire leading to the quinone cavity in the closed state, suggesting that protons could flow directly from ND1 and the quinone cavity (Figure 4.18). Since this path from ND4L/E34 and E70 towards Q-cavity is isolated from the bulk solvent, it is likely that these residues are de-protonated in the closed state because they donate two protons for the re-protonation of Q-coordinating residues Y108/H59/D160. That way "charge action" of redox reaction is brought in from the Q-site all the way to the first antiporter ND2, revealing a key missing piece in the mechanistic puzzle.

4.11. The role of charge interactions in proton pumping

Even though we did not observe significant conformational changes in the antiporters, a clear pattern of changes in charge distribution of key residues was revealed by the analysis of the cryoEM density. As noted before, charged carboxylates lose side-chain density (except when involved in salt bridges) and additionally, strong hydrogen bonds formed by lysines or glutamates when protonated are well visible in our maps (Figure 4.20). Comparison of open

and closed state MD maps revealed the clearest change in the ND4 antiporter – E123-K206 pair from the N-terminal half-channel has both residues neutral in the open state and both charged in the closed state (Figure 4.20c), while E378 from the C-terminal half-channel changes from charged in open to neutral in closed (Figure 4.20b). For some of the key residues, especially connecting lysines, it was difficult to assign a particular state, however, it was clear that N-terminal Glu-Lys pairs change in anti-phase (ND4 *vs* ND2/ND5) and the overall likely distribution is as depicted in Figure 4.21.

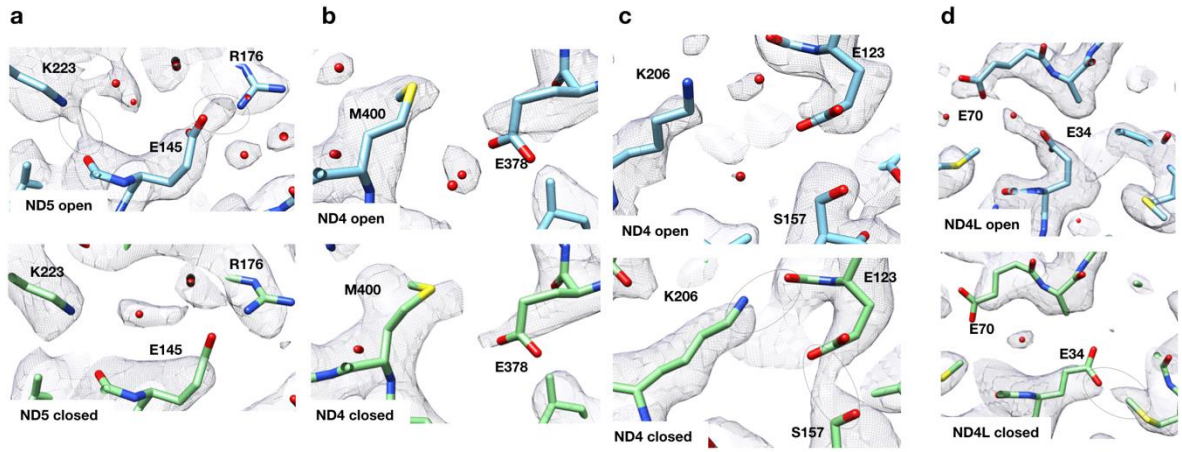


Figure 4.20. Comparison of density and hydrogen bonding patters of key proton-pumping residues between open and closed complex I during turnover.

Breaking of several hydrogen bonds and salt bridges (circled) and changes in glutamate carboxyl densities between open and closed MD. This indicates charge differences between these residues. ***a.*** *N-terminal lysine of ND5.* ***b.*** *C-terminal glutamate of ND4.* ***c.*** *N-terminal lysine of ND4.* ***d.*** *Glutamates of ND4L.*
Adapted from Kampjut & Sazanov, 2020.

The dramatic difference in hydration between ND5 and other antiporters prompted us to re-evaluate a traditional view on proton pumping in complex I, as proposed by us originally (Efremov and Sazanov, 2011; Baradaran *et al.*, 2013), where each antiporter is thought to pump one proton each. Very high proton transfer connectivity along the entire central axis all the way to the exit from ND5 towards IMS, coupled with the absence of IMS connectivity in other antiporters, suggests an alternative possibility. It is feasible that all four pumped protons are ejected via ND5, transported there along the central axis (ND5-all model). Consistently, apart from the key half-channel residues, each antiporter contains additional conserved histidines between the connecting and terminal key residue – H186 in ND2, H293 and H319 in ND4, and H328, H332 and H248 (or K336) in ND5. These increasing amounts of histidines may serve as temporary storage of protons before they can be ejected from ND5 or transferred towards ND5. That way accumulating increasing additional positive change in the C-terminal half-

channel (+2 in ND2, +3 in ND4) may drive ND5 in a mode akin to electrostatic wave model (Verkhovskaya and Bloch, 2013). Mutations of these histidines usually do not lead to dramatic results (Efremov and Sazanov, 2011), however, this is not unexpected as the only effect may be a drop in proton pumping efficiency. Such mutations need to be investigated in more detail in the future and may help to distinguish between ND5-all or one antiporter-one proton models, both of which appear feasible now.

Overall, the fact that under turnover conditions we do not see conformational changes in the antiporters but see instead a change in charge distribution in key residues would argue in support of a electrostatic wave type of the mechanism rather than conformational coupling (Baradaran *et al.*, 2013). However, the mechanism which we propose on the basis of our findings is radically different from the previous electrostatic wave model (Verkhovskaya and Bloch, 2013) because the electric "pulse" in our model originates from proton transfer from the E channel towards quinone as described above. Additionally, conformational coupling clearly plays a key role in the E-channel re-arrangements leading to this proton transfer, so our new model is of mixed type – conformational coupling in the E channel and electrostatic in the antiporters. We do not fully exclude yet a role for conformational coupling in antiporters, as it is possible that full conformational change can be achieved only in the lipid bilayer or in the presence of the proton motive force. Furthermore, conformational changes within antiporters could be temporary and high-energy and correspond to a transition state accessible only when the ND6 π-bulge is being formed (Zhang and Li, 2019).

Table 4.3. Charge distribution in membrane arm of closed and open complex I during turnover.

Densities of the key residues, as well as their hydrogen bonds, salt bridges and the number of coordinated waters were compared between the two states. Legend of symbols: Lys: + protonated, 0 neutral. Glu: 0 neutral, - deprotonated. SB: salt bridge density. HB: hydrogen bond density. SC: side chain. Bb: backbone. Empty cell – no clear information.

	ND5 (chain L)						**ND4 (chain M)**				**ND2 (chain N)**				**ND4L (chain K)**		**ND3 (chain A)**	**ND1 (chain H)**	
	K392	K336	H248	K223	D179	E145	E378	K237	K206	E123	K263	K135	K105	E34	E70	E34	D66	E143	E192
Open state				**+** HB to bb O of W144	**0**	**-** SB to SC of R176	**0** Side chain dens-ity		**0**	**0**	**0** Fewer waters		**+** More waters	**-**	**0**	**0**	**0**	**0**	**0**
Closed state				**0** Fewer waters	**-** SB to SC of R176	**0**	**-**		**+** HB to bb O of F122	**-** HB to SC of S157	**+** Stronger side-chain density		**0**	**0**	**-**	**-** HB to SC of M64	**0**	**-**	**0**

4.12. Mechanism

4.12.1. Deactivation mechanism

Here we described high-resolution structures of complex I in a wide variety of conformations which we used to propose a detailed coupling mechanism for complex I. We showed that the open complex I is a quinone-binding, i.e. "waiting", intermediate, which clarifies why ovine preparation, which exists mostly in the open conformation had high activity and showed no lag in reaching it, indicating that it was in the active state (Figure 4.14) (Letts *et al.*, 2016). Open complex I is thermodynamically more stable hence preparations of complex I which take longer to isolate, involve elevated temperature or employ specific detergents contain more open complex I (Fiedorczuk *et al.*, 2016; Letts, Fiedorczuk and Sazanov, 2016; Agip *et al.*, 2018; Blaza, Vinothkumar and Hirst, 2018). In the absence of substrates, the open waiting state can revert into the fully deactive form by ND6-TMH4 tilting and ND6 loop insertion into the PA-MD interface (Figure 4.14). The high kinetic barrier of active-to-deactive transition in mammals is thus explained by the large-scale reorganization of ND6. This is consistent with previous mouse and bovine "deactive" conformations which showed an extremely weak ND6 TMH4 density, indicating that they were partly, but not fully deactivated (or the fully deactive state class was not separated in 3D classification). They are also consistent with *Y. lipolytica* structures in which ND6 TMH4 is much less likely to shift due to its close association with ND5 and NUJM (homologue of B14.7) subunits. Hence, *Y. lipolytica* complex I quasi-deactive state is actually a waiting open state, in which the ND3 loop is disordered and so is susceptible to alkylation, but the enzyme reverts into the processive state without lag. When NADH or quinone is present, ND6 reorganization in mammals is prevented by reordering of ND1 and 49 kDa loops, which interfere with inserted ND6 loop.

4.12.2. Quinone reduction mechanism

We showed that coordinated conformational changes in the ND1, ND3, ND6, PSST and 49 kDa loops, which are initiated by opening/closing and substrate binding are critical for catalysis. PSST and 49 kDa loops need to be ordered in the retracted state to enable quinone reduction as observed in the turnover closed class. Our rotenone structures showed that binding of a ligand in the deep site is enough to reorder 49 kDa and PSST loops into this retracted conformation even when overall complex I is "open". When quinone is bound, closing could complete fully and involve ND1 and ND3 loops as well (this is supported by the turnover closed class structure, although the enzyme is reduced in this case). The ND1 loop appears to

be a critical coupling element as NADH induces large conformational changes in the ND1 loop, but only when complex I is in the open state. This is consistent with maintaining the correct stoichiometry of the reaction and with the various decoupling mutations (V208L, G214K and Y215H) observed in the ND1 loop (Kirby *et al.*, 2004; Wray *et al.*, 2013). Finally, these changes combined influence the ND3 loop conformation because unfolding or reorganization of the ND1/49 kDa loops clashes with the ordered and inserted ND3 loop (NADH open classes). ND1 and ND3 loops play a structural role in the MD-PA interface and they likely influence the open/closed transition of the complex along with the large tilts of TM helices in ND1, coupling the changes within the quinone cavity to the rest of the enzyme and to the membrane arm. Changes during opening would concomitantly drive quinol ejection as the quinone binding cavity expands at first (any open state in our structures) and then shrinks as the distal part is closed off (NADH open state).

4.12.3. Proton pumping mechanism

Rearrangements in ND1 and ND3 and opening of the whole complex likely lead to the disordering of the ND6 loop and importantly, the formation of the π-bulge in the ND6 TMH3. π-bulge acts as a gatekeeper in the E channel: in the closed state it allows a water wire to extend all the way from Q site to ND2, while in the open state this connection is broken. Its formation also allows deprotonation of critical glutamates in the E channel, ND4L/E34 and E70, as well as ND1/E192. They likely serve as proton donors to quinone (or rather to its coordinating residues). This would give a direct electrostatic pulse from quinone to antiporters, since quinone would take its protons from ND4L, in the vicinity of ND2. Such a "relocation" of charge immediately explains how redox reaction initiates proton translocation.

Analysis of water networks and electron densities in our high-resolution structures allowed us to suggest a novel model for proton translocation in complex I depicted in Figure 4.21. The switch in the orientation/openness of the proton channels is likely associated with the global change in the structure, i.e. between open and closed states. Since the open state is clearly energetically more favorable, it is likely that in this relaxed state the channels are open to the matrix side, so that protein can re-protonate the central axis from the matrix. It is also likely that the energized closed state is associated with proton expulsion into the IMS through the dry MD side and against the proton motive force (Sazanov, 2015). Since there are no obvious links to the matrix side in antiporters other than ND5, it is not clear if all of the possible four input half-channels are operational (dashed lines) although this is likely to be the case

since the protein needs a large area to gather a few available protons at the high pH of the matrix. The links to IMS, apart from ND5, are much less clear which leads to our suggestion that all four protons may be pumped out via ND5 (thus IMS outputs from other antiporters are dashed).

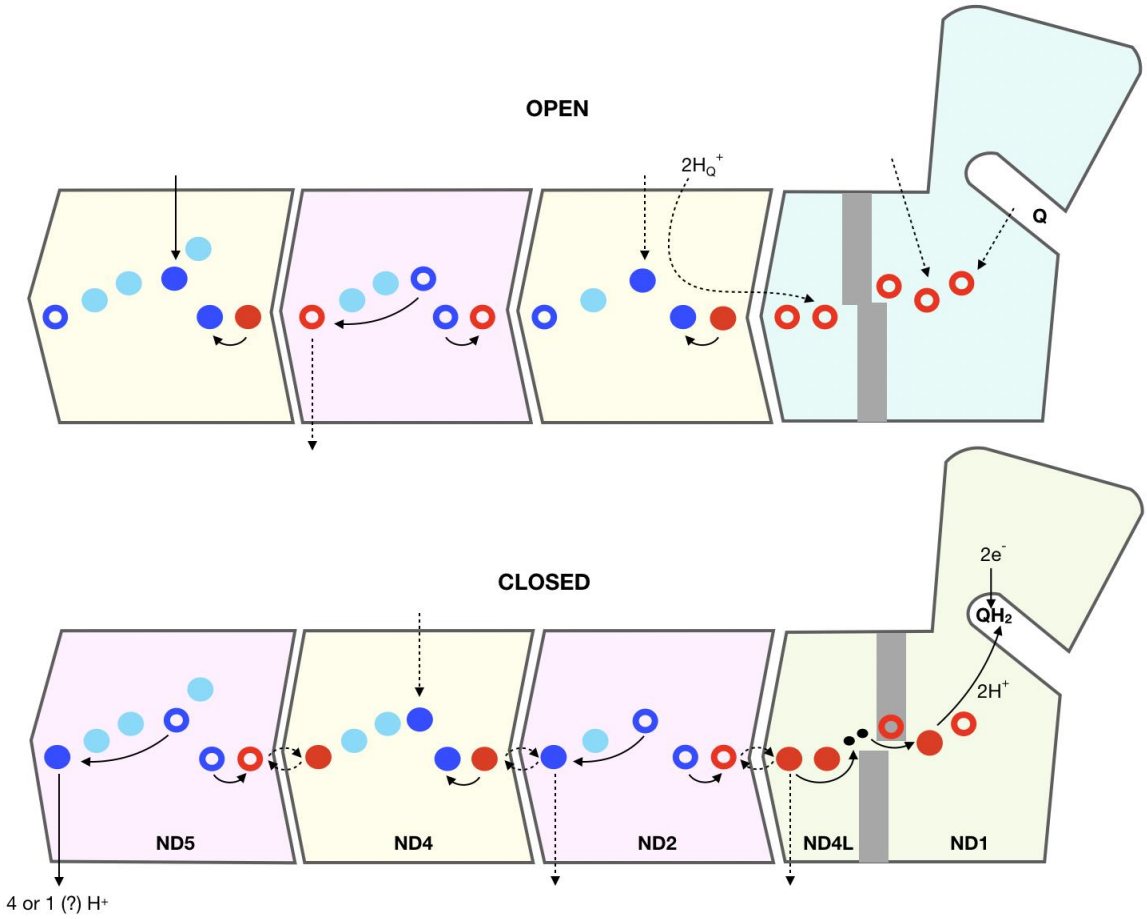


Figure 4.21. The proposed proton pumping mechanism.

Charge distribution of the proton pumping residues differs between the open and the closed states. Coloured circles correspond to glutamate/aspartate (red), lysine (blue) or histidine (cyan) and are shown in the same orientation as in Figure 1c. Full circles represent charged residues and empty circles are neutral (state of histidines is not clear). Arrows show proton movements to achieve the charge distribution as depicted in each state.
Adapted from Kampjut & Sazanov, 2020

Analysis of the putative charges of key residues based on cryoEM density of open and closed turnover states (Table 4.3) suggests an asymmetric distribution of charges between the antiporters, with ND4 being out of sync with ND2 and ND5 (Figure 4.21). In the open state in ND4 the N-terminal Glu-Lys pair is neutral, while in ND2 and ND5 this pair is charged. In the closed state the situation is completely reversed. Consistently, the C-terminal Lys residues in ND2 and ND5 appear not to be protonated in the open state, while the C-terminal Glu in ND4 is protonated. Again, the situation is reverse in the closed state. The state of connecting lysines

is not clear from the density but in the open state they are likely to be protonated from the matrix as noted above. Therefore, it is likely that they donate their protons to protonate the C-terminal Lys/Glu residue in the closed state. One exception is out-of-phase ND4, as it would have to be protonated in the closed state, accordingly with the state of other key residues. It is likely that the exchange of TM12 lysine for glutamate is a primary driver allowing for such an out-of-phase operation. In the case if all antiporters worked in sync, an undesirable excessive accumulation of charges and electrostatic imbalance around the central axis might have occurred, which is prevented by this adaptation. The charge is not clear for all of the key residues from the density itself, however, the distribution in Figure 4.21 is the most consistent with the available observations while ensuring that the overall charge within each subunit and the overall nature of interactions is conserved as described below.

