A universal coupling mechanism of respiratory complex I

Vladyslav Kravchuk, Olga Petrova, Domen Kampjut, Anna Wojciechowska-Bason, Zara Breese and Leonid Sazanov*

Affiliations: VK, OP, DK and LAS: Institute of Science and Technology Austria, Am Campus 1, 3400 Klosterneuburg, Austria. AWB and ZB: MRC Mitochondrial Biology Unit, Cambridge CB2 0XY, UK. DK current address: MRC Laboratory of Molecular Biology, Cambridge, CB2 0QH, UK. ZB current address: Sosei Heptares, Cambridge, CB21 6DG, UK.

Complex I is the first enzyme of the respiratory chain, responsible for energy production in mitochondria and bacteria[1]. It couples the transfer of two electrons from NADH to quinone and translocation of four protons across the membrane[2], with the coupling mechanism still hotly debated. Here we present high-resolution (up to 2.2 Å) cryo-EM structures of *Escherichia coli* complex I (*Ec*CI) in different redox states including catalytic turnover. *Ec*CI mostly exists in the open state, where the quinone (Q) cavity is exposed to the cytosol, allowing access for water molecules, which enable Q movements. Unlike in mammals[3], *Ec*CI can convert to the closed state only during turnover, showing that closed and open states are genuine turnover intermediates. The open-to-closed transition results in the tightly engulfed Q cavity connected to the central axis of the membrane arm, a source of substrate protons. Consistently, the proportion of closed state increases with increasing pH. We propose a detailed but straightforward and robust mechanism comprising a "domino effect" series of proton transfers and electrostatic interactions: the forward wave ("dominoes stacking") primes the pump and the reverse wave ("dominoes falling") results in the ejection of all pumped protons from the distal subunit NuoL. The mechanism naturally explains the NuoL-only proton exit pathway and is supported by our mutagenesis data. We contend that this is a universal coupling mechanism of complex I and related enzymes.

Complex I (CI) contains 14 conserved "core" subunits forming the hydrophilic peripheral arm (PA) and the membrane arm (MA), joined in an L-shape[4]. *Ec*CI consists of 13 subunits (since genes NuoC and NuoD are fused), representing the minimal (~550 kDa) CI version. Mammalian CI acquired 31 supplementary subunits,

increasing its mass to ~1 MDa[1,5-7]. PA consists of seven core subunits, where electrons are transferred from NADH to the quinone-binding (Q) site along the chain of Fe-S clusters (Figs. 1a and ED 2c). Quinone accepts two electrons from the terminal cluster N2[4,8], followed by two protons, forming quinol supplied further down the respiratory chain, and the generated proton motive force (pmf) drives ATP synthase[1]. MA consists also of seven core subunits, and the largest three (NuoL, M and N, *E. coli* nomenclature, ED Fig. 3e)) are antiporter-like subunits (ALS), homologous to each other and to the cation/H^+ MRP antiporters[9,10]. ALS are composed of two symmetry-related domains of five trans-membrane helices (TM) each, containing conserved lysines on broken TM7/12, connected by a central TM8 lysine (ED Fig. 4g). LysTM7 interacts with the conserved TM5 glutamate[4,9]. This series of protonatable residues form a central hydrophilic axis of MA, continued towards the Q site via subunits NuoK/J/A/H, forming the so-called E-channel containing many conserved glutamates.

The understanding of CI coupling mechanism evolved over time[3,8,11]. The MA architecture seems to suggest the translocation of one proton per ALS and one through the E-channel[4]. Most of redox energy is released not during Fe-S electron transfer but upon quinone reduction, suggesting its key role[8,12,13]. Initially it was proposed that quinone reactions initiate long-range conformational changes, facilitated by flexible broken helices TM7/8/12 along the central axis, leading to proton pumping[1,4,9]. Alternative theoretical proposals included forward and backward electrostatic waves, linked to formation of water wires in the ALS[14-16].

Recently, we reported structures of *Ovis aries* (ovine) CI (*Oa*CI) in several redox states including turnover, and proposed the coupling mechanism of mammalian CI involving cycling between the open and closed states[3]. In the open state key conserved loops around Q cavity (NuoCD 220-230 β1-β2 loop, NuoH 208-230 TM5-6 loop and PA/MA interface-flanking NuoA 41-62 TM1-2 loop, in *E. coli* residue numbering) unfold, while NuoB 83-90 loop changes the conformation[3]. This facilitates quinone binding, while the reduction of quinone happens only in the closed state, when key loops re-order and the cavity tightly engulfs the quinone. The transition to the closed state involves the rotation of the key ⌋TM3 (prefix indicates *E. coli* subunit), so that its π-bulge disappears and the water wire forms, therefore the protons for Q reduction come from the central axis. This brings redox "charge action"

into ALS, initiating proton pumping, driven purely electrostatically since we did not observe any conformational changes within ALS. We also tentatively suggested that a proton pathway to the periplasm is formed only in the distal ALS ND5/NuoL.

This radically novel mechanism raised debates and alternative proposals[11,17-20]. Only mammalian enzyme was observed so far in both closed and open states[6]. Other CI structures – bacterial[8,21], yeast[20], plant mitochondrial[22] and cyanobacterial NDH complex[23,24] show π-bulge in ⌟TM3, i.e. open-like state, although the degree of order of Q site loops varies. Therefore the validity of open-closed transition as part of catalytic cycle is questioned. Importantly, mammalian enzyme differs from other species by entering, in the prolonged absence of turnover, a deactive state, which can be converted back into active upon resumed sustained turnover[25]. The deactive state resembles the open state[7], but is distinct due to complete relocation of ⌟TM4[3]. Nevertheless, this similarity led to suggestions that closed/open states of mammalian enzyme should be considered as active/deactive[19]. Additionally, it is not clear how functionally important is the disorder of Q loops and how to explain the putative ND5-only proton ejection.

To answer these questions, we used *Ec*CI as an ideal model due to a rich library of mutants (Supplementary Table S8). Moreover, the bacterial enzyme is evolutionary very distant from mammalian and so the results may reveal the universal mechanistic principles. Here we present multiple structures of *Ec*CI in different redox states: native without additions (apo), with decyl-ubiquinone (DQ), with NADH, with an inhibitor piericidin A (PieA) and under turnover conditions (ED Table 1, Video S1, Supplementary figures S1-S10 and tables S1-S6). On the basis of structures and new mutagenesis data we propose a universal coupling mechanism of complex I: unexpectedly, most of proton movements happen not across but along the MA, in a series of electrostatically driven "domino effect" events.