In detail, the pumping cycle starts in the closed state with conformational changes in the ND1 area initiated by quinone binding and reduction. This leads to ND6 TMH3 rotation, freeing up a water-filled cavity linking the Q cavity to ND4L/E34 and E70. These residues donate two protons for re-protonation of D160/H59/Y108, coordinating the quinol. This creates a strong negative charge in ND4L, which is compensated by the nearby ND2/E34 taking up the proton from its partner K105 (Figure 4.18). These N-terminal Glu-Lys pairs do not form a salt bridge in any antiporter in any of the conditions that we observed (as they are slightly too far away from each other), but instead acquire more waters around them when charged. The absence of positive charge on K105 removes the electrostatic barrier and allows the nearby connecting K135 to lose its proton to the C-terminal K263. This positive charge is compensated by the nearby ND4/E123 donating its proton to its partner K206. Connecting ND4/K237 is protonated from the matrix, but the positive charge on K206 prevents the movement of its proton further on to the C-terminal E378. This glutamate loses its proton to the IMS due to electrostatic interactions with charged K237 and K206. ND4/E378 is involved in a charged network with the nearby ND5/R176-D179 and E145-K223 pairs. It is not clear whether R176-D179 pair change their charge state, but negative charge on ND4/E378 would be compensated by ND5/E145 taking up the proton from its partner K223, similarly to the situation on ND4L/ND2 interface. As in ND2, the absence of positive charge on K223 removes the electrostatic barrier and allows the nearby connecting H248 (or K336) to lose a proton to the C-terminal K392, completing a transition from open to closed state charge distribution.

As complex I opens after quinol is ejected into the lipid bilayer, protons residing on the C-terminal residues must be ejected into the IMS (except for ND4). In our scenario the N-terminal Glu-Lys pairs serve mainly as polar/charge switches, controlling proton translocation

from the connecting to the C-terminal residue. As the enzyme reverts to the open state, ND4L/E34 and E70 will be re-protonated, most likely via ND2 as the water wire to the Q site will be broken. At this stage one proton is likely to be ejected from the E channel into IMS, although the exact pathway is not yet clear. Arrival of protons to E34/E70 will cause the nearby ND2/E34 to donate its proton to ND2/K105. As the connecting K135 is protonated from the matrix in the open state, the resulting positive charge will help to drive the proton from the C-terminal lysine ND2/K263 out. However, K263 is not re-protonated again by K135 due to the electrostatic barrier from K105. As the positive charge on K263 is lost, its neighbor ND4/E123 acquires the proton back from K206. Loss of charge on ND4/K206 allows the connecting K237 to lose its proton to E378. As ND4/ E378 acquires a proton, the nearby ND5/E145 will lose its proton to ND5/K223 due to electrostatic interactions. As the connecting ND5/ H248 (or K336) is protonated from the matrix, double positive charge on K223 and H248 (or K336) will drive K392 proton out, re-setting the system to the open state in Figure 4.21. It should be noted that in this scenario simple electrostatic interactions are sufficient to drive proton translocation in a consistent manner across each antiporter, a feature which is allowed by the out-of-sync operation of ND4.

As noted above, the exact output pathways into the IMS are not yet clear. Either all four protons hop over to ND5 via the abundant central axis links using connecting histidines (cyan in Figure 4.21) as a temporary storage (ND5-all path) or they are ejected one by one by each antiporter individually. In the ND5-all option, an accumulating positive charge on histidines (from +1 in E-channel, +2 in ND2 with one His and +3 on ND4 with two His) might in principle provide the necessary driving force to eject four protons from ND5 in one go. However, such accumulation might cause significant electrostatic imbalance in MD, so the other possibility would be for ND5 to eject protons one by one in quick succession as they arrive, with histidines serving as a back-up in case of bottlenecks. Strongly in favor of the ND5-all option is an extreme difference in the IMS-side hydration between ND5 and other antiporters, revealed in this work. An additional argument is that the ND5 subunit, despite being distal, is clearly the most conserved of antiporters. On the other hand, one proton per antiporter option is easier to envisage considering the internal symmetry of antiporters and the electrostatic considerations such as above. Nevertheless, our mechanism outlined above is applicable in both modes of pumping, as the only difference in the case of the ND5-all model would be a re-distribution of protons towards the ND5 subunit (instead of directly into IMS) via many waters and charged residues, including histidines, which link the antiporters in the central axis.

The absence of any conformational changes in the antiporters in turnover conditions argues against a major role of conformational changes in proton transport, the model which we (and others) advocated previously (Baradaran *et al.*, 2013). Although we cannot exclude that such changes might be revealed by studies in liposomes under pmf, the described change in charge distribution allowed us to propose a model which is supported by our new data, is consistent with available mutagenesis data and is fully sufficient to explain the previously enigmatic mechanism of proton pumping.

Ultimately, all of these changes are initiated by the reduction of quinone in Q_d site, as experimentally shown here for the first time, which results in temporary negative charge on D160/H59/Y108 and is transmitted via charged residues into the ND1 loop and the E-channel (Figure 4.15-4.16). The release of redox energy by quinone movements between Q_d and Q_s sites can be quantified in the framework of the low-potential site Q_d and high-potential Q_s site as proposed earlier (Wikström, Sharma, Ville R I Kaila, *et al.*, 2015).

4.12.4. Entire catalytic cycle

On the basis of the above model of proton pumping and the various, many of them novel, conformations of complex I observed here, we propose a first detailed model of the entire catalytic cycle (Figure 4.22). Since the electron transfer rates are fast and the overall reaction is limited by quinone binding/quinol release (Vinogradov and Grivennikova, 2001), we assume that in most steps the enzyme will be kept reduced. The enzyme cycles between the open and the closed conformations, with most steps in the open state, consistent with its prevalence in the data. The catalytic cycle of complex I begins with quinone binding into the cavity of the open enzyme (state 1, represented in our study by open native classes). Quinone binding will promote the transition into the closed conformation in which the ND1 and 49 kDa loops reorder into the “down” and “retracted” conformations and quinone moves into the deep site, where it will get reduced (state 2, closed turnover class). This can only take place when the complex is in the closed conformation and the 49 kDa loop is retracted. Quinone reduction and protonation leads to the accumulation of negative charges on the proton donors, Y108 and D160/H59 pair. The overall result of these redox reactions is a conformational and electrostatic signal towards the E-channel and further on to the antiporters, leading to the ejection of protons into the intermembrane space as described above. This signal is mediated, at least in part, by the ND6 TMH3 rotation and the release of the π-bulge and constitutes the arrival of double negative charge towards the first antiporter ND2 via the de-protonation of key ND4L residues.

Subsequent diffusion of quinol out of the deep site towards the shallow site causes disordering of the 49 kDa, ND1 and ND3 loops as well as re-formation of the π-bulge in the ND6 subunit and the opening of the entire complex (state 3, open turnover classes). At this stage complex I gets fully reduced again and the 49 kDa loop gets extended and the ND1 loop flips upwards as observed in the NADH structures (State 4, open NADH classes). This helps to eject the quinol from the shallow binding site. The re-protonation of the key central axis residues from the matrix side can possibly be happening at States 3, 4 or 1, but we depicted it at State 4 as it is probably the most long-lived state (enzyme waiting for quinone availability). Complex I is thus reset into the "open waiting state" in which it can bind quinone again and restart the catalytic cycle (state 1, native open state). If there is no quinone or NADH present, however, this state will slowly convert into the deactive state in which the ND6 TMH4 tilt and loop insertion between the PA and MD stabilizes the open conformation (state 5, deactive open classes).

This model explains the directionality and tight coupling of the reaction, as (1) NADH-induced changes in the quinone cavity can only happen in the open conformation of the complex, (2) reduction of quinone happens only in the closed state and (3) proton pumping is induced during the transition between the two states. The key novel features of our mechanism are: firstly, the delivery of negative charge from the redox reaction site towards antiporters via concerted conformational changes propagating from the Q site towards the E channel; secondly, the description of how this negative charge drives the proton pumping in the antiporters via electrostatic wave propagation; and, thirdly, the optional, but probable, alternative pathway of proton ejection into IMS only via the ND5 subunit. Thus, our model, based on a wide range of experiments, represents a novel and rather unexpected combination of some features from previously discussed conformational coupling and electrostatic wave-type mechanisms. A consistent mechanistic picture emerges also from our X-ray, cryo-EM and MD studies on the bacterial enzyme (manuscript submitted). What remains to be done in future studies is the elucidation of the fine details of proton translocation processes, potentially by trapping complex I in a transition state, time-resolved studies or structural studies in a membrane under the proton motive force.

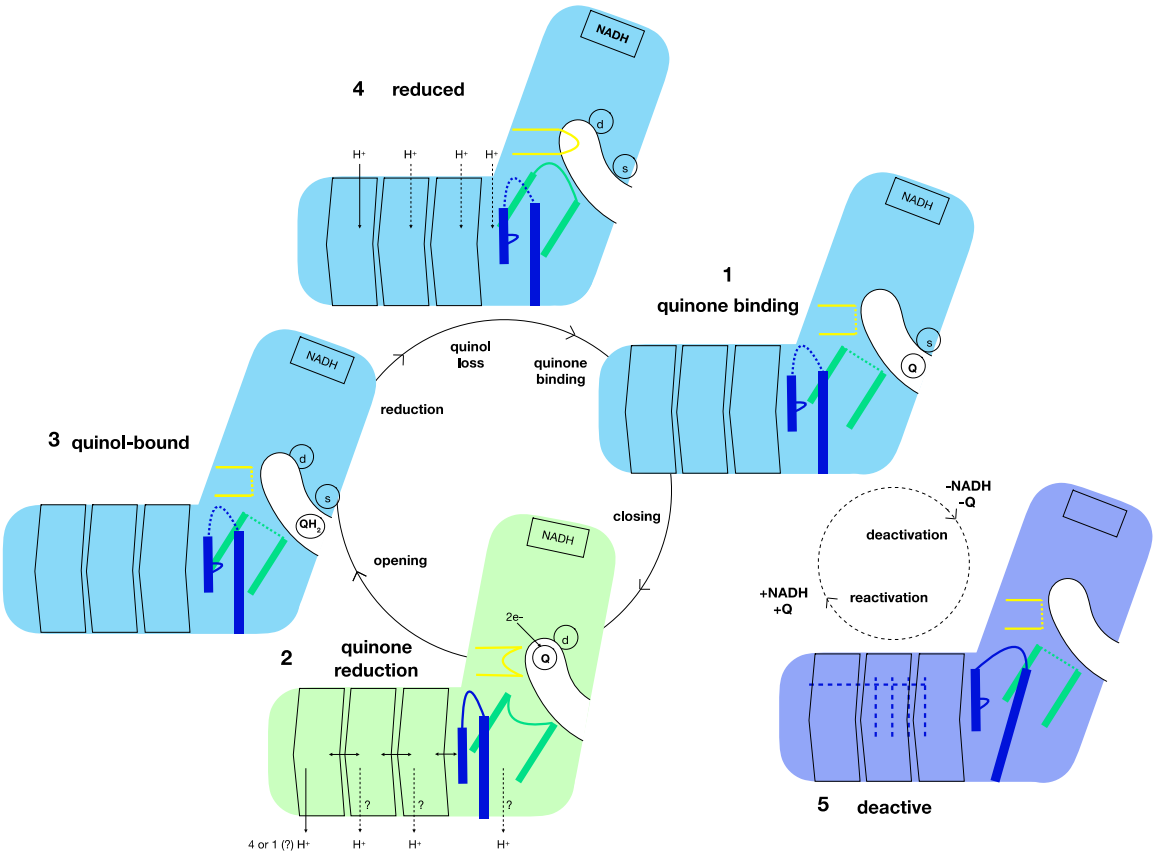


Figure 4.22. Full catalytic cycle of complex I.

Schematic representation of the main catalytic steps in complex I shows how reduction of quinone and its diffusion in and out of the cavity are coupled to loop movements, opening of the complex and pumping of protons. Deactivation pathway branches off from the main catalytic cycle at the quinone binding step when complex I is incubated in the absence of substrates for an extended period of time.
Adapted from Kampjut & Sazanov, 2020

5. Discussion

5.1.Transhydrogenase

5.1.1. Theoretical considerations

Transhydrogenase is a proton-pumping enzyme that operates by an entirely unique mechanism. This is unsurprising because it needs to solve a fundamental biophysical problem associated with its reaction: the midpoint electric potentials of the substrate and product redox pair is the same. The transhydrogenation reaction is fully reversible and depends only on the chemical potentials of substrates and products. Under the usual physiological conditions transhydrogenase produces NADPH and NAD^+ and consumes the pmf. However, bacterial cytosol or mitochondrial matrix usually contain much lower concentrations of transhydrogenase substrates (NADH and $NADP^+$) than its products (NAD^+ and NADPH). To ensure fast kinetics, K_d for the substrates (NADH and $NADP^+$) therefore needs to be much lower than for the products (NAD^+ and NADPH).

This presents a paradox as it seemingly violates the Sabatier principle. This states that the binding of substrates and products to the catalyst needs to be neither too strong nor too weak. If it is too weak, the substrates fail to associate with the enzyme, while if it is too strong, the products fail to dissociate. For biological enzymes it has been shown that binding is in fact the strongest to the transition state, which explains why enzymes speed up reactions (Dill and Bromberg, 2010).

In transhydrogenase, substrate binding must be fairly strong to scavenge the few molecules of $NADP^+$ from the mitochondrial matrix. This is indeed the case for NNT, as K_d for NADP(H) of isolated dIII is in the nanomolar range (Prasad *et al.*, 1999). It is therefore envisaged that the relative binding affinities for nucleotides change during catalysis and are driven by proton motive force-induced conformational changes ("binding-change" mechanism) (Jackson, 2012). $NADP^+$ binds mainly via its adenine, phosphate and ribose moieties, the nicotinamide remains free for reaction during the hydride transfer step. This is obvious from the structures of dI2dIII, which show that surface-exposed nicotinamides from dI and dIII are brought close together during the hydride transfer step. Substrate binding therefore does not need to massively destabilized to ensure sufficient population of the transition state. However, during the nucleotide exchange process, NADPH binding must be strongly destabilised to ensure exchange for $NADP^+$. The destabilisation is achieved by K_d lowering by "opening" of the dIII domain by its interaction with dII.

To ensure that the products do not dissociate before the proton transfer step is complete, the nucleotide exchange process is coupled to proton transfer. dIII can only dissociate after nucleotide has been exchanged and after H664 became protonated. This ensures that there is no “slipping”, as decoupling of these two processing would lead either to a futile cycle between NADH and NADPH or disruption of the proton motive force.

The mechanism for transhydrogenase proposed in this thesis is consistent with and stems from these theoretical considerations and the structures presented explain the functioning of these principles in practice. Pmf (by affecting the protonation state of H664) indeed destabilises products after the hydride and proton transfer steps are complete. Product dissociation before the proton transfer step is complete is prevented by the occlusion of dIII when in the dissociated form and opening of the NADP(H) binding pocket after dIII binds to dII and proton translocation is complete.

Out-of-phase functioning and hydride transfer have been described in earlier work and are fully consistent with the data presented here, even though the hydride transfer step conformation and out-of-step functioning of the two monomers were not observed in our structures.

5.1.2. Work arising

It would be important to trap the mammalian transhydrogenase in more conformations. Trapping mammalian NNT in a pure asymmetric state seems hard to achieve in the light of the mechanism proposed here. Face-up conformation is transient in solution which is a feature of the mechanism. Different mixtures of nucleotides or NADP(H) analogues could stabilize the face-up conformation, but ultimately it is difficult to predict which precise conformation a particular substrate or analogue mixture would stabilize.

A purely asymmetric structure can probably be obtained only in the case of NNT during active turnover, when the two monomers are catalysing reaction out-of-phase. Even then, due to the transience of dI-dIII, the asymmetric conformation might be very short-lived. Perhaps the most likely conditions for the asymmetric state would be turnover with low NADH and high NAD+ concentration, such that NADH binding to dI becomes the limiting step in the entire turnover, and the enzyme waits in the asymmetric state for NADH to bind. In the complete absence of nucleotides, NNT will revert to an auto-inhibited apo structure.

The structures of NNT during active turnover and at high resolution could allow modelling of water molecules and better understanding of the proton translocation mechanism.

It could also give rise to understanding of the ratios of different species during active turnover and identification of rate-determining steps, perhaps under different conditions.

Perhaps most interestingly such a study could help us understand the exact out-of-phase mechanism of NNT and the role of the dimer. To what extent the two monomers are coordinated is the major unsolved question in the NNT mechanism. Structures during active turnover could help to establish this.

Finally, because of the critical differences between the bacterial and mammalian transhydrogenases, the structure of bacterial NNT needs to be determined as well. This will allow the comparison with the mammalian mechanism and determine whether any differences exist. Bacterial NNT can also be mutated and the structures of the mutants solved or further biochemically characterised. In particular, the mutants of the dII-dIII, dI-dIII interfaces, such as R544, N154 or Q175 would be interesting. Secondly, mutants of the proton translocation pathway could help to determine or confirm the roles of some of the less obvious residues in the vicinity of H664, such as S799, T495 or S492. Finally, the structural ramifications of E-loop mutants such as Y1006 would be particularly interesting as they give rise to decoupled NNT without the occluded dIII conformation. Finally, the constructing a covalently linked NNT dimer would allow one to mutate only one of the monomers and the biochemical properties and structures of such mutants could be useful for probing the role of the dimer for the catalysis of NNT reaction.

5.2.Complex I

5.2.1. Principles of coupling in complex I

Similarly to NNT, complex I is also a reversible enzyme, and it can oxidise quinol and reduce NAD^+ if the driving proton-motive force is large enough. The reaction can also occur in vivo but is undesirable as it results in large amount of ROS production at the FMN site and the existence of the deactive state has been argued to be a mechanism to prevent this (Galkin *et al.*, 2008; Babot *et al.*, 2014). Also, S-nitrosylation at the C39 of the ND3 loop has been described to be connected with lowering the ROS production and backflow of electrons (Chouchani *et al.*, 2016).

The mechanism of coupling proposed here is easily reversible which is consistent with the previous observations. High pmf would promote the reverse reaction by promoting alternative antiporter conformations. Translocation of protons would create a negative electrostatic potential (probably localised on the Y108/H59/D160 triad), which would promote

quinol binding and oxidation. Negatively charged Y108/H59/D160 triad would also lower the N2 redox potential enabling reverse electron transfer from N2 to FMN and NAD^+.

As in the NNT mechanism, both electrostatic and conformational changes contribute to the coupling. Slipping is preventing by coupling proton pumping to disordering and ordering of the quinone binding cycle, thus preventing more than one quinone to bind per proton translocation cycle. Furthermore, only one quinone can fit into the cavity at any one time.

5.2.2. Work arising

Further structural work will be required to firmly establish all the intricacies of the model proposed here. In particular it will be important to differentiate between the two proton pumping models proposed in Chapter 4: four protons translocated by ND5 or one proton translocated per antiporter. The main reason for ambiguity at the moment is the different hydration pattern between the antiporter subunits with only ND5 showing clear water wires connecting the matrix and IMS sides of the membrane. It would be interesting to prepare complex I samples at different pH values or in energised proteoliposomes to potentially observe different conformations or hydration patters in the antiporter-like subunits. Additionally, higher resolution structures with only quinone (or quinol) bound could help resolve the question whether quinone/quinol on its own can promote alternative conformations.

Structures of complex I from other organisms at similar resolutions will be interesting to compare and contrast the mechanisms outlined here. Firstly, structures of *Y. lipolytica* and other related complex I proteins with similar ND6 architecture will help to understand the ND6-drive deactivation mechanism proposed here. As the coupling mechanism is probably conserved in bacteria as well as the wider complex I-like protein family, deeper structural understanding of these systems will be important to understand the most conserved features of the complex I structure and mechanism. Additionally, site-directed mutagenesis could provide evidence for many facets of the mechanism presented here. Despite several residues having already been mutated in complex, in particular the N- and C-terminal lysines and several residues of the quinone binding loops, more work is needed, in particular in the light of the specific predictions made by the mechanism here.

Ultimately, some direct observation of proton pumping will be necessary in the future to fully establish the mechanism proposed here. There is potential for time-resolved cryo-EM in this area, but numerous technical difficulties will need to be solved to make this a useful technique. Another branch of investigations in the future are also studies of complex I in situ, e.g. by cryo-electron tomography.

5.3.Grotthuss model of proton pumping

As described in the introduction, proton pumping can be direct or indirect. Direct proton pumping happens when the proton carrier itself is sequentially oxidised and reduced. Indirect proton pumping on the other hand is mediated by proton carriers which do not get themselves oxidised or reduced.

Protons can also follow two paths across the membrane, either they are carried by a membrane-soluble mobile carrier such as quinone or they follow pathways within proteins specialised for proton translocation. As the latter is the situation encountered in complex I and NNT, it warrants a more thorough discussion.

It has been proposed already in 1806 that protons could diffuse along a string of water molecules by hopping from one molecule to another (Figure 5.1). This way, formation of a covalent bond of a proton with the next water molecule and cleavage of a covalent bond with the previous water molecule results in productive translocation of a proton (Cukierman, 2006). This mode of proton translocation is believed to be the dominant mode in proteins and has been described in photosynthetic reaction centres, bacteriorhodopsin, cytochrome c oxidases. It probably occurs also in proton channels (Li *et al.*, 2015). This mode of proton translocation allows a tight control over translocation of single protons, which are otherwise highly mobile in solution (protons have the highest electrical mobility of all biologically important ions and are 10-fold more mobile than cations of alkali metals for example).

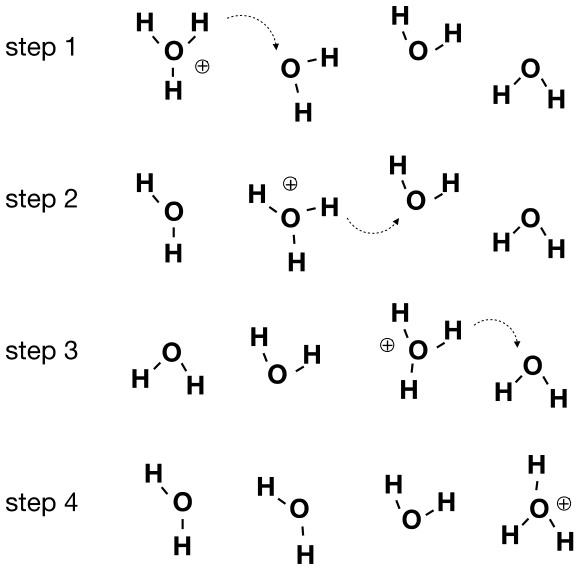


Figure 5.1. Grotthuss mechanism of proton transfer.
Protons do not exist in the free form in solution, but are always bound to water in the form of a hydronium ion H_3O^+.

The water wire in the figure above presents a proton-permeable pathway but it is passive and cannot be controlled. In proteins instead of a long string of waters permeating the membrane, some positions are usually exchanged for protonatable side chains, such as lysine, histidine, aspartate and glutamate or hydroxyl group-containing side chains such as tyrosine, serine and threonine (Figure 5.2). This does not change the underlying mechanism of proton translocation but allows more control over the process and allows the proton pumps to couple proton pumping to other enzymatic processes by adjusting pK_a values of the intervening residues. Aspartate and glutamate have pK_a 4.0 and 4.5 respectively and are negatively charged at neutral pH, while lysine has pKa 10.4 and is positively charged at neutral pH (Fersht, 1985). Histidine has pK_a around 6.0 and can be either neutral or positively charged at neutral pH. The pK_a values depend strongly on the electrostatic environments around the residues and can vary widely (Nakamura, 1996). In Figure 5.2 increasing the pK_a value of glutamate, such that it is neutral in step 1 closes the proton translocation channel. Similarly, increasing the pK_a of lysine such that it is protonated in ground state would stop the proton translocation at the step 3. Fluctuations in protein conformation, electrostatic influences of substrate binding or redox reactions can modulate transfer of protons by changing the pK_a of side chains which are parts of the water wires. This was confirmed both by NNT and complex I structures, which had at least two protonatable "gate-keeping" residues in each proton chain.

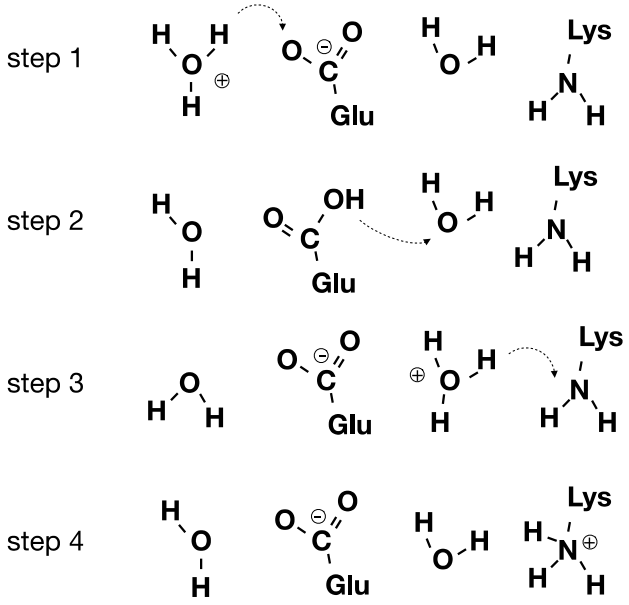


Figure 5.2. Extended Grotthuss mechanism.

pK_a values of hydroxyl-group containing residues are too high for them to exist in stable negatively or positively charged states, hence they can only serve as water-chain mediators but probably not as gate-keepers providing appreciable control over the proton translocating pathways.

The remaining questions in proton pumping mechanism are to what extent the diffusion of hydronium ions and water is allowed along the path, which probably depends on the exact protein environment surrounding them. Deep within the dry regions of NNT proton path and complex I antiporter channels there is probably little diffusion due to steric occlusion. On the other hand, complex I contains quite large water-filled pockets between the antiporters and the E-channel/quinone cavity which can accommodate several water molecules that should in principle be laterally mobile. Whether proton transfer to reduced quinone happens by H3O+ diffusion or by proton hopping along a chain of water molecules will need to be established in future studies.

Another crucial open question with regards to water wires is how transient they are. If there are intervening protonatable protein residues the pKa of which is controlled, then water wires could in principle exist all the time without disrupting the pmf. Indeed, a complete water wire was observed in ND5 subunit of complex I. On the other hand, the formation of the water wire could happen in an otherwise dry region of the protein only at a decisive point during the reaction which would add another layer of control over proton translocation. Molecular dynamics simulations suggests that water wire formation is transient in transhydrogenase dII (Padayatti *et al.*, 2017). Understanding to what extent the water wires are transient will be necessary to distinguish between the two models of proton translocation proposed for complex I: the ND5-all and one-antiporter-one-proton models. Currently it is not possible to say whether the differences in the hydration of antiporters in complex I mean that only ND5 translocates protons or that water wire formation in ND4 and ND2 is transient and was not observed experimentally.

5.4.Redox-coupled proton pumping

5.4.1. Role of electrostatic interactions

Electrostatic interactions are the core of the proton pumping mechanisms in both NNT and complex I. They have also been proposed for photosystem II and cytochrome c oxidase (complex IV) (Ugur, Rutherford and Kaila, 2016; Saura *et al.*, 2019). Because proton gating as described above necessarily requires changes of pK_a of critical residues, electrostatic control is probably a universal principle for redox-coupled proton pumping.

Electrostatic interactions are strong long-range interactions, because the Coulombic interactions between charged objects diminishes with distance only as r^{-1}. Charge separation is therefore rare at the macroscopic scale because it is too high-energy, however, at the microscopic scale it can drive processes. Electrostatic potential energy, U(r) of a single point charge in the presence of another charge can be expressed as $U(r) = C\ q_1\ q_2 / D\ r$, where C is the Coulomb's constant, q_1 and q_2 are the charges, D is a dielectric constant and r is the distance between the charges. Alternatively, at the molecular scale it is useful to think about the above relationship as the Bjerrum length, l_B, which is the charge separation at which the Coulombic electrostatic potential between the charges just equals the thermal energy. $l_B = C\ q_1\ q_2 / D\ k_B\ T$.

In water, which is a polar solvent, electrostatic ions are well shielded, which is reflected by a high dielectric constant of water D = 80 (at room temperature). Dielectric constant reflects 1) the polarizability of atoms and molecules of the solvent themselves, 2) their permanent dipole moment, and 3) the polarizability from a shift in the network of hydrogen bond. Additionally, when salts are dissolved in water, they will also shield electrostatic objects.

In protein cores, the dielectric constant is much smaller, because there is no network of hydrogen bonds, and only polarizability and dipole moments make contributions, and even the dipole moments in protein cores are usually aligned such as in alpha helices, to cancel each other out and assume the most stable lowest energy conformation. D within proteins has been estimated to be around 4.

This difference in dielectric constants means that the Bjerrum length for two charges within the protein at 37ºC is 134 Å, while in water it is only 7 Å. Within a typical protein, nearly all the charges will "feel" each other and will be carefully balanced. Any perturbation to the network will cause long-range changes, which if appropriately tuned and harnessed can be used to drive processes. In NNT, 13 Å separation between the nicotinamide and histidine-664 of opposing or equal charges therefore results in electrostatic potential 10-fold larger than the thermal energy and can drive dIII dissociation. Equally, in complex I, most charged residues along the central axis are within 10 Å and influence each other electrostatically as suggested by the mechanism. The mechanisms suggested in this thesis are thermodynamically feasible. Electrostatic mechanisms probably occur to some extent in all proton pumping, as they provide a large driving force and are easy to generate by adjusting pK_a of protein residues or redox reactions of electron or proton carriers.

5.4.2. Role of cooperativity

Cooperativity is often employed in biology to make transitions between states sharper. The role of cooperativity in complex I and NNT mechanisms is still speculative. Nevertheless, in NNT, the two monomers appear to be negatively cooperatively linked which ensures anti-phase action. The best evidence for this comes from measurements of K_d, which show that binding of a substrate on one monomer increases the K_d for the same substrate on the other monomer (Venning *et al.*, 2000). The structures presented here suggest many possible ways of communications between the monomers that could explain this cooperativity, such as dIII-dIII interaction or regulating dI-dIII formation. In complex I the role of cooperativity is even more speculative, but it is not hard to envisage that the row of three antiporters in complex I membrane domain does display some cooperativity. As the mechanism suggests, once one antiporter switches the conformation, it is more energetically favourable for the other antiporters to switch into the same conformation as well.

In proton pumps in general, the role of cooperativity could perhaps help to ensure the correct stoichiometry between the redox reactions and proton translocation and thus prevent slipping. In NNT, action as a dimer likely adds another layer of control: hydride transfer step in one monomer is coupled to proton transfer in the other and the reaction cannot proceed until both are complete. Thus, the two monomers keep each other in check. Similarly, in complex I, antiporters seem to undergo transitions between the alternative states cooperatively. Apart from slipping this could also ensure the correct stoichiometry of proton pumping in complex I as the reaction can proceed to the next step only after all the antiporters have switched and translocated a proton. This could be especially important for pumps with more than one translocation pathway.

5.4.3. Role of conformational changes

Conformational changes do not necessarily have to have a role in proton pumping, as the Grotthuss mechanism can readily operate on a purely electrostatic basis. However, conformational changes can open and close substrate binding sites and proton translocating pathways. In both complex I and NNT conformational changes play a crucial role. The reason could be that conformational changes can easily provide directionality to the reaction and thus make the system more robust and less prone to errors such a slipping. Additionally, a conformational change mechanism might be easier to evolve and be less sensitive to mutations compared to a purely electrostatic mechanism where finely tuned electrostatic potentials are required.

It seems that a conformational change mechanism is employed to control the substrate exchange in both complex I and transhydrogenase. In complex I quinone binding cavity changes shape dramatically and controls quinone binding in the deep and the shallow sites. Controlled binding to the two sites is connected to proton pumping and thus the coupling between the redox reaction and proton pumping is controlled. If conformational changes were not involved and quinone binding and reduction in the deep site could happen at any time, it is easy to imagine the uncoupling of the redox reaction from proton pumping.

In transhydrogenase, E loop closes over the nucleotide and prevents its dissociation until dIII binds dII and opens the proton channel. Again, substrate exchange is coupled to proton pumping and, in many ways, the two mechanisms are analogous. Should NADP(H) be allowed to exchange freely with the solution, transhydrogenation reaction would easily become decoupled from the proton pumping.

Whether slipping could be prevented by purely electrostatic means is difficult to ascertain. Since complex I and transhydrogenase are completely reversible enzymes, conformational change is perhaps necessary to drive the reaction in a single direction. In irreversible enzymes, however, formation of the product could drive the reaction forward and maintain the correct sequence of events and stoichiometry without the need for large conformational changes.

5.5.Comparison to other redox-coupled proton pumps

Many other redox-coupled proton pumps exist, some distant relatives of the enzymes described here and some completely unrelated. Complex I is part of the large family of proteins, including Mbh, Mrp and Ndh, with shared broad structural features and evolutionary history. Nevertheless, the structures of these related proteins are only now starting to be described. It will be important to see to what extent the mechanism is conserved between these enzymes. Interestingly, they can contain between one and four antiporter-like subunits and thus adjust the stoichiometry of proton pumping. Furthermore, proton pumping modules do not necessarily have to be coupled to quinone reduction but can also be connected to sodium translocation with seemingly minor structural changes in Mrp and it will be interesting to understand how the modular architecture of complex I is compatible with the mechanism.

Transhydrogenase does not have any known homologous proteins, and in particular no known homologues of the proton-translocating dII domain. However, comparison between the mammalian and bacterial versions will almost surely uncover further intricacies of its mechanism.

Among unrelated proton pumps, Complex III is the only mitochondrial proton pump that employs a redox loop mechanism to pump protons across the membrane. Hence, it does not operate by a Grotthuss mechanism, but conformational changes and electrostatics nevertheless play an important role. Firstly, electrostatic potentials of redox groups are carefully tuned to allow electron bifurcation after each oxidation of quinol. Conformational changes, on the other hand, are involved in electron transfer between haems b and c as the swivelling of the Rieske protein shuttles the electrons between the two haemes.

Finally, complex IV is the only "purely electrostatic" proton pump which employs a direct proton pumping mechanism. From the discussion in the preceding chapter it appears that this is only possible because it is also the only enzyme which catalyses an irreversible reaction. Delivery of each electron to the haem a-Cu_B site results in pumping of a proton. Finally, after the delivery of the fourth electron, oxygen double bond is split and reduced, which is an essentially irreversible step of the reaction. This resets the enzyme and kinetically restricts its return to the previous step.