Three conformations of *E. coli* complex I

With the enzyme purified entirely in DDM detergent[26] we observed only one state of *Ec*CI (Fig. 1bc), similar to a recent report[21]. In contrast to other CI structures, parts of the PA/MA interface were disordered, resulting in a completely exposed Q cavity. Unlike mammalian CI, bacterial enzyme does not show a pronounced deactive state[27], however, *Ec*CI can enter a more transient "resting" state of lower activity, from which

it can quickly recover under turnover[28]. DDM inhibits *Ec*CI[26] (ED Fig. 4e), therefore DDM datasets may show the "resting" state, where activity is uncoupled from proton pumping[28] due to disruption of Q cavity. The resting state has similarities to the deactive state of mammalian CI, but the disorder around Q cavity encompasses a larger area in *Ec*CI. However, in contrast to deactive *Oa*CI, $_{H}$TM5-6 loop is ordered in resting *Ec*CI, adopting "up" conformation (Fig. 2c and ED Fig. 4d) and pushing NuoCD so that PA is shifted about 8 Å away from MA (Fig. 1b).

To observe more native conformations, we diluted DDM-solubilised *Ec*CI stock in in a milder detergent LMNG, with added *E. coli* lipids, which showed optimal activity (ED Fig. 4e). We collected six such DDM/LMNG datasets in order to identify any specific features induced only by the turnover (or by reduction/Q/inhibitor) (ED Table 1). In all DDM/LMNG datasets the "resting" and "open" (not reported previously) states were observed. The resting state was similar to the DDM structures, while in the "open" state, the PA has joined back to the MA and the Q cavity was mostly reformed. The key loops ($_{H}$TM5-6 and $_{A}$TM1-2) were disordered and $_{J}$TM3 had a π-bulge, hence we termed this an "open" *Ec*CI state (Fig. 1d).

Exclusively under turnover (confirmed by several lines of experiments, Supplementary discussion §6) we observed a third *Ec*CI state. We termed it "closed" as it contained an enclosed Q cavity with all the key loops ordered, and rotated $_{J}$TM3 without π-bulge (Fig. 1e and Video S2). Notably, the fact that closed state was not observed in NADH, DQ or PieA *Ec*CI datasets suggests that neither reduction nor quinone/inhibitor binding alone can induce it, thus in *E. coli* the closed state is a higher energy intermediate and energy input during turnover is required for it to be observed.

In contrast to the mammalian enzyme, where the open state reflects the larger angle between the PA and MA due to the PA tilt, in *Ec*CI open and closed states differ mainly by the rotation of the PA (Fig. 1b). This suggests that similar conformational changes around the Q site do not necessarily lead to similar PA movements in different species. The "open" and "closed" states terms can still be applied for *Ec*CI and other species, referring mainly to the open and closed, respectively, Q cavity.

To verify whether the resting state could be a DDM-induced artefact, we solubilised and purified *Ec*CI entirely in LMNG. In apo and turnover datasets the proportion of the resting state dropped but remained significant (ED Table 1). Its

structure was unchanged, suggesting that the resting state is not an artefact of DDM exposure, but can be partially (DDM/LMNG datasets) or strongly (DDM datasets) promoted by DDM and associated de-lipidation (ED Fig. 4ab and Supplementary discussion §1). LMNG turnover structures also revealed a higher proportion of closed state (24%) than in DDM/LMNG (4%), suggesting overall stabilisation of *Ec*CI in milder detergent. As all non-mammalian species studied so far show only open state in the absence of turnover, the appearance of closed *Ec*CI state only under turnover, in three independent datasets, is a definite proof that closed/open states are true catalytic intermediates. High similarity of open-to-closed transition between *Ec*CI and *Oa*CI also confirms that open/closed states of mammalian enzyme should be considered as catalytic intermediates.

Quinone binding cavity

Quinone binds within an elongated cavity at the PA/MA interface. The description of the PA structure is provided in the Supplementary Discussion §2 and ED Figs. 2-3, since the overall structure of *Ec*CI was recently published[21]. CI-bound quinone previously was observed at the deep (Q_d) and shallow (Q_s) sites within the cavity, ~12 Å and ~24 Å from cluster N2, respectively[4,8,20]. In *Oa*CI both sites were occupied in the closed state under turnover[3], which is possible with short-tailed DQ, as the native quinone occupies the entire length of the cavity[19]. Notably, in the open state *Ec*CI (DDM/LMNG datasets) the only observed quinone was bound in the additional site ~ 16 Å from N2, with headgroup interacting with $_{CD}$Q328 and stacked against $_{B}$L86 (ED Fig. 5b). DQ, native UQ8 or piericidin A were bound in the same site, which we termed median, or Q_m (ED Fig. 6). The Q_d binding is prevented in all *Ec*CI open state structures by the extended conformation of the NuoCD β1-β2 loop, blocking the deep end of the cavity (Fig. 2b). In *Oa*CI open states this loop is disordered, with extended conformation only in the NADH-reduced state[3]. An extended loop is observed in the NDH complex[29], with plastoquinone (PQ) bound in Q_m position (ED Fig. 5c). A conserved NuoB alanine is replaced by $_{B}$Y65 in *E. coli* or by F54 in NDH, facing the cavity, which may create a bottleneck responsible for Q_m site, not present in other species (in open states of *Oa*CI site Q_s is occupied). This may explain why *Ec*CI shows lower affinities to many inhibitors than the mitochondrial enzyme[30].

In the closed *Ec*CI under turnover, NuoCD loop is retracted (ED Fig. 5a), allowing for DQ to bind in Q_d site. A second DQ molecule is observed close to Q_s site,

probably shifted from Q_m by the tail of Q_d-bound DQ (Fig. 1e and ED Fig. 7b). At the Q_d site DQ accepts electrons from cluster N2, forms H-bond with the conserved $_{CD}$Y277 and stacks against $_{CD}$H228 from NuoCD loop, similar to *Oa*CI[3] and *Tt*CI[8]. Along with NuoCD loop retraction, open-to-closed state transition includes ordering of $_{H}$TM5-6 and $_{A}$TM1-2 loops, rotation of $_{J}$TM3 and $_{H}$TM4 with flip of Y156 (Fig. 2a), and tilting of NuoH helices (ED Fig. 4c). These features faithfully reproduce the open-to-closed state transition in *Oa*CI[3].