The crucial difference to the mechanisms of transhydrogenase and complex I is that if complex IV fails to couple proton pumping to redox reaction in one cycle, the enzyme is reset before the beginning of the next step and cannot run in the opposite direction. Consumption of pmf to produce oxygen and reduced cytochrome c is not possible. In transhydrogenase or complex I, on the other hand, failure to couple redox and proton pumping reactions, can however result in a futile cycle of substrates or uncoupling of pmf across the inner mitochondrial membrane. Decoupled transhydrogenase would catalyse a reaction between NADH and NADPH until the electric potentials of the two nucleotide pools equilibrated. Even worse, reverse operation of complex I would result in disruption of the proton motive force and ROS production.

This notion could perhaps explain why complex I and transhydrogenase employ both electrostatic and conformational mechanisms for redox reaction coupling to proton pumping. Additionally, cooperativity between the two monomers of transhydrogenase and between the three antiporter-like subunits of complex I could exert further control on the transitions between the conformational states of these enzymes. We are beginning to understand how different unrelated enzymes have evolved similar or different mechanisms to solve the same underlying biophysical problem.

References

Adams, P. D. *et al.* (2010) '*PHENIX* : a comprehensive Python-based system for macromolecular structure solution', *Acta Crystallographica Section D Biological Crystallography*. International Union of Crystallography, 66(2), pp. 213–221. doi: 10.1107/S0907444909052925.

Agip, A. N. A. *et al.* (2018) 'Cryo-em structures of complex i from mouse heart mitochondria in two biochemically defined states', *Nature Structural and Molecular Biology*. Nature Publishing Group, 25(7), pp. 548–556. doi: 10.1038/s41594-018-0073-1.

Albracht, S. P. J., Meijer, A. J. and Rydström, J. (2011) 'Mammalian NADH: Ubiquinone oxidoreductase (Complex I) and nicotinamide nucleotide transhydrogenase (Nnt) together regulate the mitochondrial production of H 2O 2-Implications for their role in disease, especially cancer', *Journal of Bioenergetics and Biomembranes*. Springer US, 43(5), pp. 541–564. doi: 10.1007/s10863-011-9381-4.

Allen, J. F. (2015) 'Why chloroplasts and mitochondria retain their own genomes and genetic systems: Colocation for redox regulation of gene expression', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 112(33), pp. 10231–10238. doi: 10.1073/pnas.1500012112.

Althage, M., Bizouarn, T. and Rydström, J. (2001) 'Identification of a region involved in the communication between the NADP(H) binding domain and the membrane domain in proton pumping E. coli transhydrogenase', *Biochemistry*. American Chemical Society, 40(33), pp. 9968–9976. doi: 10.1021/bi0103157.

Arkblad, E. L. *et al.* (2002) 'Expression of proton-pumping nicotinamide nucleotide transhydrogenase in mouse, human brain and C. elegans', *Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology*, 133(1), pp. 13–21. doi: 10.1016/S1096-4959(02)00107-0.

Attané, C. *et al.* (2016) 'Differential Insulin Secretion of High-Fat Diet-Fed C57BL/6NN and C57BL/6NJ Mice: Implications of Mixed Genetic Background in Metabolic Studies', *PLOS ONE*. Edited by A. Abderrahmani, 11(7), p. e0159165. doi: 10.1371/journal.pone.0159165.

Babot, M. *et al.* (2014) 'ND3, ND1 and 39 kDa subunits are more exposed in the de-active form of bovine mitochondrial complex I', *Biochimica et Biophysica Acta (BBA) - Bioenergetics*. Elsevier, 1837(6), pp. 929–939. doi: 10.1016/J.BBABIO.2014.02.013.

Barad, B. A. *et al.* (2015) 'EMRinger: Side chain-directed model and map validation for 3D cryo-electron microscopy', *Nature Methods*. Nature Publishing Group, 12(10), pp. 943–946. doi: 10.1038/nmeth.3541.

Baradaran, R. *et al.* (2013) 'Crystal structure of the entire respiratory complex I', *Nature*. Nature Research, 494(7438), pp. 443–448. doi: 10.1038/nature11871.

Bayburt, T. H. and Sligar, S. G. (2010) 'Membrane protein assembly into Nanodiscs', *FEBS Letters*, 584(9), pp. 1721–1727. doi: 10.1016/j.febslet.2009.10.024.

Bergkvist, A. *et al.* (2000) 'Interactions of the NADP(H)-binding domain III of proton-translocating transhydrogenase from Escherichia coli with NADP(H) and the NAD(H)-binding domain I studied by NMR and site-directed mutagenesis', *Biochemistry*. American Chemical Society, 39(41), pp. 12595–12605. doi: 10.1021/bi0004091.

Berrisford, J. M. and Sazanov, L. A. (2009) 'Structural Basis for the Mechanism of Respiratory Complex I', *Journal of Biological Chemistry*, 284(43), pp. 29773–29783. doi: 10.1074/jbc.M109.032144.

Bizouarn, T. *et al.* (1997) 'Mutations at tyrosine-235 in the mobile loop region of domain I protein of transhydrogenase from Rhodospirillum rubrum strongly inhibit hydride transfer', *Biochimica et Biophysica Acta - Bioenergetics*, 1320(3), pp. 265–274. doi: 10.1016/S0005-2728(97)00030-3.

Bizouarn, T. *et al.* (2005) 'Nucleotide binding affinities of the intact proton-translocating transhydrogenase from Escherichia coli', *Biochimica et Biophysica Acta - Bioenergetics*, 1708(3), pp. 404–410. doi: 10.1016/j.bbabio.2005.04.004.

Bizouarn, T. and Jackson, J. B. (1993) 'The ratio of protons translocated/hydride ion equivalent transferred by nicotinamide nucleotide transhydrogenase in chromatophores from Rhodospirillum rubrum', *European Journal of Biochemistry*. Blackwell Publishing Ltd, 217(2), pp. 763–770. doi: 10.1111/j.1432-1033.1993.tb18304.x.

Blaza, J. N., Vinothkumar, K. R. and Hirst, J. (2018) 'Structure of the Deactive State of Mammalian Respiratory Complex I', *Structure*. Cell Press, 26(2), pp. 312-319.e3. doi: 10.1016/j.str.2017.12.014.

Blomberg, M. R. A. (2016) 'Mechanism of Oxygen Reduction in Cytochrome c Oxidase and the Role of the Active Site Tyrosine', *Biochemistry*. American Chemical Society, 55(3), pp.

489–500. doi: 10.1021/acs.biochem.5b01205.

Bokori-Brown, M. *et al.* (2016) 'Cryo-EM structure of lysenin pore elucidates membrane insertion by an aerolysin family protein', *Nature Communications*. Nature Publishing Group, 7(1), pp. 1–7. doi: 10.1038/ncomms11293.

Bragg, P. D. and Hou, C. (1999) 'Effect of NBD chloride (4-chloro-7-nitrobenzo-2-oxa-1,3-diazole) on the pyridine nucleotide transhydrogenase of Escherichia coli', *Biochimica et Biophysica Acta - Bioenergetics*, 1413(3), pp. 159–171. doi: 10.1016/S0005-2728(99)00090-0.

Bragg, P. D. and Hou, C. (2001) 'Characterization of mutants of βHistidine91, βAspartate213, and βAsparagine222, possible components of the energy transduction pathway of the proton-translocating pyridine nucleotide transhydrogenase of Escherichia coli', *Archives of Biochemistry and Biophysics*, 388(2), pp. 299–307. doi: 10.1006/abbi.2001.2298.

Brandt, U. (2006) 'Energy Converting NADH: Quinone Oxidoreductase (Complex I)', *Annual Review of Biochemistry*. Annual Reviews, 75(1), pp. 69–92. doi: 10.1146/annurev.biochem.75.103004.142539.

Brandt, U. (2011) 'A two-state stabilization-change mechanism for proton-pumping complex i', in *Biochimica et Biophysica Acta - Bioenergetics*. Elsevier, pp. 1364–1369. doi: 10.1016/j.bbabio.2011.04.006.

Bridges, H. R., Bill, E. and Hirst, J. (2012) 'Mössbauer spectroscopy on respiratory complex I: The iron-sulfur cluster ensemble in the NADH-reduced enzyme is partially oxidized', *Biochemistry*. American Chemical Society, 51(1), pp. 149–158. doi: 10.1021/bi201644x.

Brilot, A. F. *et al.* (2012) 'Beam-induced motion of vitrified specimen on holey carbon film', *Journal of Structural Biology*. J Struct Biol, 177(3), pp. 630–637. doi: 10.1016/j.jsb.2012.02.003.

Cabrera-Orefice, A. *et al.* (2018) 'Locking loop movement in the ubiquinone pocket of complex I disengages the proton pumps', *Nature Communications*. Nature Publishing Group, 9(1), p. 4500. doi: 10.1038/s41467-018-06955-y.

Carroll, J. *et al.* (2005) 'The post-translational modifications of the nuclear encoded subunits of complex I from bovine heart mitochondria.', *Molecular & cellular proteomics : MCP*.

American Society for Biochemistry and Molecular Biology, 4(5), pp. 693–9. doi: 10.1074/mcp.M500014-MCP200.

Carroll, J., Fearnley, I. M. and Walker, J. E. (2006) 'Definition of the mitochondrial proteome by measurement of molecular masses of membrane proteins', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 103(44), pp. 16170–16175. doi: 10.1073/pnas.0607719103.

Chari, A. *et al.* (2015) 'ProteoPlex: stability optimization of macromolecular complexes by sparse-matrix screening of chemical space.', *Nature methods*. Nature Research, 12(9), pp. 859–865. doi: 10.1038/nmeth.3493.

Chen, J. *et al.* (2019) 'Eliminating effects of particle adsorption to the air/water interface in single-particle cryo-electron microscopy: Bacterial RNA polymerase and CHAPSO', *Journal of Structural Biology: X*. Academic Press Inc., 1, p. 100005. doi: 10.1016/j.yjsbx.2019.100005.

Chen, V. B. *et al.* (2010) 'MolProbity: all-atom structure validation for macromolecular crystallography.', *Acta crystallographica. Section D, Biological crystallography*. International Union of Crystallography, 66(Pt 1), pp. 12–21. doi: 10.1107/S0907444909042073.

Chouchani, E. T. *et al.* (2016) 'A unifying mechanism for mitochondrial superoxide production during ischemia-reperfusion injury', *Cell Metabolism*. Cell Press, pp. 254–263. doi: 10.1016/j.cmet.2015.12.009.

Colowick, S. P. *et al.* (1952) 'Pyridine Nucleotide Transhydrogenase. 1. Indirect evidence for the reaction and purification of the enzyme', *Journal of Biological Chemistry*, 195, pp. 95–106. Available at: http://www.jbc.org/content/195/1/95.full.pdf (Accessed: 23 April 2017).

Cooley, R. B., Arp, D. J. and Karplus, P. A. (2010) 'Evolutionary Origin of a Secondary Structure: π-Helices as Cryptic but Widespread Insertional Variations of α-Helices That Enhance Protein Functionality', *Journal of Molecular Biology*. Academic Press, 404(2), pp. 232–246. doi: 10.1016/J.JMB.2010.09.034.

Corpas, F. J. and Barroso, J. B. (2014) 'NADPH-generating dehydrogenases: their role in the mechanism of protection against nitro-oxidative stress induced by adverse environmental conditions', *Frontiers in Environmental Science*. Frontiers Media S.A., 2(DEC), p. 55. doi:

10.3389/fenvs.2014.00055.

Cotton, N. P. J. *et al.* (2001) 'The Crystal Structure of an Asymmetric Complex of the Two Nucleotide Binding Components of Proton-Translocating Transhydrogenase', *Structure*, 9, pp. 165–176. Available at: http://ac.els-cdn.com/S0969212601005718/1-s2.0-S0969212601005718-main.pdf?_tid=749f4b6c-235a-11e7-8a34-00000aab0f26&acdnat=1492425886_fa9791dec2fde4bdb6d55e59a5839311 (Accessed: 17 April 2017).

Cukierman, S. (2006) 'Et tu, Grotthuss! and other unfinished stories', *Biochimica et Biophysica Acta - Bioenergetics*. Elsevier, pp. 876–885. doi: 10.1016/j.bbabio.2005.12.001.

Degli Esposti, M. (1998) 'Inhibitors of NADH-ubiquinone reductase: An overview', *Biochimica et Biophysica Acta - Bioenergetics*. Elsevier, pp. 222–235. doi: 10.1016/S0005-2728(98)00029-2.

Diggle, C. *et al.* (1996) 'Properties of the Purified, Recombinant, NADP(H)-Binding Domain III of the Proton-Translocating Nicotinamide Nucleotide Transhydrogenase from Rhodospirillum Rubrum', *European Journal of Biochemistry*. Blackwell Science Ltd, 241(1), pp. 162–170. doi: 10.1111/j.1432-1033.1996.0162t.x.

Dill, K. and Bromberg, S. (2010) *Molecular Driving Forces: Statistical Thermodynamics in Chemistry, Physics, Biology, and Nanoscience*. 2nd edn. Garland Science.

Dogar, I. *et al.* (2020) 'C57BL/6J mice upregulate catalase to maintain the hydrogen peroxide buffering capacity of liver mitochondria', *Free Radical Biology and Medicine*. Elsevier Inc., 146, pp. 59–69. doi: 10.1016/j.freeradbiomed.2019.10.409.

Dröse, S., Galkin, A. and Brandt, U. (2009) 'Measurement of Superoxide Formation by Mitochondrial Complex I of Yarrowia Lipolytica', *Methods in Enzymology*, pp. 475–490. doi: 10.1016/S0076-6879(08)04426-1.

Dröse, S., Stepanova, A. and Galkin, A. (2016) 'Ischemic A/D transition of mitochondrial complex i and its role in ROS generation', *Biochimica et Biophysica Acta - Bioenergetics*. The Authors, 1857(7), pp. 946–957. doi: 10.1016/j.bbabio.2015.12.013.

Dzioba-Winogrodzki, J. *et al.* (2009) 'The Vibrio cholerae Mrp system: Cation/proton antiport properties and enhancement of bile salt resistance in a heterologous host', *Journal of Molecular Microbiology and Biotechnology*. J Mol Microbiol Biotechnol, 16(3–4), pp. 176–

186. doi: 10.1159/000119547.

Earley, F. G. P. *et al.* (1987) 'Photolabelling of a mitochondrially encoded subunit of NADH dehydrogenase with [3 H]dihydrorotenone', *FEBS Letters*. No longer published by Elsevier, 219(1), pp. 108–112. doi: 10.1016/0014-5793(87)81200-0.

Earley, F. G. and Ragan, C. I. (1984) 'Photoaffinity labelling of mitochondrial NADH dehydrogenase with arylazidoamorphigenin, an analogue of rotenone.', *The Biochemical journal*, 224(2), pp. 525–534. doi: 10.1042/bj2240525.

Efremov, R. G. and Sazanov, L. A. (2011) 'Structure of the membrane domain of respiratory complex i', *Nature*. Nature Publishing Group, 476(7361), pp. 414–421. doi: 10.1038/nature10330.

Efremov, R. G. and Sazanov, L. A. (2012) 'The coupling mechanism of respiratory complex i - A structural and evolutionary perspective', in *Biochimica et Biophysica Acta - Bioenergetics*. Elsevier, pp. 1785–1795. doi: 10.1016/j.bbabio.2012.02.015.

Elurbe, D. M. and Huynen, M. A. (2016) 'The origin of the supernumerary subunits and assembly factors of complex I: A treasure trove of pathway evolution', *Biochimica et Biophysica Acta - Bioenergetics*. Elsevier B.V., 1857(7), pp. 971–979. doi: 10.1016/j.bbabio.2016.03.027.

Emsley, P. *et al.* (2010) 'Features and development of Coot.', *Acta crystallographica. Section D, Biological crystallography*. International Union of Crystallography, 66(Pt 4), pp. 486–501. doi: 10.1107/S0907444910007493.

Esfandi, F. *et al.* (2019) 'Downregulation of nicotinamide nucleotide transhydrogenase and its naturally occurring antisense RNA in gastric cancer', *Asia-Pacific Journal of Clinical Oncology*. Blackwell Publishing Ltd, 15(5), pp. e191–e196. doi: 10.1111/ajco.13230.

Euro, L. *et al.* (2008) 'Electrostatic interactions between FeS clusters in NADH:Ubiquinone oxidoreductase (complex I) from Escherichia coli', *Biochemistry*. American Chemical Society, 47(10), pp. 3185–3193. doi: 10.1021/bi702063t.

Evgeny, K. (2010) 'Crystal contacts as nature's docking solutions', *Journal of Computational Chemistry*, 31(1), pp. 133–143. doi: 10.1002/jcc.21303.

Eytan, G. D., Eytan, E. and Rydström, J. (1987) 'Energy-linked nicotinamide-nucleotide transhydrogenase. Light-driven transhydrogenase catalyzed by transhydrogenase from beef

heart mitochondria reconstituted with bacteriorhodopsin.', *Journal of Biological Chemistry*, 262(11), pp. 5015–5019. Available at: http://www.jbc.org/content/262/11/5015.full.pdf (Accessed: 8 August 2017).