*Ec*CI shows additional re-arrangements not observed in *Oa*CI (Fig. 2a and Videos S2-S3). $_{J}$TM3-4 loop is completely changes conformation, with W87 flip, while in *Oa*CI it was disordered in the open state. $_{CD}$LHL element, linking NuoC and NuoD, thus absent in other species, is also almost completely rebuilt, resulting in the rotation and shift of its helix. Together these re-arrangements help to "push in" the NuoA loop into the crevice between NuoCD and NuoB in the closed state, ordering the loop and closing Q cavity (Video S2). The conformation of the flexible $_{H}$TM5-6 loop, containing many conserved charged residues, is exactly the same in the closed states of *Ec*CI and *Oa*CI, consistent with its essential mechanistic role.

The open state, uniquely so far to *E. coli*, could be separated by 3D classification into two states - one "open" (described above) and another we termed "open-ready" (ED Fig. 1a), which likely represents an additional, previously not resolved, intermediate in the catalytic cycle (Supplementary discussion §7).

Proton translocation pathways

High resolution of the structures allowed us to identify a large number of water molecules (~860 in PA and ~430 in MA) and to reveal for the first time the MA hydration pattern in the bacterial CI. Apart from the protein surface, waters clustered around the central axis, connecting the key charged residues (ED Fig. 7a). The MA structures for resting, open and open-ready states did not show significant differences in the overall structure, therefore we discuss the highest resolution (2.3 Å in MA) open-ready LMNG turnover structure (Fig. 3a) and compare it to the closed LMNG turnover state (2.5 Å in MA).

Analysis of proton translocation pathways affirms[3] that the E-channel and NuoN/M ALS lack connections to the periplasm, blocked by large hydrophobic residues. Subunit NuoL is unique in having a highly hydrated exit to the periplasm, connecting LysTM12 (K399) to $_{L}$D400 and polar residues nearby. We also do not

observe any conformational changes in ALS between open/closed turnover states. Therefore, in all states of the complex, proton exit pathway into the periplasm is formed only in the distal ALS NuoL. This counter-intuitive feature is thus conserved from bacterial to mitochondrial[3,20] enzyme.

Following proton pathways along the central axis we see a highly hydrated connection all the way from key $_{L}$K399 to $_{N}$E133 via repeating Lys(Glu)TM12-Lys(His)TM8-LysTM7-GluTM5 series, linked by additional charged residues (labelled on Fig. 3a). Only NuoL and NuoM have a link to the cytoplasm via branching residues sitting on broken TM8. These residues ($_{L}$H254 and $_{M}$K265) could be able to switch the conformation (not linked to the redox state, ED Fig. 7de and Supplementary discussion §3) and thus help to re-distribute incoming protons along the central axis. The connection continues till the essential[9] NuoK residues: $_{N}$GluTM5(E133)-$_{K}$E72-$_{K}$E36. Then in all *Ec*CI open states there is a long (~13 Å) break in hydration, with hydrophobic residues from $_{J}$TM3 blocking the path from $_{K}$E36 to $_{A}$D79. However, in the closed state the cytoplasmic half of the $_{J}$TM3 rotates, π-bulge disappears and small residues G61-A62 replace A62-I63, which opens the cavity for waters to fill in the break (Fig. 3bc), creating the connection to $_{A}$D79. $_{H}$TM4 helix also rotates, so that the invariant Y156, sitting on another π-bulge, flips over ~180^{o} from facing the lipids in the open state directly into the created water path, helping to establish a firm Grotthuss connection (similarly to Y142 in *Oa*CI[3]). The conformation of this tyrosine is a distinct feature helping to easily recognise open or closed state (Figs. 2a and 3bc). From $_{A}$D79 a pathway continues via NuoH residues E157-Y156-H208-E216-D213-E218 and some ordered waters in the Q cavity (Fig. 3a, ED Fig. 7b) towards the $_{CD}$D329/H228 pair, a likely source of two substrate protons for quinone[3,4].

Since we did not observe any conformational changes within ALS under turnover conditions, electrostatics are likely to drive proton transfer within these subunits. Consistently, judging from density for carboxylate side-chains (almost disappearing in charged state), $_{H}$E157, $_{A}$D79, $_{K}$E36 and $_{N}$E133 were all neutral in open state but charged in closed turnover state (ED Fig. 7c), as also observed in *Oa*CI[3].

Because key TM12 residue in NuoM is a glutamate (E407) instead of lysine, we suggested previously that NuoM might pump protons in anti-phase with NuoL/N[3]. To test this hypothesis we mutated $_{M}$E407 to lysine and found that *Ec*CI was still active,

which suggests that in fact all three ALS act in a similar fashion. We also mutated residues around the main Q entry site, confirming that it is the only entry used by quinone, refuting recent proposals that Q may also enter from the cytosol[31] (ED Figs. 8-9, Supplementary discussion §4).

Open-to-closed state transition

All the hallmarks of the open-to-closed state transition are conserved between the evolutionarily distant bacterial and mammalian enzymes[3], suggesting that they are the key to the universal mechanism of complex I. In the open state, due to disorder of key loops, the Q site is open to the cytoplasm via the wide opening of the cavity (or branching tunnel) emanating roughly from the Q_m site, both in bacterial and mammalian enzymes (W site in Fig. 1d and ED Fig. 5ef). As a key novel feature of the mechanism that explains experimental observations (Supplementary discussion §5), we propose that this additional W site is essential for exit or entry of water molecules accompanying the entry or exit, respectively, of quinone via the main entry point from the lipid bilayer (Q in Fig. 1de). Before quinone entry the cavity is pre-filled with waters, therefore, without the extra "hole" (W) it will be difficult for quinone to get an access into the cavity, since its tail would be blocking the Q entry like a cork in the bottle (ED Fig. 5d). Similarly, when quinol exits the cavity, because of the blocking tail the waters must come in through the W site to fill in the vacated space within the cavity. In stark contrast, in the closed state the key loops get ordered, sealing off the cavity and tightly engulfing the bound quinone (Fig. 1e), so that waters cannot get into the cavity and protons for quinone protonation have to come from the central MA axis. This process would be optimal with the native long-tailed quinone but will also work with DQ, as the limited number of waters in the cavity (sealed by the lipids at the Q entry) will not be able to provide two protons.

Another defining feature of the open-to-closed transition is $_J$TM3 rotation. It is likely caused by a combination of tilting of NuoH TM helices, freeing up space for rotation, and a large shift of NuoCD β-sheet with β1-β2 loop (Fig. 2b). This sheet pulls along with it the tightly interacting $_A$TM1-2 and $_J$TM3-4 loops, which probably causes re-winding of $_J$TM3. Importantly, the pattern of two small residues followed by a large hydrophobic residue in NuoJ (Gly61-Ala62-Ile63) is fully conserved in all species. The π-bulges on $_H$TM4 and neighbouring $_H$TM8, allowing $_H$Y156 flipping,

are also conserved (Fig. 3bc). As this pattern is responsible for the creation of the water wire, this suggests that the transfer of "charge action" of quinone oxidoreduction towards ALS is conserved.

Since key glutamates in the E-channel are unprotonated in the closed state, the proportion of closed state could be expected to increase with increasing pH. Therefore we collected *Ec*CI turnover dataset at pH 8, for comparison with the initial pH 6 data (ED Table 1). Strikingly, the proportion of closed state indeed increased dramatically (from 4% to 15%). We also collected apo datasets (since mammalian enzyme shows apo closed state) for *Oa*CI at three pH values (ED Table 1), and the proportion of closed state was also observed to increase with pH. Importantly, the activity of *Ec*CI and *Oa*CI has actually decreased with pH (ED Fig. 4ef), firmly establishing that closed state is not equivalent to active state (since its proportion does not follow the activity), but instead is a part of catalytic cycle along with open state. The pattern of changes in ED Table 1 suggests that the pKa of key residues involved in open-to-closed transition is probably close to 8. The pH in mitochondrial matrix is about 8.0, and in *E. coli* cytoplasm about 7.6-7.8 [32], therefore the increased proportion of closed state at higher pH likely reflects on the *in vivo* situation.

Coupling mechanism

On the basis of a compendium of our bacterial and mammalian enzyme structures and mutagenesis data (Supplementary discussion §4 and ED Table S8) we propose the universal, applicable to all species, "domino effect" mechanism of complex I, depicted in Fig. 4 and fully described in Supplementary discussion §8. In essence, quinone/quinol binding and release happen in the open state (Steps 1, 4-5), enabled by waters coming via site W. The cycle starts with Step 1, where quinone binds and initiates the transition to the closed state (Step 2). Quinone is reduced and two protons are taken from the central MA axis to complete the reaction, which results in the re-distribution of protons, so that key TM12 residues are protonated and the charge is switched between TM5 and TM7 residues (Step 3). This is a highly energised state, akin to stacked dominos ready to fall. In the transition to open state, TM8 residues are protonated from the cytoplasm (Step 4), and $_{L}$TM12 proton is ejected into the periplasm due to electrostatic interactions. This initiates a series of proton transfers along the central axis due to appearance of a "vacancy" on the "left" of the chain and the electrostatic "pressure" of the incoming proton from the "right" (Step 5), akin to

stacked dominoes falling. This results in four protons in total ejected from NuoL and the cycle re-starting. Crucially, for the mechanism to work, $_{N/M}$TM12 protons must be transferred to the neighbouring $_{M/L}$TM5 and not directly to the periplasm, as otherwise the process will not be initiated in the next subunit (i.e. a domino will fall without tripping the next one), explaining NuoL-only exit.

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

Despite NuoL-only exit, all three ALS and the E-channel are essential, being responsible for the eventual transfer of one pumped proton each. Therefore, the varying number of ALS is related to the number of protons pumped per cycle in each of evolutionary-related complexes, such as MRP[10], MBH and MBS, according to the available redox energy[33]. The mechanism appears to be conserved: the Q-like cavity encloses different substrates, such as sodium ions, plastoquinone, hydrogen or polysulfide, while the principle of the redox charge action via the lateral proton transfer along the central axis remains fully applicable[10].

References

1 Sazanov, L. A. A giant molecular proton pump: structure and mechanism of respiratory complex I. *Nature reviews. Molecular cell biology* **16**, 375-388, (2015).

2 Jones, A. J., Blaza, J. N., Varghese, F. & Hirst, J. Respiratory Complex I in Bos taurus and Paracoccus denitrificans Pumps Four Protons across the Membrane for Every NADH Oxidized. *J Biol Chem* **292**, 4987-4995, (2017).

3 Kampjut, D. & Sazanov, L. A. The coupling mechanism of mammalian respiratory complex I. *Science* **370**, (2020).

4 Baradaran, R., Berrisford, J. M., Minhas, G. S. & Sazanov, L. A. Crystal structure of the entire respiratory complex I. *Nature* **494**, 443-448, (2013).

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

6 Fiedorczuk, K. *et al.* Atomic structure of the entire mammalian mitochondrial complex I. *Nature* **538**, 406-410, (2016).

7 Agip, A. A. *et al.* Cryo-EM structures of complex I from mouse heart mitochondria in two biochemically defined states. *Nat Struct Mol Biol* **25**, 548-556, (2018).

8 Gutierrez-Fernandez, J. *et al.* Key role of quinone in the mechanism of respiratory complex I. *Nature communications* **11**, 4135, (2020).

9 Efremov, R. G. & Sazanov, L. A. Structure of the membrane domain of respiratory complex I. *Nature* **476**, 414-420, (2011).

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

11 Muhlbauer, M. E. *et al.* Water-Gated Proton Transfer Dynamics in Respiratory Complex I. *J Am Chem Soc* **142**, 13718-13728, (2020).

12 Wikstrom, M., Sharma, V., Kaila, V. R., Hosler, J. P. & Hummer, G. New perspectives on proton pumping in cellular respiration. *Chem Rev* **115**, 2196-2221, (2015).

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

14 Verkhovskaya, M. & Bloch, D. A. Energy-converting respiratory Complex I: on the way to the molecular mechanism of the proton pump. *The international journal of biochemistry & cell biology* **45**, 491-511, (2013).

15 Kaila, V. R. I. Long-range proton-coupled electron transfer in biological energy conversion: towards mechanistic understanding of respiratory complex I. *Journal of the Royal Society, Interface* **15**, (2018).

16 Di Luca, A., Gamiz-Hernandez, A. P. & Kaila, V. R. I. Symmetry-related proton transfer pathways in respiratory complex I. *Proc Natl Acad Sci U S A* **114**, E6314-E6321, (2017).

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

18 Nuber, F. *et al.* A Quinol Anion as Catalytic Intermediate Coupling Proton Translocation With Electron Transfer in E. coli Respiratory Complex I. *Frontiers in chemistry* **9**, 672969, (2021).

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

20 Parey, K. *et al.* High-resolution structure and dynamics of mitochondrial complex I-Insights into the proton pumping mechanism. *Science advances* **7**, eabj3221, (2021).

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

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

23 Laughlin, T. G., Bayne, A. N., Trempe, J. F., Savage, D. F. & Davies, K. M. Structure of the complex I-like molecule NDH of oxygenic photosynthesis. *Nature* **566**, 411-414, (2019).