Eytans, G. D. *et al.* (1987) 'Energy-linked Nicotinamide-nucleotide Transhydrogenase. Characterization of reconstituted ATP-driven transhydrogenase from beef heart mitochondria', *THE JOURNAL OF BIOLOGICAL CHEMISTRY*, 262(11), pp. 5008–5014. Available at: http://www.jbc.org/content/262/11/5008.full.pdf (Accessed: 8 August 2017).

Feja, B. and Aebi, U. (1999) 'Determination of the inelastic mean free path of electrons in vitrified ice layers for on-line thickness measurements by zero-loss imaging', *Journal of Microscopy*. John Wiley & Sons, Ltd, 193(1), pp. 15–19. doi: 10.1046/j.1365-2818.1999.00436.x.

Fergusson, G. *et al.* (2014) *Defective insulin secretory response to intravenous glucose in C57Bl/6J compared to C57Bl/6N mice*, *Molecular Metabolism*. doi: 10.1016/j.molmet.2014.09.006.

Fiedorczuk, K. *et al.* (2016) 'Atomic structure of the entire mammalian mitochondrial complex I', *Nature*. Nature Publishing Group, pp. 1–21. doi: 10.1038/nature19794.

Fiedorczuk, K. (2017) *Cryo-electron microscopy studies on ovine mitochondrial complex I*. University of Cambridge. Available at: https://www.repository.cam.ac.uk/handle/1810/270318 (Accessed: 2 May 2020).

Fiedorczuk, K. and Sazanov, L. A. (2018) 'Mammalian Mitochondrial Complex I Structure and Disease-Causing Mutations', *Trends in Cell Biology*. Elsevier Current Trends, pp. 835–867. doi: 10.1016/j.tcb.2018.06.006.

Finel, M. (1998) 'Organization and evolution of structural elements within complex I', *Biochimica et Biophysica Acta - Bioenergetics*. Elsevier B.V., 1364(2), pp. 112–121. doi: 10.1016/S0005-2728(98)00022-X.

Fjellstrom, O. *et al.* (1999) 'Catalytic Properties of Hybrid Complexes of the NAD (H) -Binding and NADP (H) -Binding Domains of the Proton-Translocating Transhydrogenases from Escherichia coli and Rhodospirillum rubrum †'. American Chemical Society, pp. 415–422. doi: 10.1021/bi9817111.

Fjellström, O., Johansson, C. and Rydström, J. (1997) 'Structural and catalytic properties of

the expressed and purified NAD(H)- and NADP(H)-binding domains of proton-pumping transhydrogenase from Escherichia coli', *Biochemistry*. American Chemical Society, 36(38), pp. 11331–11341. doi: 10.1021/bi970958f.

Flück, C. E. (2017) 'Mechanisms in endocrinology: Update on pathogenesis of primary adrenal insufficiency: Beyond steroid enzyme deficiency and autoimmune adrenal destruction', *European Journal of Endocrinology*. BioScientifica Ltd., pp. R99–R111. doi: 10.1530/EJE-17-0128.

Fontaine, E. (2018) 'Metformin-Induced Mitochondrial Complex I Inhibition: Facts, Uncertainties, and Consequences', *Frontiers in Endocrinology*. Frontiers Media SA, 9. doi: 10.3389/fendo.2018.00753.

Francisco, A. *et al.* (2018) 'Nicotinamide nucleotide transhydrogenase is required for brain mitochondrial redox balance under hampered energy substrate metabolism and high-fat diet', *Journal of Neurochemistry*. John Wiley & Sons, Ltd (10.1111), 147(5), pp. 663–677. doi: 10.1111/jnc.14602.

Frotscher, E. *et al.* (2015) 'A fluorinated detergent for membrane-protein applications', *Angewandte Chemie - International Edition*, 54(17), pp. 5069–5073. doi: 10.1002/anie.201412359.

Fu, Q., Ma, R. and Fioravanti, C. F. (2019) 'Purification of Adult Hymenolepis diminuta (CESTODA) Mitochondrial NADPH→NAD+ Transhydrogenase', *Journal of Parasitology*. American Society of Parasitologists, 105(2), p. 321. doi: 10.1645/18-55.

Gabaldón, T., Rainey, D. and Huynen, M. A. (2005) 'Tracing the evolution of a large protein complex in the eukaryotes, NADH:ubiquinone oxidoreductase (Complex I)', *Journal of Molecular Biology*. Academic Press, 348(4), pp. 857–870. doi: 10.1016/j.jmb.2005.02.067.

Galemou Yoga, E. *et al.* (2019) 'Mutations in a conserved loop in the PSST subunit of respiratory complex I affect ubiquinone binding and dynamics', *Biochimica et Biophysica Acta - Bioenergetics*. Elsevier, 1860(7), pp. 573–581. doi: 10.1016/j.bbabio.2019.06.006.

Galkin, A. *et al.* (2008) 'Identification of the mitochondrial ND3 subunit as a structural component involved in the active/deactive enzyme transition of respiratory complex I', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology, 283(30), pp. 20907–20913. doi: 10.1074/jbc.M803190200.

Gargari, S. S. *et al.* (2019) 'Transcription Levels of nicotinamide nucleotide transhydrogenase and Its Antisense in Breast Cancer Samples', *Cell Journal*. Royan Institute (ACECR), 21(3), pp. 331–336. doi: 10.22074/cellj.2019.6238.

Glaeser, R. M. *et al.* (2016) 'Factors that Influence the Formation and Stability of Thin, Cryo-EM Specimens', *Biophysical Journal*. Biophysical Society, pp. 749–755. doi: 10.1016/j.bpj.2015.07.050.

Glaeser, R. M. (2016) 'Specimen Behavior in the Electron Beam', in *Methods in Enzymology*. Academic Press Inc., pp. 19–50. doi: 10.1016/bs.mie.2016.04.010.

Glavas, N. A. and Bragg, P. D. (1995) 'The mechanism of hydride transfer between NADH and 3-acetylpyridine adenine dinucleotide by the pyridine nucleotide transhydrogenase of Escherichia coli', *Biochimica et Biophysica Acta*, 1231, pp. 297–303. Available at: http://ac.els-cdn.com/0005272895000892/1-s2.0-0005272895000892-main.pdf?_tid=de77969c-2d1a-11e7-8338-00000aab0f26&acdnat=1493498088_b56725d6e6bad353a15d3bdd879552a1 (Accessed: 29 April 2017).

Glavas, N. A., Hou, C. and Bragg, P. D. (1995) 'Involvement of Histidine-91 of the β Subunit in Proton Translocation by the Pyridine Nucleotide Transhydrogenase of Escherichia coli', *Biochemistry*. American Chemical Society, 34(23), pp. 7694–7702. doi: 10.1021/bi00023a016.

Goujon, M. *et al.* (2010) 'A new bioinformatics analysis tools framework at EMBL-EBI', *Nucleic Acids Research*. Oxford University Press, 38(SUPPL. 2), pp. W695–W699. doi: 10.1093/nar/gkq313.

Grant, T., Rohou, A. and Grigorieff, N. (2018) 'CisTEM, user-friendly software for single-particle image processing', *eLife*. eLife Sciences Publications Ltd, 7. doi: 10.7554/eLife.35383.

Grivennikova, V. G. *et al.* (1997) 'Interaction of the mitochondrial NADH-ubiquinone reductase with rotenone as related to the enzyme active/inactive transition', *Biochimica et Biophysica Acta - Bioenergetics*. Elsevier, 1319(2–3), pp. 223–232. doi: 10.1016/S0005-2728(96)00163-6.

Groß, C. J. *et al.* (2016) 'K+ Efflux-Independent NLRP3 Inflammasome Activation by Small

Molecules Targeting Mitochondria', *Immunity*, 45(4), pp. 1–13. doi: 10.1016/j.immuni.2016.08.010.

Gu, Y. *et al.* (2019) 'High expression of long non-coding RNA NNT-AS1 facilitates progression of cholangiocarcinoma through promoting epithelial-mesenchymal transition.', *American journal of translational research*. e-Century Publishing Corporation, 11(9), pp. 5438–5456. Available at: http://www.ncbi.nlm.nih.gov/pubmed/31632521 (Accessed: 12 February 2020).

Guerrero-Castillo, S. *et al.* (2017) 'The Assembly Pathway of Mitochondrial Respiratory Chain Complex I', *Cell Metabolism*. Cell Press, 25(1), pp. 128–139. doi: 10.1016/j.cmet.2016.09.002.

Guo, R. *et al.* (2017) 'Architecture of Human Mitochondrial Respiratory Megacomplex I2III2IV2', *Cell*. Cell Press, 170(6), pp. 1247-1257.e12. doi: 10.1016/j.cell.2017.07.050.

Haapanen, O., Djurabekova, A. and Sharma, V. (2019) 'Role of Second Quinone Binding Site in Proton Pumping by Respiratory Complex I', *Frontiers in Chemistry*. Frontiers, 7, p. 221. doi: 10.3389/fchem.2019.00221.

Haapanen, O. and Sharma, V. (2018) 'A modeling and simulation perspective on the mechanism and function of respiratory complex I', *Biochimica et Biophysica Acta - Bioenergetics*. Elsevier, pp. 510–523. doi: 10.1016/j.bbabio.2018.04.001.

Hauer, F. *et al.* (2015) 'GraDeR: Membrane Protein Complex Preparation for Single-Particle Cryo-EM', *Structure*. Elsevier Ltd, 23(9), pp. 1769–1775. doi: 10.1016/j.str.2015.06.029.

Hayashi, T. and Stuchebrukhov, A. A. (2011) 'Electron tunneling in respiratory complex i', *Journal of Physical Chemistry B*. National Academy of Sciences, 115(18), pp. 5354–5364. doi: 10.1021/jp109410j.

He, W., Zhang, Y. and Xia, S. (2020) 'LncRNA NNT-AS1 promotes non-small cell lung cancer progression through regulating miR-22-3p/YAP1 axis', *Thoracic Cancer*. John Wiley and Sons Inc., pp. 1759-7714.13280. doi: 10.1111/1759-7714.13280.

Heiker, J. T. *et al.* (2013) 'Nicotinamide nucleotide transhydrogenase mRNA expression is related to human obesity', *Obesity*, 21(3), pp. 529–534. doi: 10.1002/oby.20095.

Heinemann, M. *et al.* (2016) 'Distinct transcriptional regulation of the two Escherichia coli transhydrogenases PntAB and UdhA', *Microbiology*. Microbiology Society, 162(9), pp.

1672–1679. doi: 10.1099/mic.0.000346.

Henderson, R. (1995) ‘The Potential and Limitations of Neutrons, Electrons and X-Rays for Atomic Resolution Microscopy of Unstained Biological Molecules’, *Quarterly Reviews of Biophysics*. Cambridge University Press, 28(2), pp. 171–193. doi: 10.1017/S003358350000305X.

Henderson, R. and Unwin, P. N. T. (1975) ‘Three-dimensional model of purple membrane obtained by electron microscopy’, *Nature*. Nature, 257(5521), pp. 28–32. doi: 10.1038/257028a0.

Ho, H.-Y. *et al.* (2017) ‘Nicotinamide nucleotide transhydrogenase (NNT) deficiency dysregulates mitochondrial retrograde signaling and impedes proliferation’. doi: 10.1016/j.redox.2017.04.035.

Hu, X. *et al.* (1995) ‘Characterization of the interaction of NADH with proton pumping E. coli transhydrogenase reconstituted in the absence and in the presence of bacteriorhodopsin’, *BBA - Bioenergetics*, 1229(1), pp. 64–72. doi: 10.1016/0005-2728(94)00187-A.

Hu, X. *et al.* (1999) ‘Site-directed mutagenesis of charged and potentially proton-carrying residues in the β subunit of the proton-translocating nicotinamide nucleotide transhydrogenase from escherichia coli. Characterization of the βH91, βD392, and βK424 mutants’, *Biochemistry*. American Chemical Society, 38(5), pp. 1652–1658. doi: 10.1021/bi981654b.

Hu, X., Zhang, J. and Rydström, J. (1998) ‘Interactions of reduced and oxidized nicotinamide mononucleotide with wild-type and αD195E mutant proton-pumping nicotinamide nucleotide transhydrogenases from Escherichia coli’, *Biochimica et Biophysica Acta - Bioenergetics*, 1367(1–3), pp. 134–138. doi: 10.1016/S0005-2728(98)00141-8.

Huang, L. *et al.* (2019) ‘Long non-coding RNA NNT-AS1 functions as an oncogenic gene through modulating MIR-485/ BCl9 in cholangiocarcinoma’, *Cancer Management and Research*. Dove Medical Press Ltd, 11, pp. 7739–7749. doi: 10.2147/CMAR.S207801.

Hutton, M. *et al.* (1994) ‘Kinetic resolution of the reaction catalysed by proton-translocating transhydrogenase from Escherichia coli as revealed by experiments with analogues of the nucleotide substrates’, *European Journal of Biochemistry*. Blackwell Publishing Ltd, 219(3), pp. 1041–1051. doi: 10.1111/j.1432-1033.1994.tb18587.x.

Iwata, S. *et al.* (1995) 'Structure at 2.8 Å resolution of cytochrome c oxidase from paracoccus denitrificans', *Nature*. Nature, 376(6542), pp. 660–669. doi: 10.1038/376660a0.

Iwata, S. *et al.* (1998) 'Complete structure of the 11-subunit bovine mitochondrial cytochrome bc1 complex', *Science*. American Association for the Advancement of Science, 281(5373), pp. 64–71. doi: 10.1126/science.281.5373.64.

Jackson, J. B. (2012) 'A review of the binding-change mechanism for proton-translocating transhydrogenase', *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1817(10), pp. 1839–1846. doi: 10.1016/j.bbabio.2012.04.006.

Jackson, J. B. *et al.* (2015) 'Review and Hypothesis. New insights into the reaction mechanism of transhydrogenase: Swivelling the dIII component may gate the proton channel', *FEBS Letters*, pp. 2027–2033. doi: 10.1016/j.febslet.2015.06.027.

Jan, J. *et al.* (2013) 'Metabolic engineering and transhydrogenase effects on NADPH availability in *escherichia coli*', *Biotechnology Progress*, 29(5), pp. 1124–1130. doi: 10.1002/btpr.1765.

Jeeves, M. *et al.* (2000) 'Solution structure of the NADP(H)-binding component (dIII) of proton-translocating transhydrogenase from Rhodospirillum rubrum', *Biochimica et Biophysica Acta - Bioenergetics*, 1459(2–3), pp. 248–257. doi: 10.1016/S0005-2728(00)00159-6.

Jin, P. *et al.* (2017) 'Electron cryo-microscopy structure of the mechanotransduction channel NOMPC', *Nature Publishing Group*. Nature Research, 547(7661), pp. 118–122. doi: 10.1038/nature22981.

Jo, S. *et al.* (2008) 'CHARMM-GUI: A web-based graphical user interface for CHARMM', *Journal of Computational Chemistry*. John Wiley & Sons, Ltd, 29(11), pp. 1859–1865. doi: 10.1002/jcc.20945.

Johansson, T. *et al.* (2005) 'X-ray structure of domain I of the proton-pumping membrane protein transhydrogenase from Escherichia coli', *Journal of Molecular Biology*, 352(2), pp. 299–312. doi: 10.1016/j.jmb.2005.07.022.

Kaila, V. R. I. (2018) 'Long-range proton-coupled electron transfer in biological energy conversion: Towards mechanistic understanding of respiratory complex i', *Journal of the Royal Society Interface*. The Royal Society, p. 20170916. doi: 10.1098/rsif.2017.0916.

Kaledhonkar, S. *et al.* (2019) 'Late steps in bacterial translation initiation visualized using time-resolved cryo-EM', *Nature*. Nature Publishing Group, pp. 400–404. doi: 10.1038/s41586-019-1249-5.

Kampjut, D. and Sazanov, L. A. (2019) 'Structure and mechanism of mitochondrial proton-translocating transhydrogenase', *Nature*. Nature Publishing Group, 573(7773), pp. 291–295. doi: 10.1038/s41586-019-1519-2.

Kaplan, N. O., Colowick, S. P. and Neufeld, E. F. (1952) 'Pyridine Nucleotide Transhydrogenase. 2. Direct evidence for and mechanism of the transhydrogenase reaction', *Journal of Biological Chemistry*, 195, pp. 107–119. Available at: http://www.jbc.org/content/195/1/107.full.pdf (Accessed: 23 April 2017).

Kaplan, N. O., Colowick, S. P. and Neufeld, E. F. (1953) 'Pyridine nucleotide transhydrogenase: III. Animal tissue transhydrogenases', *Journal of Biological Chemistry*, pp. 1–15. Available at: http://www.jbc.org/content/205/1/1.full.pdf (Accessed: 23 April 2017).

Kelley, L. A. *et al.* (2015) 'The Phyre2 web portal for protein modeling, prediction and analysis.', *Nature protocols*. Europe PMC Funders, 10(6), pp. 845–58. doi: 10.1038/nprot.2015.053.

Kim, A. *et al.* (2010) 'Genetic modifier of mitochondrial superoxide dismutase-deficient mice delays heart failure and prolongs survival', *Mammalian Genome*. Springer-Verlag, 21(11–12), pp. 534–542. doi: 10.1007/s00335-010-9299-x.