24 Schuller, J. M. *et al.* Structural adaptations of photosynthetic complex I enable ferredoxin-dependent electron transfer. *Science* **363**, 257-260, (2019).

25 Kotlyar, A. B. & Vinogradov, A. D. Slow active/inactive transition of the mitochondrial NADH-ubiquinone reductase. *Biochim. Biophys. Acta* **1019**, 151-158, (1990).

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

27 Kotlyar, A. B., Albracht, S. P. & van Spanning, R. J. Comparison of energization of complex I in membrane particles from Paracoccus denitrificans and bovine heart mitochondria. *Biochim. Biophys. Acta* **1365**, 53-59, (1998).

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

29 Pan, X. *et al.* Structural basis for electron transport mechanism of complex I-like photosynthetic NAD(P)H dehydrogenase. *Nature communications* **11**, 610, (2020).
30 Friedrich, T. *et al.* Two binding sites of inhibitors in NADH: ubiquinone oxidoreductase (complex I). Relationship of one site with the ubiquinone-binding site of bacterial glucose:ubiquinone oxidoreductase. *European journal of biochemistry / FEBS* **219**, 691-698, (1994).
31 Masuya, T., Uno, S., Murai, M. & Miyoshi, H. Pinpoint Dual Chemical Cross-Linking Explores the Structural Dynamics of the Ubiquinone Reaction Site in Mitochondrial Complex I. *Biochemistry* **60**, 813-824, (2021).
32 Zilberstein, D., Agmon, V., Schuldiner, S. & Padan, E. Escherichia coli intracellular pH, membrane potential, and cell growth. *J Bacteriol* **158**, 246-252, (1984).
33 Yu, H., Schut, G. J., Haja, D. K., Adams, M. W. W. & Li, H. Evolution of complex I-like respiratory complexes. *J Biol Chem*, 100740, (2021).

Methods

Sample preparation and imaging

*Ec*CI was purified from BL21 (DDM and DDM/LMNG datasets) or MC4100 (LMNG datasets) *E. coli* cells by anion exchange and size exclusion chromatography as described before[26]. This procedure gave a pure and active native protein preparation without any added purification tags. Concentrated *Ec*CI stocks were stored under liquid nitrogen in small aliquots in 20 mM Bis-Tris pH 6.0, 50 mM NaCl, 2 mM $CaCl_2$, 18% glycerol and ~2% DDM in case of DDM and DDM/LMNG datasets while LMNG purified protein was stored in 20 mM MES-HCl pH 6.0, 200 mM NaCl, 30% glycerol and ~0.7% LMNG. The amount of detergent in protein stocks was high because of the final concentrating procedure.

In the case of *Ec*CI datasets in DDM, the protein was run through Superose 6 SEC column immediately before grid preparation in 20 mM MES pH 6.0, 2 mM $CaCl_2$, 250 mM NaCl, 0.02% DDM and concentrated to ~10 mg/ml (~0.4% DDM final). The high protein concentration was required to achieve adequate coverage of the holes with particles. Additionally, an increased NaCl concentration in the buffer was required to prevent protein aggregation on the grid. For DDM_NADH dataset, 5 mM NADH was added immediately before freezing (20 seconds between NADH addition and vitrification). Quantifoil Cu/carbon 0.6/1 grids were used. Before using, grids were glow discharged (0.7 mbar and 30 mA for 2 min in the ELMO Glow Discharge unit, Agar Scientific, Stansted, UK). Blotting was done at 4°C, 100% humidity with 25 blotting force and 5s blotting time. All blotting and freezing procedures were done with 2.7 µL sample applied on a grid using FEI Vitrobot Mark IV.

Earlier, we have showed that *Ec*CI is inhibited by high DDM concentrations[26]. However, high protein concentration (therefore high concentration of DDM accumulated during the concentration) was required for optimal particle distribution in ice over grid holes. Therefore, to have highly active samples on a grid, we changed the detergent to LMNG and used carbon-support grids, which require much less concentrated sample.

In the case of datasets in DDM/LMNG, protein stocks were not subject to SEC but rather diluted (from ~17 mg/ml to ~0.5 mg/ml) in dilution buffer DB (20 mM MES-HCl pH 6.0, 250 mM NaCl, 2 mM $CaCl_2$ 0.01% LMNG) and incubated for at least 4 hours on ice. In case of LMNG-only preparation, the protein was used right

after purification, diluted in DB. *E. coli* total lipid extract (ETL; Avanti Polar Lipids Inc., Alabama, USA) was added to the protein sample before blotting and freezing. ETL stock was prepared by drying the initial stock (25 mg/ml in chloroform) under nitrogen gas stream, washing with diethyl ether, drying under N_2 and dissolving at 7.5 mg/ml in 10 mM MES pH6, 3% CHAPS buffer.

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

Since the use of carbon-coated grids allowed us to minimise protein concentration, under turnover substrate depletion was avoided, ensuring full turnover conditions at the time of snap-freezing. For the turnover datasets *Ec*CI (0.16 mg/ml for DDM/LMNG pH 6 and pH 8 and 0.25 mg/ml for LMNG) was mixed with ETL (0.25 mg/ml, 0.1% CHAPS) and DQ (410 μM for DDM/LMNG pH 6 and 750 μM for DDM/LMNG pH 8 and for LMNG) and incubated for 2 min at room temperature (22°C). Then NADH (1.5 mM) was rapidly added with mixing by aspiration, followed by application on a grid. It took ~20 s from NADH addition to protein

freezing and most of this time the sample was in the Vitrobot chamber kept at 15°C. Since the enzymatic activity at 18°C in lipids/DDM/LMNG is 5.3 μmol NADH/min/mg prot (ED Fig. 4e), the maximal amount of substrates to be used within 20 s (1/3 of min) would be ~ 5.3 x 0.16 x 1000 / 3 ~ 280 μM, ensuring that neither DQ or NADH are used up by the time of plunge-freezing. In addition, the exact same sample mixture and in the same conditions as used for grid application, was applied to a NanoDrop (Denovix DS-11) spectrophotometer kept at 22°C and the reaction kinetics was followed at 340 nm, confirming that it proceeded well beyond 20 s, significantly slowing only at about 40 s. The turnover of *Ec*CI in all cases at the time of freeze plunging is clearly confirmed by the presence of strong cryoEM density for FMN, NADH, NuoF and NuoE (ED Fig. 3b), which would otherwise be absent in the presence of NADH and absence of turnover (ED Fig. 3c). Finally, only under turnover we observe closed *Ec*CI and Q_d-bound quinone (in three independent datasets).

For *Oa*CI data, CI was purified as described previously[3]. After final size-exclusion purification step in a buffer, containing 50mM NaCl, 1mM EDTA, 0.002% LMNG and 20mM HEPES pH 7.4, CAPS pH9, or Sodium Acetate buffer pH 5.5 (referred to as pH 7.4, pH 9, and pH 5.5 conditions, respectively), *Oa*CI was concentrated to 3 mg/ml and used immediately for cryo-EM grid preparation. 0.2% CHAPS was added to the protein sample before grid preparation to improve ice quality and particle distribution. 2.7 μL sample was applied to a freshly glow-discharged Quantifoil 0.6/1 copper grid and blotted for 6-8 s with blotting force 25 at 4°C and 100% humidity in a FEI Vitrobot Mark IV. Grids were flash-frozen in liquid ethane and stored in liquid nitrogen.

Around 3000-3500 images were collected for most *Ec*CI datasets, while for DDM-LMNG pH8 and LMNG turnover datasets around 8000 and 11000 images were collected, respectively. All *Ec*CI datasets in DDM were collected with TF Krios TEM at CEITEC electron microscope facility in Brno, while all datasets in DDM/LMNG and LMNG were collected with TF Krios TEM at IST Austria electron microscope facility. *OaCI* datasets, of around 3000 images each, were collected with TF Glacios TEM at IST. Image collection settings and equipment are summarized in Supplementary Tables S2-7.

Activity measurements

509 Enzyme activity assays for *Ec*CI were done using a Shimadzu UV-2600 UV-VIS
510 spectrophotometer at 30 °C or 18 °C (lower limit) and with magnetic stirring (320
511 rpm). NADH:DQ and NADH:FeCy (ferricyanide) oxidoreduction activities were
512 measured by following NADH (ε = 6.1 mM^{-1} cm^{-1}) oxidation at 340 nm. Assay buffer
513 was very similar to the one used for the preparation of grids (20 mM MES pH6.0,
514 2mM $CaCl_2$, 250mM NaCl) except for the presence or absence of ETL (0.25 mg/ml,
515 with 0.1% CHAPS carry-over from ETL stock) and different detergent concentrations
516 (ED Fig. 4e). *Ec*CI was equilibrated for 3 min with 100 μM DQ before NADH
517 addition and NADH:DQ activity measurement. In all cases, the reaction was started
518 by NADH addition. Control experiments were performed in the presence of 30 μM
519 piericidin A inhibitor.
520 For *OaCI*, enzyme activity was measured in a similar manner in a buffer
521 containing 50mM NaCl, 1mM EDTA, 10% glycerol, 0.05% LMNG, 0.25 mg/mL
522 DOPC:cardiolipin (4:1) lipids, 1 mg/mL bovine serum albumin, and 20mM of either
523 HEPES pH 7.4, CAPS pH 9 or Sodium Acetate pH 5.5. Rotenone, when used, was
524 added to 5 μM final concentration.
525
526 Cryo-EM data processing

527 All data processing was done in RELION 3.0 and 3.1[34]. All datasets had a similar
528 processing strategy for the sake of consistency. Movies were aligned and dose-
529 weighted using MotionCor2[35]. CtfFind4 was used for CTF estimation on non-dose-
530 weighted micrographs[36]. About 3k movies were collected per dataset ((Supplementary
531 Figures S1-S10, Tables S2-6). Poor micrographs were filtered based on
532 AccumMotionTotal, CtfFigOfMerit and CtfMaxResolution parameters and manually
533 by CTF image (e.g., removing micrographs with the ice ring). Our general processing
534 procedure implemented four steps: picking, cleaning (using 2D and 3D classification),
535 refinement (using CtfRefine and Polishing tools in RELION[37] and class separation
536 using focus-revert-classify (FRC) strategy. The FRC strategy comprises initial
537 focused alignment with a mask around the peripheral arm, followed by classification
538 (with no angular or transitional searches) with a mask around the membrane arm[38].
539 Given the diversity in hinge angles between the peripheral and membrane arms,
540 aligning all the particles by the peripheral arm maximizes the differences during
541 classification focused on membrane arm. The peripheral arm was chosen for initial

alignment because it contains heavy FeS clusters, helping the alignment. 2D classifications were done with k = 100 and T = 2 unless otherwise stated.

For DDM_Apo (Supplementary Fig. S1), the first dataset collected, 2D classes from manually picked 3k particles were used as a template to initially pick 500k particles using RELION's Autopick. Particles were extracted 2X binned and classified using 2D and 3D classifications. Low-pass filtered (40 Å) *Tt*CI (PDB 4HEA [4]) structure was used as the initial model at the very first 3D classification. Best 3D class was used as a template to pick 531k particles with Gautomatch v0.56 (K. Zhang software) and 525k particles with RELION's Autopick, leaving 755k unique particles after duplicates removal. Particles were extracted 2X binned and cleaned with one round of 2D and four rounds of 3D classifications. The first 3D classification was performed with k = 6 and T = 4 in two steps: 50 iterations with 7.5° global angular search and 10 iterations with 3.7° global angular search (569k particles remained). Further we performed 2D classification (560k particles remained). The second 3D classification was performed with k = 4 and T = 4 in two steps: 50 iterations with 7.5° global angular search and 20 iterations with 3.7° global angular search. Next, the duplicates were again removed and 467k particles remained. The third 3D classification was performed with k = 4 and T = 8 in two steps: 50 iterations with 7.5° global angular search, 10 iterations with 3.7° global angular search and 30 iterations with 1.8° local angular search (371k particles remained). The fourth 3D classification was done without angular searches with k = 5 and T = 6 (264k particles remained). Before the fourth 3D classification, the particles were refined with the loose mask around the molecule. After classifications, clean particles were refined into a single "consensus" structure, followed by per-particle defocus and per-particle trajectory refinement using CtfRefine and Polishing tools (CtfRefine – Bayesian polishing – CtfRefine). Refined particles were subjected to FRC classification with k = 6 and T = 10. This resulted in six resting state classes that differed from each other only by the degree of openness. The three best classes were combined into one consensus resting class. The final density map was formed from focus-refined PA and MA maps. The PA and MA maps were aligned on the most populated class and merged into the final composite map using the "vop max" command in Chimera software[39].