Kirby, D. M. *et al.* (2004) 'Mutations of the mitochondrial ND1 gene as a cause of MELAS', *Journal of Medical Genetics*, 41(10), pp. 784–789. doi: 10.1136/jmg.2004.020537.

Kotlyar, A. B. and Vinogradov, A. D. (1990) 'Slow active/inactive transition of the mitochondrial NADH-ubiquinone reductase', *BBA - Bioenergetics*. Elsevier, 1019(2), pp. 151–158. doi: 10.1016/0005-2728(90)90137-S.

Krissinel, E. and Henrick, K. (2007) 'Inference of macromolecular assemblies from crystalline state.', *Journal of Molecular Biology*, 372(3), pp. 774–97. doi: 10.1016/j.jmb.2007.05.022.

Kucukelbir, A., Sigworth, F. J. and Tagare, H. D. (2014) 'Quantifying the local resolution of cryo-EM density maps', *Nature Methods*. Nature Publishing Group, 11(1), pp. 63–65. doi:

10.1038/nmeth.2727.

Kühlbrandt, W. (2014) ‘The resolution revolution’, *Science*. American Association for the Advancement of Science, pp. 1443–1444. doi: 10.1126/science.1251652.

Landau, M. *et al.* (2005) ‘ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures’, *Nucleic Acids Research*, 33(Web Server), pp. W299–W302. doi: 10.1093/nar/gki370.

Larkin, M. A. *et al.* (2007) ‘Clustal W and Clustal X version 2.0’, *Bioinformatics*. Oxford University Press, 23(21), pp. 2947–2948. doi: 10.1093/bioinformatics/btm404.

Laughlin, T. G. *et al.* (2019) ‘Structure of the complex I-like molecule NDH of oxygenic photosynthesis’, *Nature*. Nature Publishing Group, pp. 411–414. doi: 10.1038/s41586-019-0921-0.

Lepault, J., Booy, F. P. and Dubochet, J. (1983) ‘Electron microscopy of frozen biological suspensions’, *Journal of Microscopy*. J Microsc, 129(1), pp. 89–102. doi: 10.1111/j.1365-2818.1983.tb04163.x.

Letts, J. A. *et al.* (2016) ‘Purification of ovine respiratory complex i results in a highly active and stable preparation’, *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology, 291(47), pp. 24657–24675. doi: 10.1074/jbc.M116.735142.

Letts, J. A. *et al.* (2019) ‘Structures of Respiratory Supercomplex I+III2 Reveal Functional and Conformational Crosstalk’, *Molecular Cell*. Cell Press, 75(6), pp. 1131-1146.e6. doi: 10.1016/j.molcel.2019.07.022.

Letts, J. A., Fiedorczuk, K. and Sazanov, L. A. (2016) ‘The architecture of respiratory supercomplexes’, *Nature*. Nature Publishing Group, 537(7622), pp. 644–648. doi: 10.1038/nature19774.

Leung, J. H. *et al.* (2015) ‘Structural biology. Division of labor in transhydrogenase by alternating proton translocation and hydride transfer.’, *Science (New York, N.Y.)*, 347(6218), pp. 178–81. doi: 10.1126/science.1260451.

Li, Q. *et al.* (2015) ‘Resting state of the human proton channel dimer in a lipid bilayer’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 112(44), pp. E5926–E5935. doi: 10.1073/pnas.1515043112.

Li, S. *et al.* (2018) 'Nicotinamide nucleotide transhydrogenase-mediated redox homeostasis promotes tumor growth and metastasis in gastric cancer', *Redox Biology*. Elsevier, 18, pp. 246–255. doi: 10.1016/J.REDOX.2018.07.017.

Lobo-Jarne, T. and Ugalde, C. (2018) 'Respiratory chain supercomplexes: Structures, function and biogenesis', *Seminars in Cell and Developmental Biology*. Elsevier Ltd, pp. 179–190. doi: 10.1016/j.semcdb.2017.07.021.

Maklashina, E., Kotlyar, A. B. and Cecchini, G. (2003) 'Active/de-active transition of respiratory complex I in bacteria, fungi, and animals', *Biochimica et Biophysica Acta - Bioenergetics*. Elsevier, 1606(1–3), pp. 95–103. doi: 10.1016/S0005-2728(03)00087-2.

Marreiros, B. C. *et al.* (2016) 'Exploring membrane respiratory chains', *Biochimica et Biophysica Acta - Bioenergetics*. Elsevier B.V., 1857(8), pp. 1039–1067. doi: 10.1016/j.bbabio.2016.03.028.

Martin, W., Sousa, F. and Lane, N. (2014) 'Energy at life's origin', *Science*, 344(6188), pp. 1092–1093. doi: 10.1111/soru.12033.11.

Mastronarde, D. N. (2005) 'Automated electron microscope tomography using robust prediction of specimen movements', *Journal of Structural Biology*, 152(1), pp. 36–51. doi: 10.1016/j.jsb.2005.07.007.

Mather, O. C. *et al.* (2004) 'Active-site conformational changes associated with hydride transfer in proton-translocating transhydrogenase', *Biochemistry*. American Chemical Society, 43(34), pp. 10952–10964. doi: 10.1021/bi0497594.

Mathiesen, C. and Hägerhäll, C. (2002) 'Transmembrane topology of the NuoL, M and N subunits of NADH:quinone oxidoreductase and their homologues among membrane-bound hydrogenases and bona fide antiporters', *Biochimica et Biophysica Acta - Bioenergetics*. Biochim Biophys Acta, 1556(2–3), pp. 121–132. doi: 10.1016/S0005-2728(02)00343-2.

Mathiesen, C. and Hägerhäll, C. (2003) 'The "antiporter module" of respiratory chain Complex I includes the MrpC/NuoK subunit - A revision of the modular evolution scheme', *FEBS Letters*. Elsevier, 549(1–3), pp. 7–13. doi: 10.1016/S0014-5793(03)00767-1.

Matsuzaki, S. and Humphries, K. M. (2015) 'Selective inhibition of deactivated mitochondrial complex i by biguanides', *Biochemistry*. American Chemical Society, 54(11), pp. 2011–2021. doi: 10.1021/bi501473h.

McMullan, G. *et al.* (2014) ‘Comparison of optimal performance at 300keV of three direct electron detectors for use in low dose electron microscopy’, *Ultramicroscopy*. Elsevier, 147, pp. 156–163. doi: 10.1016/j.ultramic.2014.08.002.

Meimaridou, E. *et al.* (2012) ‘Mutations in NNT encoding nicotinamide nucleotide transhydrogenase cause familial glucocorticoid deficiency.’, *Nature genetics*. Nature Publishing Group, 44(7), pp. 740–2. doi: 10.1038/ng.2299.

Mercer-Haines, N. and Fioravanti, C. F. (2008) ‘Hymenolepis diminuta: Mitochondrial transhydrogenase as an additional site for anaerobic phosphorylation’, *Experimental Parasitology*, 119(1), pp. 24–29. doi: 10.1016/j.exppara.2007.12.006.

Mitchell, P. (1961) ‘Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism’, *Nature*, 191, pp. 144–149.

Moparthi, V. K. and Hägerhäll, C. (2011) ‘The evolution of respiratory chain complex i from a smaller last common ancestor consisting of 11 protein subunits’, *Journal of Molecular Evolution*. Springer, 72(5–6), pp. 484–497. doi: 10.1007/s00239-011-9447-2.

Murai, M. and Miyoshi, H. (2016) ‘Current topics on inhibitors of respiratory complex i’, *Biochimica et Biophysica Acta - Bioenergetics*. Elsevier B.V., 1857(7), pp. 884–891. doi: 10.1016/j.bbabio.2015.11.009.

Nakamura, H. (1996) ‘Roles of electrostatic interaction in proteins’, *Quarterly Reviews of Biophysics*. Cambridge University Press (CUP), 29(1), pp. 1–90. doi: 10.1017/s0033583500005746.

Nickel, A. G. G. *et al.* (2015) ‘Reversal of mitochondrial transhydrogenase causes oxidative stress in heart failure’, *Cell Metabolism*, 22(3), pp. 472–484. doi: 10.1016/j.cmet.2015.07.008.

Nogales, E. (2015) ‘The development of cryo-EM into a mainstream structural biology technique’, *Nature Methods*. Nature Publishing Group, 13(1), pp. 24–27. doi: 10.1038/nmeth.3694.

Obiozo, U. M. *et al.* (2007) ‘Substitution of tyrosine 146 in the dI component of proton-translocating transhydrogenase leads to reversible dissociation of the active dimer into inactive monomers’, *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology, 282(50), pp. 36434–36443. doi: 10.1074/jbc.M705433200.

Olausson, T. *et al.* (1993) 'Site-directed mutagenesis of tyrosine residues at nicotinamide nucleotide binding sites of Escherichia coli transhydrogenase', *Biochemistry*. American Chemical Society, 32(48), pp. 13237–13244. doi: 10.1021/bi00211a036.

Olsson, M. H. M. *et al.* (2011) 'PROPKA3: Consistent treatment of internal and surface residues in empirical p K a predictions', *Journal of Chemical Theory and Computation*. American Chemical Society, 7(2), pp. 525–537. doi: 10.1021/ct100578z.

Orlova, E. V. and Saibil, H. R. (2011) 'Structural analysis of macromolecular assemblies by electron microscopy', *Chemical Reviews*, 111(12), pp. 7710–7748. doi: 10.1021/cr100353t.

Padayatti, P. S. *et al.* (2017) 'Critical Role of Water Molecules in Proton Translocation by the Membrane-Bound Transhydrogenase', *Structure*. Elsevier Ltd., pp. 1–9. Available at: http://dx.doi.org/10.1016/j.str.2017.05.022.

Parey, A. K. *et al.* (2019) 'High-resolution cryo-EM structures of respiratory complex I - mechanism , assembly and disease', *Science Advances*. American Association for the Advancement of Science, 5(12), pp. 1–47. doi: 10.1126/sciadv.aax9484.

Parey, K. *et al.* (2018) 'Cryo-EM structure of respiratory complex I at work', *eLife*, 7. doi: 10.7554/eLife.39213.

Penzkofer, A. *et al.* (2007) 'Protein aggregation studied by forward light scattering and light transmission analysis', *Chemical Physics*. North-Holland, 342(1–3), pp. 55–63. doi: 10.1016/j.chemphys.2007.09.014.

Pettersen, E. F. *et al.* (2004) 'UCSF Chimera - A visualization system for exploratory research and analysis', *Journal of Computational Chemistry*, 25(13), pp. 1605–1612. doi: 10.1002/jcc.20084.

Phelps, D. C. and Hatefi, Y. (1984) 'Interaction of purified nicotinamidenucleotide transhydrogenase with dicyclohexylcarbodiimide', *Biochemistry*, 23(19), pp. 4475–4480. doi: 10.1021/BI00314A037.

Prasad, G. S. *et al.* (1999) 'Crystal structure of transhydrogenase domain III at 1.2 A resolution.', *Nature structural biology*. Nature Publishing Group, 6(12), pp. 1126–31. doi: 10.1038/70067.

Prasad, G. S. *et al.* (2002) 'Crystal structures of transhydrogenase domain I with and without bound NADH', *Biochemistry*. American Chemical Society, 41(42), pp. 12745–12754. doi:

10.1021/bi020251f.

Prasad, R *et al.* (2014) 'Oxidative stress and adrenocortical insufficiency', *Journal of Endocrinology*, 221, pp. 63–73. doi: 10.1530/JOE-13-0346.

Prasad, Rathi *et al.* (2014) 'Thioredoxin reductase 2 (TXNRD2) mutation associated with familial glucocorticoid deficiency (FGD)', *Journal of Clinical Endocrinology and Metabolism*, 99(8), pp. E1556–E1563. doi: 10.1210/jc.2013-3844.

Pravda, L. *et al.* (2018) 'MOLEonline: a web-based tool for analyzing channels, tunnels and pores (2018 update)', *Nucleic Acids Research*. Oxford University Press, 46(W1), pp. W368–W373. doi: 10.1093/nar/gky309.

Prisant, M. G. *et al.* (2020) 'New tools in MolProbity validation: CaBLAM for CryoEM backbone, UnDowser to rethink "waters," and NGL Viewer to recapture online 3D graphics', *Protein Science*. Blackwell Publishing Ltd, 29(1), pp. 315–329. doi: 10.1002/pro.3786.

Punjani, A. *et al.* (2017) 'CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination', *Nature Methods*. Nature Publishing Group, 14(3), pp. 290–296. doi: 10.1038/nmeth.4169.

Ripoll, V. M. *et al.* (2012) 'Nicotinamide nucleotide transhydrogenase (NNT) acts as a novel modulator of macrophage inflammatory responses', *The FASEB Journal*, 26(8), pp. 3550–3562. doi: 10.1096/fj.11-199935.

Roberts, P. G. and Hirst, J. (2012) 'The deactive form of respiratory complex I from mammalian mitochondria is a Na+/H+ antiporter.', *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 287(41), pp. 34743–51. doi: 10.1074/jbc.M112.384560.

Rodenburg, R. J. (2016) 'Mitochondrial complex I-linked disease', *Biochimica et Biophysica Acta - Bioenergetics*. Elsevier B.V., 1857(7), pp. 938–945. doi: 10.1016/j.bbabio.2016.02.012.

Rohou, A. and Grigorieff, N. (2015) 'CTFFIND4: Fast and accurate defocus estimation from electron micrographs', *Journal of Structural Biology*. Academic Press, 192(2), pp. 216–221. doi: 10.1016/J.JSB.2015.08.008.

Ronchi, J. A. *et al.* (2013) 'A spontaneous mutation in the nicotinamide nucleotide transhydrogenase gene of C57BL/6J mice results in mitochondrial redox abnormalities', *Free*

Radical Biology and Medicine, 63, pp. 446–456. doi: 10.1016/j.freeradbiomed.2013.05.049.

Ronchi, J. A. *et al.* (2016) 'The Contribution of Nicotinamide Nucleotide Transhydrogenase to Peroxide Detoxification Is Dependent on the Respiratory State and Counterbalanced by Other Sources of NADPH in Liver Mitochondria', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology, 291(38), pp. 20173–87. doi: 10.1074/jbc.M116.730473.

Rosiersl, C. Des *et al.* (1994) 'Reversibility of the Mitochondrial Isocitrate Dehydrogenase Reaction in the Perfused Rat Liver EVIDENCE FROM ISOTOPOMER ANALYSIS OF CITRIC ACID CYCLE INTERMEDIATES*', 269(44), pp. 27179–27182. Available at: http://www.jbc.org/content/269/44/27179.full.pdf (Accessed: 14 August 2017).

Roucher-Boulez, F. *et al.* (2016) 'NNT mutations: A cause of primary adrenal insufficiency, oxidative stress and extraadrenal defects', *European Journal of Endocrinology*, 175(1), pp. 73–84. doi: 10.1530/EJE-16-0056.

Runswick, M. J. *et al.* (2013) 'The affinity purification and characterization of ATP synthase complexes from mitochondria.', *Open biology*, 3, p. 120160. doi: 10.1098/rsob.120160.

Russo, C. J. and Passmore, L. A. (2016) *Progress towards an optimal specimen support for electron cryomicroscopy*, *Current Opinion in Structural Biology*. doi: 10.1016/j.sbi.2015.12.007.

Rydstrom, J. (1972) 'Site-Specific Inhibitors of Mitochondrial Nicotinamide-Nucleotide Transhydrogenase', *European Journal of Biochemistry*. Blackwell Publishing Ltd, 31(3), pp. 496–504. doi: 10.1111/j.1432-1033.1972.tb02557.x.

Rydström, J. *et al.* (1976) 'Evidence for a lipid dependence of mitochondrial nicotinamide nucleotide transhydrogenase', *Biochimica et Biophysica Acta (BBA) - Bioenergetics*. Elsevier, 430(3), pp. 419–425. doi: 10.1016/j.aqpro.2013.07.003.

Rydström, J. (2006) 'Mitochondrial NADPH, transhydrogenase and disease', *Biochimica et Biophysica Acta - Bioenergetics*, pp. 721–726. doi: 10.1016/j.bbabio.2006.03.010.

Saeed, S. *et al.* (2020) '<scp>NAD</scp> (P) transhydrogenase has vital non-mitochondrial functions in malaria parasite transmission', *EMBO reports*. John Wiley & Sons, Ltd, p. e47832. doi: 10.15252/embr.201947832.

Santos, L. R. B. *et al.* (2017) 'NNT reverse mode of operation mediates glucose control of

mitochondrial NADPH and glutathione redox state in mouse pancreatic β-cells', *Molecular Metabolism*, 6(6), pp. 535–547. doi: 10.1016/j.molmet.2017.04.004.

Sauer, U. *et al.* (2004) 'The soluble and membrane-bound transhydrogenases UdhA and PntAB have divergent functions in NADPH metabolism of Escherichia coli.', *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 279(8), pp. 6613–9. doi: 10.1074/jbc.M311657200.

Saura, P. *et al.* (2019) 'Electric field modulated redox-driven protonation and hydration energetics in energy converting enzymes', *Chemical Communications*. Royal Society of Chemistry, 55(43), pp. 6078–6081. doi: 10.1039/c9cc01135h.

Sazanov, L. A. (2015) 'A giant molecular proton pump: Structure and mechanism of respiratory complex I', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, pp. 375–388. doi: 10.1038/nrm3997.