For DDM_NADH dataset (Supplementary Fig. S2), 303k particles were picked with Gautomatch and 306k particles with RELION's Autopick using *Ec*CI structure as a 3D template and joined with duplicates removal (441k unique particles). Particles

were extracted 2X binned and cleaned using one round of 2D (290k particles remained) and one round of 3D classifications (226k particles remained). The 3D classification was done with k = 4 and T = 8 in four steps: 45 iterations with 7.5° global angular search, 10 iterations with 3.7° global angular search, 25 iterations with 1.8° local angular search and 10 iterations with 0.9° local angular search. Good classes were re-extracted at the full pixel size, followed by duplicates removal (194k particles remained). Next, particles were refined into a single "consensus" structure, followed by per-particle defocus and per-particle trajectory refinement using CtfRefine and Polishing tools. Refined particles were subject to FRC classification with k = 6 and T = 4. This resulted in six resting state classes that differed from each other only by the degree of openness. The four best classes were combined into one consensus resting class. The final density map was formed from focus-refined PA and MA maps. The PA and MA maps were aligned on the most populated class (class 4) and merged using "vop max" command in Chimera software.

For DDM/LMNG_Turnover_pH6 dataset (Supplementary Fig. S3), 773k particles were picked with RELION's Autopick using *Ec*CI structure as a 3D template. Particles were extracted 2X binned and cleaned using one round of 2D (629k particles remained) and one round of 3D classifications (269k particles remained). The 3D classification was done with k = 6 and T = 4 in three steps: 25 iterations with 7.5° global angular search, 5 iterations with 3.7° global angular search and 10 iterations with 1.8° local angular search. Good classes were re-extracted at the full pixel size. Next, particles were refined into a single "consensus" structure, followed by per-micrograph aberrations (two rounds), per-particle defocus (two rounds) and per-particle trajectory refinement using CtfRefine and Bayesian polishing tools. Another cleaning step was performed: particles were aligned on the PA and classified without searches with a loose mask around PA; then same was done for MD (204k particles remained). Afterwards, we performed FRC classification with k = 4 and T = 4. This resulted in two resting classes and two open classes that differed from each other by degree of openness. The two resting classes were combined into one consensus resting class. The final density map was formed from focus-refined PA and MA maps. The PA and MA maps were aligned on the most populated resting class (class 3) and merged using "vop max" command in Chimera software. The particles from the two open classes were joined and FRC classified with k = 6 and T = 10. This resulted in three good classes: one closed and two open. The two open classes were combined

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

For DDM/LMNG_Apo dataset (Supplementary Fig. S4), 918k particles were picked with RELION's Autopick using *Ec*CI structure as a 3D template. Particles were extracted at 2x the physical pixel size and cleaned using two rounds of 2D (487k particles remained) and one round of 3D classifications (387k particles remained). The first round of 2D classification was done with k =100 and T = 2 and the second round was done with k = 20 and T = 2.5 on each good class (obtained from the first round) separately. Afterwards, all good 2D classes were joined together. The 3D classification was done with k = 4 and T = 4 in three steps: 20 iterations with 7.5° global angular search, 10 iterations with 3.7° global angular search and 15 iterations with 1.8° local angular search. Good classes were re-extracted at the full pixel size. Next, particles were refined into a single "consensus" structure, followed by per-micrograph aberrations (two rounds), per-particle defocus (two rounds) and per-particle trajectory refinement using CtfRefine and Bayesian polishing tools. Another cleaning step was performed: particles were aligned on the PA and classified without searches with a loose mask around PA; then same was done for MA (366k particles remained). Afterwards, we performed FRC classification with k = 4 and T = 4. This resulted in two resting state classes and two open state classes that differed from each other by the degree of openness. The two resting classes were combined into one consensus resting class. The final density map was formed from focus-refined PA and MA maps. The PA and MA maps were aligned on the most populated resting state class and merged using "vop max" command in chimera software. The same procedure was done to obtain the final composite open class map.

For DDM/LMNG_PieA dataset (Supplementary Fig. S5), 463k particles were picked with Gautomatch using 2D projections of 3D *Ec*CI structure (obtained from previous datasets) as a template. Particles were extracted 2X binned and cleaned using

two rounds of 2D (256k particles remained) and one round of 3D classifications (233k particles remained). The first round of 2D classification was done with k =100 and T = 2 and the second round was done with k = 20 and T = 2.5 on each good class (obtained from the first round) separately. Afterwards, all good 2D classes were joined together and duplicates removed. The 3D classification was done with k = 5 and T = 4 in three steps: 25 iterations with 7.5° global angular search, 12 iterations with 3.7° global angular search and 7 iterations with 1.8° local angular search. Good classes were re-extracted at the full pixel size. Next, particles were refined into a single "consensus" structure, followed by per-micrograph aberrations (two rounds), per-particle defocus (two rounds) and per-particle trajectory refinement using CtfRefine and Bayesian polishing tools. Another cleaning step was performed: particles were aligned on the PA and classified without searches with a loose mask around PA; then same was done for MA (214k particles remained). Afterwards, we performed FRC classification with k = 4 and T = 4. This resulted in three poor (low resolution) resting state classes and one open state class. Because of low resolution (> 4 Å) and absence of differences between the other resting classes, we did not model DDM/LMNG_PieA resting class. The final density map for the open class was formed from focus-refined PA and MA maps. The PA and MA maps were aligned on the entire initial map and merged using "vop max" command in Chimera software.

For DDM/LMNG_NADH+FMN dataset (Supplementary Fig. S6), 563k particles were picked with AutoPick using *Ec*CI structure as a 3D template. Particles were extracted 2X binned and cleaned using one round of 2D (498k particles remained) and one round of 3D classifications (145k particles remained). The 3D classification was done with k = 6 and T = 4 in three steps: 20 iterations with 7.5° global angular search, 5 iterations with 3.7° global angular search and 15 iterations with 1.8° local angular search. Good classes were re-extracted at the full pixel size. Next, particles were refined into a single "consensus" structure, followed by per-micrograph aberrations (two rounds), per-particle defocus (two rounds) and per-particle trajectory refinement using CtfRefine and Bayesian polishing tools. Another cleaning step was performed: particles were aligned on the PA and classified without searches with a loose mask around PA; then same was done for MA (122k particles remained). Afterwards, we performed FRC classification with k = 4 and T = 4. This resulted in three poor (low resolution) resting classes and one open class. Because of low resolution (> 4 Å) and absence of differences between the other resting classes, we did not model

DDM/LMNG_NADH+FMN resting class. The final density map for the open class was formed from focus-refined PA and MA maps. The PA and MA maps were aligned on the entire initial map and merged using "vop max" command in Chimera software.