Sazanov, L. A. A. and Jackson, J. B. B. (1994) 'Proton-translocating transhydrogenase and NAD- and NADP-linked isocitrate dehydrogenases operate in a substrate cycle which contributes to fine regulation of the tricarboxylic acid cycle activity in mitochondria', *FEBS Letters*, 344(2–3), pp. 109–116. doi: 10.1016/0014-5793(94)00370-X.

Sazanov, L. A. and Hinchliffe, P. (2006) 'Structure of the hydrophilic domain of respiratory complex I from Thermus thermophilus', *Science*. Science, 311(5766), pp. 1430–1436. doi: 10.1126/science.1123809.

Sazanov, L. A. and Jackson, J. B. (1993) 'Activation and inhibition of mitochondrial transhydrogenase by metal ions', *Biochimica et Biophysica Acta*. Elsevier, pp. 225–228. doi: 10.1016/0005-2728(93)90177-H.

Sazanov, L. A. and Jackson, J. B. (1995) 'Cyclic reactions catalysed by detergent-dispersed and reconstituted transhydrogenase from beef-heart mitochondria; implications for the mechanism of proton translocation', *Biochimica et Biophysica Acta*, 1231, pp. 304–312. Available at: http://ac.els-cdn.com/0005272895000962/1-s2.0-0005272895000962-main.pdf?_tid=79bb3be2-2d19-11e7-9123-00000aacb362&acdnat=1493497489_f0ad49058d88d54f85504a6a165b0f8d (Accessed: 29 April 2017).

Schagger, H. and Pfeiffer, K. (2000) 'Supercomplexes in the respiratory chains of yeast and

mammalian mitochondria', *The EMBO Journal*, 19(8), pp. 1777–1783. doi: 10.1093/emboj/19.8.1777.

Scheres, S. H. W. *et al.* (2007) 'Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization', *Nature Methods*. Nature Publishing Group, 4(1), pp. 27–29. doi: 10.1038/nmeth992.

Scheres, S. H. W. (2012a) 'A bayesian view on cryo-EM structure determination', *Journal of Molecular Biology*. Elsevier Ltd, 415(2), pp. 406–418. doi: 10.1016/j.jmb.2011.11.010.

Scheres, S. H. W. (2012b) 'RELION: Implementation of a Bayesian approach to cryo-EM structure determination', *Journal of Structural Biology*. Elsevier Inc., 180(3), pp. 519–530. doi: 10.1016/j.jsb.2012.09.006.

Schilling, B. *et al.* (2005) 'Mass spectrometric identification of a novel phosphorylation site in subunit NDUFA10 of bovine mitochondrial complex I', *FEBS Letters*. Elsevier, 579(11), pp. 2485–2490. doi: 10.1016/j.febslet.2005.03.061.

Schuller, J. M. *et al.* (2019) 'Structural adaptations of photosynthetic complex I enable ferredoxin-dependent electron transfer', *Science*. American Association for the Advancement of Science, 363(6424), pp. 257–260. doi: 10.1126/science.aau3613.

Schulte, M. *et al.* (2019) 'A mechanism to prevent production of reactive oxygen species by Escherichia coli respiratory complex I', *Nature Communications*. Nature Publishing Group, 10(1), p. 2551. doi: 10.1038/s41467-019-10429-0.

Scott, R., Van Vliet, G. and Deladoëy, J. (2017) 'Association of adrenal insufficiency with insulin-dependent diabetes mellitus in a patient with inactivating mutations in nicotinamide nucleotide transhydrogenase: a phenocopy of the animal model.', *European journal of endocrinology*. European Society of Endocrinology, 176(3), pp. C1–C2. doi: 10.1530/EJE-16-0970.

Sharma, V. *et al.* (2015) 'Redox-induced activation of the proton pump in the respiratory complex i', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 112(37), pp. 11571–11576. doi: 10.1073/pnas.1503761112.

Sharpley, M. S. and Hirst, J. (2006) 'The inhibition of mitochondrial complex I (NADH:ubiquinone oxidoreductase) by Zn2+.', *The Journal of biological chemistry*.

American Society for Biochemistry and Molecular Biology, 281(46), pp. 34803–9. doi: 10.1074/jbc.M607389200.

Shi, A. *et al.* (2013) 'Activating transhydrogenase and NAD kinase in combination for improving isobutanol production', *Metabolic Engineering*, 16, pp. 1–10. doi: 10.1016/j.ymben.2012.11.008.

Singh, A. *et al.* (2003) 'Interactions between transhydrogenase and thio-nicotinamide analogues of NAD(H) and NADP(H) underline the importance of nucleotide conformational changes in coupling to proton translocation', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology, 278(35), pp. 33208–33216. doi: 10.1074/jbc.M303061200.

Smith, A. L. (1967) '[13] Preparation, properties, and conditions for assay of mitochondria: Slaughterhouse material, small-scale', *Methods in Enzymology*. Academic Press, 10, pp. 81–86. doi: 10.1016/0076-6879(67)10016-5.

Smith, P. M., Fox, J. L. and Winge, D. R. (2012) 'Biogenesis of the cytochrome bc 1 complex and role of assembly factors', *Biochimica et Biophysica Acta - Bioenergetics*. NIH Public Access, pp. 276–286. doi: 10.1016/j.bbabio.2011.11.009.

Stroud, D. A. *et al.* (2016) 'Accessory subunits are integral for assembly and function of human mitochondrial complex I', *Nature*. Nature Publishing Group, pp. 1–17. doi: 10.1038/nature19754.

Studley, W. K. *et al.* (1999) 'Phylogenetic analyses of proton-translocating transhydrogenases.', *Microbial & comparative genomics*, 4(3), pp. 173–186. doi: 10.1089/omi.1.1999.4.173.

Sundaresan, V. *et al.* (2003) 'Conformational Change in the NADP(H) Binding Domain of Transhydrogenase Defines Four States', *Biochemistry*. American Chemical Society, 42(42), pp. 12143–12153. doi: 10.1021/bi035006q.

Sundaresan, V. *et al.* (2005) 'Conformational Diversity in NAD(H) and Interacting Transhydrogenase Nicotinamide Nucleotide Binding Domains', *Journal of Molecular Biology*, 346(2), pp. 617–629. doi: 10.1016/j.jmb.2004.11.070.

Timischl, B. *et al.* (2008) 'Development of a quantitative, validated Capillary electrophoresis-time of flight-mass spectrometry method with integrated high-confidence

analyte identification for metabolomics', *Electrophoresis*, 29(10), pp. 2203–2214. doi: 10.1002/elps.200700517.

Tong, R. C. W., Glavas, N. A. and Bragg, P. D. (1991) 'Topological analysis of the pyridine nucleotide transhydrogenase of Escherichia coli using proteolytic enzymes', *Biochimica et Biophysica Acta*, 1080, pp. 19–28. Available at: http://ac.els-cdn.com/016748389190106A/1-s2.0-016748389190106A-main.pdf?_tid=aa377536-5200-11e7-a9a6-00000aacb360&acdnat=1497555027_c84cd5b707e6ec40eae2b289ac927cae (Accessed: 15 June 2017).

Toye, A. A. *et al.* (2005) 'A genetic and physiological study of impaired glucose homeostasis control in C57BL/6J mice', *Diabetologia*. Springer-Verlag, 48(4), pp. 675–686. doi: 10.1007/s00125-005-1680-z.

Tynan-Connolly, B. M. and Nielsen, J. E. (2006) 'Redesigning protein pKa values', *Protein Science*. Wiley, 16(2), pp. 239–249. doi: 10.1110/ps.062538707.

Ugur, I., Rutherford, A. W. and Kaila, V. R. I. (2016) 'Redox-coupled substrate water reorganization in the active site of Photosystem II - The role of calcium in substrate water delivery', *Biochimica et Biophysica Acta - Bioenergetics*. Elsevier B.V., 1857(6), pp. 740–748. doi: 10.1016/j.bbabio.2016.01.015.

Uhlén, M. *et al.* (2015) 'Tissue-based map of the human proteome', *Science*, 347(6220). Available at: http://science.sciencemag.org/content/347/6220/1260419/tab-pdf (Accessed: 8 August 2017).

Uno, S. *et al.* (2019) 'Exploring the quinone/inhibitor-binding pocket in mitochondrial respiratory complex I by chemical biology approaches', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology, 294(2), pp. 679–696. doi: 10.1074/jbc.RA118.006056.

Uno, S. *et al.* (2020) 'Oversized ubiquinones as molecular probes for structural dynamics of the ubiquinone reaction site in mitochondrial respiratory complex I', *Journal of Biological Chemistry*, p. jbc.RA119.012347. doi: 10.1074/jbc.RA119.012347.

Urbani, A. and Warne, T. (2005) 'A colorimetric determination for glycosidic and bile salt-based detergents: Applications in membrane protein research', *Analytical Biochemistry*. Anal Biochem, 336(1), pp. 117–124. doi: 10.1016/j.ab.2004.09.040.

Uribe, E. G. and Jagendorf, A. T. (1968) ‘Membrane permeability and internal volume as factors in ATP synthesis by spinach chloroplasts’, *Archives of Biochemistry and Biophysics*. Academic Press, 128(2), pp. 351–359. doi: 10.1016/0003-9861(68)90041-6.

Vandock, K. P. *et al.* (2011) ‘Phospholipid Dependence of the Reversible, Energy-Linked, Mitochondrial Transhydrogenase in Manduca sexta’, *The Journal of Membrane Biology*. Springer-Verlag, 242(2), pp. 89–94. doi: 10.1007/s00232-011-9379-1.

Vartak, R. *et al.* (2015) ‘Redefining the roles of mitochondrial DNA-encoded subunits in respiratory Complex I assembly’, *Biochimica et Biophysica Acta - Molecular Basis of Disease*. Elsevier, 1852(7), pp. 1531–1539. doi: 10.1016/j.bbadis.2015.04.008.

Venning, J. D. *et al.* (1998) ‘Stopped-flow kinetics of hydride transfer between nucleotides by recombinant domains of proton-translocating transhydrogenase’, *European Journal of Biochemistry*. Springer-Verlag, 257(1), pp. 202–209. doi: 10.1046/j.1432-1327.1998.2570202.x.

Venning, J. D. *et al.* (2000) ‘Stopped-flow reaction kinetics of recombinant components of proton-translocating transhydrogenase with physiological nucleotides.’, *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 275(26), pp. 19490–7. doi: 10.1074/jbc.M000577200.

Verkhovskaya, M. and Bloch, D. A. (2013) ‘Energy-converting respiratory Complex I: On the way to the molecular mechanism of the proton pump’, *The International Journal of Biochemistry & Cell Biology*. Pergamon, 45(2), pp. 491–511. doi: 10.1016/j.biocel.2012.08.024.

Vinogradov, A. D. and Grivennikova, V. G. (2001) ‘The Mitochondrial Complex I: Progress in Understanding of Catalytic Properties’, *IUBMB Life (International Union of Biochemistry and Molecular Biology: Life)*. Taylor and Francis Inc., 52(3–5), pp. 129–134. doi: 10.1080/15216540152845920.

Vinothkumar, K. R., Zhu, J. and Hirst, J. (2014) ‘Architecture of mammalian respiratory complex I’, *Nature*. Nature Publishing Group, 515(7525), pp. 80–84. doi: 10.1038/nature13686.Architecture.

Walker, J. E. (2012) ‘The ATP synthase: The understood, the uncertain and the unknown’, *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1817, p. S1. doi:

10.1016/j.bbabio.2012.06.013.

Wang, J. (2017) ‘On the appearance of carboxylates in electrostatic potential maps’, *Protein Science*. Blackwell Publishing Ltd, 26(3), pp. 396–402. doi: 10.1002/pro.3093.

Warnau, J. *et al.* (2018) ‘Redox-coupled quinone dynamics in the respiratory complex I.’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 115(36), pp. E8413–E8420. doi: 10.1073/pnas.1805468115.

White, S. A. *et al.* (2000) ‘The high-resolution structure of the NADP(H)-binding component (dIII) of proton-translocating transhydrogenase from human heart mitochondria’, *Structure*, 8(1), pp. 1–12. doi: 10.1016/S0969-2126(00)00075-7.

Whitehead, S. J. *et al.* (2005) ‘Zinc ions selectively inhibit steps associated with binding and release of NADP(H) during turnover of proton-translocating transhydrogenase’, *FEBS Letters*, 579(13), pp. 2863–2867. doi: 10.1016/j.febslet.2005.04.026.

Wikström, M., Sharma, V., Kaila, Ville R I, *et al.* (2015) ‘New perspectives on proton pumping in cellular respiration’, *Chemical Reviews*, 115(5), pp. 2196–2221. doi: 10.1021/cr500448t.

Wikström, M., Sharma, V., Kaila, Ville R.I., *et al.* (2015) ‘New perspectives on proton pumping in cellular respiration’, *Chemical Reviews*. American Chemical Society, pp. 2196–2221. doi: 10.1021/cr500448t.

Wikström, M. and Hummer, G. (2012) ‘Stoichiometry of proton translocation by respiratory complex I and its mechanistic implications’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 109(12), pp. 4431–4436. doi: 10.1073/pnas.1120949109.

Wikström, M. and Sharma, V. (2018) ‘Proton pumping by cytochrome c oxidase – A 40 year anniversary’, *Biochimica et Biophysica Acta - Bioenergetics*. Elsevier B.V., pp. 692–698. doi: 10.1016/j.bbabio.2018.03.009.

Wolf, S. *et al.* (2016) ‘Brain damage resulting from postnatal hypoxic-ischemic brain injury is reduced in C57BL/6J mice as compared to C57BL/6N mice’, *Brain Research*, 1650, pp. 224–231. doi: 10.1016/j.brainres.2016.09.013.

Wray, C. D. *et al.* (2013) ‘A new mutation in MT-ND1 m.3928G>C p.V208L causes Leigh disease with infantile spasms’, *Mitochondrion*, 13(6), pp. 656–661. doi:

10.1016/j.mito.2013.09.004.

Wu, D. *et al.* (2019) 'Long noncoding RNA NNT-AS1 enhances the malignant phenotype of bladder cancer by acting as a competing endogenous RNA on microRNA-496 thereby increasing HMGB1 expression', *Aging*. Impact Journals LLC, 11(24), pp. 12624–12640. doi: 10.18632/aging.102591.

Wu, L. N. Y., Alberta, J. A. and Fisher, R. R. (1986) '[33] Purification and reconstitution of bovine heart mitochondrial transhydrogenase', *Methods in Enzymology*, 126, pp. 353–360. doi: 10.1016/S0076-6879(86)26035-8.

Wu, M. *et al.* (2016) 'Structure of Mammalian Respiratory Supercomplex I1III2IV1', *Cell*. Cell Press, 167(6), pp. 1598-1609.e10. doi: 10.1016/j.cell.2016.11.012.

Yachdav, G. *et al.* (2014) 'PredictProtein - An open resource for online prediction of protein structural and functional features', *Nucleic Acids Research*, 42(W1). doi: 10.1093/nar/gku366.

Yamaguchi, M. and Hatefi, Y. (1989) 'Mitochondrial nicotinamide nucleotide transhydrogenase: NADPH binding increases and NADP binding decreases the acidity and susceptibility to modification of cysteine-893.', *Biochemistry*, 28(14), pp. 6050–6. doi: 10.1021/BI00440A049.

Yamaguchi, M. and Hatefi, Y. (1993) 'Energy-transducing nicotinamide nucleotide transhydrogenase. Nucleotide binding properties of the purified enzyme and proteolytic fragments', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology, 268(24), pp. 17871–17877. Available at: http://www.jbc.org/content/268/24/17871.full.pdf (Accessed: 4 January 2017).

Yamaguchi, M. and Hatefi, Y. (1995) 'Proton-translocating nicotinamide nucleotide transhydrogenase. Reconstitution of the extramembranous nucleotide-binding domains', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology, 270(47), pp. 28165–28168. doi: 10.1074/jbc.270.47.28165.

Yamaguchi, M. and Hatefi, Y. (1997) 'High cyclic transhydrogenase activity catalyzed by expressed and reconstituted nucleotide-binding domains of Rhodospirillum rubrum transhydrogenase', *Biochimica et Biophysica Acta - Bioenergetics*, 1318(1–2), pp. 225–234. doi: 10.1016/S0005-2728(96)00139-9.

Yamaguchi, M. and Stout, C. D. (2003) 'Essential Glycine in the Proton Channel of Escherichia coli Transhydrogenase', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology, 278(46), pp. 45333–45339. doi: 10.1074/jbc.M308236200.

Yamaguchi, M., Stout, D. C. and Hatefi, Y. (2002) 'The proton channel of the energy-transducing nicotinamide nucleotide transhydrogenase of Escherichia coli', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology, 277(37), pp. 33670–33675. doi: 10.1074/jbc.M204170200.

Yamaguchi, M., Wakabayashi, S. and Hatefi, Y. (1990) 'Mitochondrial Energy-Linked Nicotinamide Nucleotide Transhydrogenase: Effect of Substrates on the Sensitivity of the Enzyme to Trypsin and Identification of Tryptic Cleavage Sites', *Biochemistry*, 29(17), pp. 4136–4143. doi: 10.1021/bi00469a017.

Yip, C. Y. *et al.* (2011) 'Evolution of respiratory complex I "Supernumerary" subunits are present in the α-proteobacterial enzyme', *Journal of Biological Chemistry*. J Biol Chem, 286(7), pp. 5023–5033. doi: 10.1074/jbc.M110.194993.

Yu, H. *et al.* (2018) 'Structure of an Ancient Respiratory System', *Cell*. Elsevier, 0(0). doi: 10.1016/j.cell.2018.03.071.

Zhang, C. *et al.* (2020) 'Structural insights into NDH-1 mediated cyclic electron transfer', *Nature Communications*. Nature Research, 11(1), pp. 1–13. doi: 10.1038/s41467-020-14732-z.

Zhang, L. and Hermans, J. (1996) 'Hydrophilicity of cavities in proteins', *Proteins: Structure, Function and Genetics*. Proteins, 24(4), pp. 433–438. doi: 10.1002/(SICI)1097-0134(199604)24:4<433::AID-PROT3>3.0.CO;2-F.

Zhang, X. C. and Li, B. (2019) 'Towards understanding the mechanisms of proton pumps in Complex-I of the respiratory chain', *Biophysics Reports*. Springer Science and Business Media LLC, 5(5–6), pp. 219–234. doi: 10.1007/s41048-019-00094-7.

Zheng, H. *et al.* (2008) 'Data mining of metal ion environments present in protein structures', *Journal of Inorganic Biochemistry*. J Inorg Biochem, 102(9), pp. 1765–1776. doi: 10.1016/j.jinorgbio.2008.05.006.

Zheng, S. Q. *et al.* (2017) 'MotionCor2: anisotropic correction of beam-induced motion for

improved cryo-electron microscopy', *Nature Methods*. Nature Publishing Group, 14(4), pp. 331–332. doi: 10.1038/nmeth.4193.

Zhu, Jiapeng, Vinothkumar, K. R. and Hirst, J. (2016) 'Structure of mammalian respiratory complex I', *Nature*, 536. doi: 10.1038/nature13686.

Zickermann, V. *et al.* (2015) 'Mechanistic insight from the crystal structure of mitochondrial complex I', *Science*. American Association for the Advancement of Science, 347(6217), pp. 44–49. doi: 10.1126/science.1259859.

Zivanov, J. *et al.* (2018) 'New tools for automated high-resolution cryo-EM structure determination in RELION-3', *eLife*, 7. doi: 10.7554/eLife.42166.

Appendix 1. Model and data collection statistics for NNT datasets

	NNT-NADP$^+$ "double face-down" (EMD-4635) (PDB 6QTI)	NNT-NADP$^+$ "single face-down" (EMD-4637) (PDB 6QUE)	Apo-NNT "double face-down" (EMDB-10099) (PDB 6S59)
Data collection and processing			
Magnification	130000x	130000x	130000x
Voltage (kV)	300	300	300
Electron exposure (e–/Å^2)	72	72	90
Defocus range (μm)	~ -1 to -2.5	~ -1 to -2.5	~ -1 to -2.5
Pixel size (Å)	1.065	1.065	0.84
Symmetry imposed	C1	C1	C1
Initial particle images (no.)	1076677	1076677	500001
Final particle images (no.)	98294	58441	67908
Map resolution (Å)	dI_2 – 3.2 Å, dII_2dIII_2 – 2.9 Å	dI_2 – 4.2 Å, dII_2dIII – 3.7 Å	dI_2 – 4.5 Å, dII_2dIII_2 – 3.7 Å
FSC threshold	0.143	0.143	0.143
Refinement			
Initial model used (PDB code)	dI and dII built de novo, dIII based on 1D4O	6QTI	6QTI
Model resolution (Å)	dI_2 – 3.4 Å, dII_2dIII_2 – 3.1 Å	dI_2 – 4.3 Å, dII_2dIII – 4.0 Å	dII_2dIII_2 – 3.8 Å
FSC threshold	0.5	0.5	0.5
Map sharpening *B* factor (Å^2)	dII_2dIII_2 -36.2 dI_2 -50.2	dII_2dIII -55 dI_2 -77.0	dII_2dIII_2 -67.3 dI_2 -197.5
Model composition		13791	
Non-hydrogen atoms	15996	1865	13397
Protein residues	2076	3	2066
Ligands	14		8
B factors (Å^2)			
Protein	74.0	128.0	115.8
Ligand	114.8	172.8	143.5
R.m.s. deviations			
Bond lengths (Å)	0.0071	0.008	0.0085
Bond angles (°)	1.29	1.30	1.36
Validation	1.56		
MolProbity score	3.58	1.66	1.59
EMRinger score	3.37	1.73	2.53
Clashscore	0	4.26	3.50
Poor rotamers (%)		0.21	0.44
Ramachandran plot			
Favored (%)	93.44	92.85	92.95
Allowed (%)	6.56	7.05	7.05
Disallowed (%)	0	0.1	0

Appendix 2 – Model and data collection statistics for native complex I dataset

	Native CxI – closed (EMD-11256) (PDB 6ZKO)	Native CxI – open1 (EMD-11257) (PDB 6ZKP)	Native CxI – open2 (EMD-11258) (PDB 6ZKQ)	Native CxI – open3 (EMD-11259) (PDB 6ZKR)	PA-focused refinement	MD-focused refinement
Data collection and processing						
Magnification	165000x	165000x	165000x	165000x	165000x	165000x
Voltage (kV)	300	300	300	300	300	300
Electron exposure (e–/Å^2)	98	98	98	98	98	98
Defocus range (μm)	~ -1 to -2	~ -1 to -2	~ -1 to -2	~ -1 to -2	~ -1 to -2	~ -1 to -2
Pixel size (Å)	1.061	1.061	1.061	1.061	1.061	1.061
Symmetry imposed	C1	C1	C1	C1	C1	C1
Initial particle images (no.)	128952	128952	128952	128952	128952	128952
Final particle images (no.)	10024	32982	22086	20316	22086	20316
Map resolution (Å) at 0.143 FSC threshold	3.8	3.2	3.3	3.5	3.3	3.5
Refinement						
Initial model used (PDB code)	5LNK	5LNK	5LNK	5LNK	5LNK	5LNK
Model resolution (Å) at 0.5 FSC threshold						
Map sharpening *B* factor (Å^2)	-72	-56	-61	-67		
Model composition						
Non-hydrogen atoms	67263	66769	66965	66796	-	-
Protein residues	66080	65436	65442	65462		
Ligands	1183	1333	1523	1334		
B factors (Å^2)						
Protein	47.4	54.4	50.7	53.5	-	-
Ligand	68.3	84.5	78.8	82.0		
R.m.s. deviations						
Bond lengths (Å)	0.0079	0.0081	0.0091	0.0085	-	-
Bond angles (°)	1.76	1.48	1.77	1.55		
Validation						
MolProbity score	1.84	1.63	1.67	1.82	-	-
EMRinger score	2.19	4.08	3.92	3.15		
Clashscore	7.63	5.34	5.60	7.68		
Poor rotamers (%)	1	0.04	0.13	0.28		
Ramachandran plot						
Favored (%)	93.66	95.02	94.76	94.13	-	-
Allowed (%)	6.32	4.96	5.20	5.81		
Disallowed (%)	0.02	0.02	0.04	0.06		

Appendix 3 – Model and data collection statistics for NADH complex I dataset

	CxI-NADH – closed (EMD-11248) (PDB 6ZKG)	CxI-NADH – open1 (EMD-11249) (PDB 6ZKH)	CxI-NADH – open2 (EMD-11250) (PDB 6ZKI)	CxI-NADH – open3 (EMD-11251) (PDB 6ZKJ)	PA-focused refinement	MD-focused refinement
Data collection and processing						
Magnification	130000x	130000x	130000x	130000x	130000x	130000x
Voltage (kV)	300	300	300	300	300	300
Electron exposure (e–/Å^2)	89	89	89	89	89	89
Defocus range (μm)	~ -1 to -2	~ -1 to -2	~ -1 to -2	~ -1 to -2	~ -1 to -2	~ -1 to -2
Pixel size (Å)	1.061	1.061	1.061	1.061	1.061	1.061
Symmetry imposed	C1	C1	C1	C1	C1	C1
Initial particle images (no.)	182053	182053	182053	182053	182053	182053
Final particle images (no.)	14451	28253	64314	33612	64314	33612
Map resolution (Å) at 0.143 FSC threshold	3.4	3.0	2.8	3.0	2.8	3.0
Refinement						
Initial model used (PDB code)	5LNK	5LNK	5LNK	5LNK	5LNK	5LNK
Model resolution (Å)						
FSC threshold						
Map sharpening *B* factor (Å^2)	-61	-50	-43	-43		
Model composition						
Non-hydrogen atoms	67308	66906	67284	66949	-	-
Protein residues	66081	65506	65548	65548		
Ligands	1227	1400	1590	1401		
B factors (Å^2)						
Protein	47.7	50.3	46.1	52.1	-	-
Ligand	68.4	77.6	70.1	75.9		
R.m.s. deviations						
Bond lengths (Å)	0.0079	0.0091	0.0072	0.0083	-	-
Bond angles (°)	1.66	1.51	1.27	1.63		
Validation						
MolProbity score	1.74	1.62	1.57	1.61	-	-
EMRinger score	3.16	3.86	4.21	4.05		
Clashscore	6.54	5.54	5.51	5.38		
Poor rotamers (%)	0.13	0.04	0.06	0.11		
Ramachandran plot						
Favored (%)	94.44	95.45	96.05	95.43	-	-
Allowed (%)	5.54	4.46	3.93	4.52		
Disallowed (%)	0.02	0.09	0.02	0.05		

Appendix 4 – Model and data collection statistics for deactive complex I dataset

	Deactive CxI – open1 (EMD-11260) (PDB 6ZKS)	Deactive CxI – open2 (EMD-11261) (PDB 6ZKT)	Deactive CxI – open3 (EMD-11262) (PDB 6ZKU)	Deactive CxI – open4 (EMD-11263) (PDB 6ZKV)	PA-focused refinement	MD-focused refinement
Data collection and processing						
Magnification	81000x	81000x	81000x	81000x	81000x	81000x
Voltage (kV)	300	300	300	300	300	300
Electron exposure (e–/Å^2)	79	79	79	79	79	79
Defocus range (μm)	~ -1 to -2	~ -1 to -2	~ -1 to -2	~ -1 to -2	~ -1 to -2	~ -1 to -2
Pixel size (Å)	1.055	1.055	1.055	1.055	1.055	1.055
Symmetry imposed	C1	C1	C1	C1	C1	C1
Initial particle images (no.)	272990	272990	272990	272990	272990	272990
Final particle images (no.)	69703	76479	65539	61269	65539	61269
Map resolution (Å) at 0.143 FSC threshold	3.1	2.8	3.0	2.9	3.0	2.9
Refinement						
Initial model used (PDB code)	5LNK	5LNK	5LNK	5LNK	5LNK	5LNK
Model resolution (Å)						
FSC threshold						
Map sharpening B factor (Å^2)	-86.9	-83.8	-73.2	-81.0		
Model composition						
Non-hydrogen atoms	66105	66561	65296	65051	-	-
Protein residues	64931	64931	64122	63877		
Ligands	1174	1174	1174	1174		
B factors (Å^2)						
Protein	46.4	50.8	48.7	46.8	-	-
Ligand	71.1	76.7	70.6	68.8		
R.m.s. deviations						
Bond lengths (Å)	0.0079	0.0104	0.0091	0.0084	-	-
Bond angles (°)	1.78	1.39	1.82	1.80		
Validation						
MolProbity score	1.61	1.64	1.61	1.60	-	-
EMRinger score	3.43	4.04	3.65	3.67		
Clashscore	5.37	5.54	5.29	5.17		
Poor rotamers (%)	0.16	0.28	0.07	0.04		
Ramachandran plot						
Favored (%)	95.40	95.16	95.29	95.34	-	-
Allowed (%)	4.59	4.80	4.67	4.62		
Disallowed (%)	0.01	0.04	0.04	0.04		

Appendix 5 – Model and data collection statistics for rotenone complex I dataset

	CxI-rotenone – closed (EMD-11252) (PDB 6ZKK)	CxI-rotenone – open1 (EMD-11253) (PDB 6ZKL)	CxI-rotenone – open2 (EMD-11254) (PDB 6ZKM)	CxI-rotenone – open3 (EMD-11255) (PDB 6ZKN)	PA-focused refinement	MD-focused refinement
Data collection and processing						
Magnification	130000x	130000x	130000x	130000x	130000x	130000x
Voltage (kV)	300	300	300	300	300	300
Electron exposure (e–/Å^2)	89	89	89	89	89	89
Defocus range (μm)	~ -1 to -2	~ -1 to -2	~ -1 to -2	~ -1 to -2	~ -1 to -2	~ -1 to -2
Pixel size (Å)	1.061	1.061	1.061	1.061	1.061	1.061
Symmetry imposed	C1	C1	C1	C1	C1	C1
Initial particle images (no.)	142565	142565	142565	142565	142565	142565
Final particle images (no.)	7162	21296	61945	55837		
Map resolution (Å) at 0.143 FSC threshold	3.7	3.1	2.8	2.9		
Refinement						
Initial model used (PDB code)	5LNK	5LNK	5LNK	5LNK	5LNK	5LNK
Model resolution (Å)						
FSC threshold						
Map sharpening *B* factor (Å^2)	-100	-44	-71	-48		
Model composition						
Non-hydrogen atoms	67463	65983	65067	64771	-	-
Protein residues	66081	64649	63782	63486		
Ligands	1382	1336	1285	1283		
B factors (Å^2)						
Protein	49.4	49.8	47.5	50.3	-	-
Ligand	72.8	74.6	72.3	74.4		
R.m.s. deviations						
Bond lengths (Å)	0.0077	0.0086	0.0135	0.0082	-	-
Bond angles (°)	1.79	1.81	1.82	1.84		
Validation						
MolProbity score	1.72	1.65	1.56	1.67	-	-
EMRinger score	2.45	3.21	4.02	3.47		
Clashscore	6.22	5.80	5.61	6.18		
Poor rotamers (%)	0.06	0.09	0.04	0.14		
Ramachandran plot						
Favored (%)	94.49	95.13	96.27	95.18	-	-
Allowed (%)	5.47	4.83	3.72	4.76		
Disallowed (%)	0.04	0.04	0.01	0.06		

Appendix 6 – Model and data collection statistics for turnover complex I dataset

	Turnover CxI – closed (EMD-11244) (PDB 6ZKC)	Turnover CxI – open1 (EMD-11245) (PDB 6ZKD)	Turnover CxI – open2 (EMD-11246) (PDB 6ZKE)	Turnover CxI – open3 (EMD-11247) (PDB 6ZKF)	PA-focused refinement (EMD-11241) (PDB 6ZK9)	Open MD-focused refinement (EMD-11242) (PDB 6ZKA)	Closed MD-focused refinement (EMD-11243) (PDB 6ZKB)
Data collection and processing							
Magnification	130000x	130000x	130000x	130000x	130000x	130000x	130000x
Voltage (kV)	300	300	300	300	300	300	300
Electron exposure (e–/Å^2)	89	89	89	89	89	89	89
Defocus range (μm)	~ -1 to -2	~ -1 to -2	~ -1 to -2	~ -1 to -2	~ -1 to -2	~ -1 to -2	~ -1 to -2
Pixel size (Å)	1.061	1.061	1.061	1.061	1.061	1.061	1.061
Symmetry imposed	C1	C1	C1	C1	C1	C1	C1
Initial particle images (no.)	344541	344541	344541	344541	344541	344541	344541
Final particle images (no.)	15769	81780	98436	48396	344541	315484	29057
Map resolution (Å) at 0.143 FSC threshold	3.1	2.7	2.6	2.8	2.3	2.5	2.5
Refinement							
Initial model used (PDB code)	5LNK	5LNK	5LNK	5LNK	5LNK	5LNK	5LNK
Model resolution (Å)							
FSC threshold							
Map sharpening *B* factor (Å^2)	-50	-41	-39	-48	-38	-50	-50
Model composition							
Non-hydrogen atoms	67480	67343	66930	66601	29588	40817	39934
	66081	65487	65507	65507	27640	38066	38132
Protein residues	1399	1043	1301	1094	457	1455	1360
Ligands					1473	1296	442
Waters							
B factors (Å^2)							
Protein	46.2	47.5	45.7	53.3	65.5	46.3	80.7
Ligand	72.8	69.1	70.3	80.5	82.0	73.4	103.2
Waters					54.3	37.6	58.4
R.m.s. deviations							
Bond lengths (Å)	0.0078	0.0078	0.0084	0.0076	0.0096	0.0082	0.0076
Bond angles (°)	1.78	1.76	1.61	1.75	1.30	1.26	1.25
Validation							
MolProbity score	1.73	1.55	1.60	1.62	1.45	1.33	1.49
EMRinger score	3.61	4.12	4.40	4.09	6.62	6.03	4.91
Clashscore	7.48	5.59	6.60	6.45	4.84	4.41	5.03
Poor rotamers (%)	0.1	0.03	0.03	0.07	0.10	0.17	0.07
Ramachandran plot							
Favored (%)	95.39	96.32	96.41	96.10	96.79	97.47	96.51
Allowed (%)	4.57	3.67	3.55	3.89	3.15	2.53	3.47
Disallowed (%)	0.04	0.01	0.04	0.01	0.06	0	0.02