For DDM/LMNG_DQ dataset (Supplementary Fig. S7), 825k particles were picked with RELION's Autopick using *Ec*CxI structure as a 3D template. Particles were extracted 2X binned and cleaned using two rounds of 3D classifications. The first round of 3D classification was done with k = 6 and T = 4, 20 iterations with 7.5° global angular search (417k particles remained). The second round of 3D classification was done with k = 4 and T = 4 in three steps: 15 iterations with 7.5° global angular search, 10 iterations with 3.7° global angular search and 15 iterations with 1.8° local angular search (137k particles remained). Good classes were re-extracted at the full pixel size. Next, particles were refined into a single "consensus" structure, followed by per-micrograph aberrations (two rounds), per-particle defocus (two rounds) and per-particle trajectory refinement using CtfRefine and Bayesian polishing tools. Another cleaning step was performed: particles were aligned on the PA and classified without searches with a loose mask around PA; then same was done for MD (123k particles remained). Afterwards, we performed FRC classification with k = 4 and T = 4. This resulted in three resting state classes and one open state class that differed from each other by the degree of openness. The three resting state classes were combined into one consensus resting state class. The final density map was formed from focus-refined PA and MA maps. The PA and MA maps were aligned on the most populated resting state class and merged using "vop max" command in Chimera software. The same procedure was done to obtain the final composite open state class map.

For DDM/LMNG_Turnover_pH8 dataset (Supplementary Fig. S8), 1.65 million particles were picked with RELION's Autopick using *Ec*CI structure as a 3D template. Particles were extracted 2X binned and cleaned using one round of 2D (870k particles remained) and one round of 3D classifications (325k particles remained). The 3D classification was done with k = 4 and T = 4 in three steps: 20 iterations with 7.5° global angular search, 10 iterations with 3.7° global angular search and 10 iterations with 1.8° local angular search. Good classes were re-extracted at the full pixel size. Next, particles were refined into a single "consensus" structure, followed by per-micrograph aberrations (two rounds), per-particle defocus (two

rounds) and per-particle trajectory refinement using CtfRefine and Bayesian polishing tools. Afterwards, we performed FRC classification with k = 6 and T = 10. This resulted in two resting classes, two open classes, one closed and one junk class. The two resting classes were combined into one consensus resting class. The final density map was formed from focus-refined PA and MA maps. The PA and MA maps were aligned on the most populated resting class and merged using "vop max" command in Chimera software. The particles from the two open and the closed classes were joined and FRC classified with k = 6 and T = 64. This resulted in three open, one closed and two open-ready classes. Same type classes were combined into consensus classes. The final density map for each class was formed from focus-refined PA and MA maps filtered by local resolution. The PA and MD maps were aligned on the most populated class and merged using "vop max" command in chimera software.

For LMNG_Apo dataset (Supplementary Fig. S9), ~1 million particles were picked with RELION's Autopick using *Ec*CI structure as a 3D template. Particles were extracted 2X binned and cleaned using one round of 2D (550k particles remained) and one round of 3D classifications (240k particles remained). The 3D classification was done with k = 4 and T = 4 in three steps: 20 iterations with 7.5° global angular search, 10 iterations with 3.7° global angular search and 10 iterations with 1.8° local angular search. Good classes were re-extracted at the full pixel size. Next, particles were refined into a single "consensus" structure, followed by per-micrograph aberrations (two rounds), per-particle defocus (two rounds) and per-particle trajectory refinement using CtfRefine and Bayesian polishing tools. Afterwards, we performed FRC classification with k = 6 and T = 10. This resulted in two resting classes, two open-ready classes, one open and one junk class. Structures of the resting and open classes were not built because of poor density (4.5-5 Å resolution). The two open-ready classes were combined into one consensus class. The final density map was formed from globally filtered (to the value of 0.143 FSC cut-off) focus-refined maps of NuoFEG, NuoAJKHCD (the PA-MA junction subunits) and NuoNML combined using "vop max" command in chimera software.

For LMNG_Turnover dataset (Supplementary Fig. S10), ~4.5 million particles were picked with RELION's Autopick using *Ec*CI structure as a 3D template. Particles were extracted 2X binned and cleaned using one round of 2D (1.6 million particles remained) and one round of 3D classifications (708k particles remained). The 3D classification was done with k = 4 and T = 4 in three steps: 20 iterations with

7.5° global angular search, 10 iterations with 3.7° global angular search and 10 iterations with 1.8° local angular search. Good classes were re-extracted at the full pixel size. Next, particles were refined into a single "consensus" structure, followed by per-micrograph aberrations (two rounds), per-particle defocus (two rounds) and per-particle trajectory refinement using CtfRefine and Bayesian polishing tools. Afterwards, we performed FRC classification with k = 6 and T = 10. This resulted in 5 mixed (open, open-ready, resting) classes, and one closed class. The final density map for the closed class was formed from focus-refined PA and MA maps. The particles from the five classes were joined and FRC classified with k = 4 and T = 16. This resulted in two resting and two open-ready classes. The resting classes were combined into one consensus class. Open-ready classes were merged and subjected of another FRC classification with k = 6 and T = 64. This resulted in 5 open-ready classes with slightly different PA-MA angle and one open class. Open-ready classes were merged into one consensus class. The final density maps for each class were formed from focus-refined PA and MA maps filtered by local resolution. The PA and MA maps were aligned on the most populated class and merged using "vop max" command in chimera software.

All *Oa*CI datasets were processed in a similar manner to ensure that different pH data can be compared directly (Supplementary Figures S11-S13, Table S7). Processing was done in RELION 3.1. Movie frames were aligned using MotionCor2 and initial CTF parameters were estimated from averaged images using CTFFIND 4.1.14. For the further processing steps, only the micrographs where the Thon rings extended to 7 Å were included. Autopicking was done in RELION using *OaCI* structure as a 3D template and resulted in 557585 particles from 2931 micrographs for pH 7.4 conditions; 557664 particles from 2097 micrographs for pH 5.5 conditions; 660526 particles from 2799 micrographs for pH 9 conditions. Particles were extracted 2X binned and cleaned using one round of 2D classification (547k particles, 540k particles, and 615k particles remained for pH 7.4, pH5.5, and pH 9 conditions, respectively), followed by one round of 3D classification, performed in the following manner: k = 6 and T = 4 were used in three steps: 25 iterations with 7.5° global angular search, 25 iterations with 3.7° global angular search and 25 iterations with 1.8° local angular search. Good classes were then extracted at full pixel size, resulting in 459315 particles, 258453 particles, and 371884 particles for pH 7.4, pH5.5, and pH 9 conditions, respectively. CTF parameters and per-particle trajectories were then

refined in an iterative manner: one round of CTF refinement followed by Bayesian polishing and another round of CTF refinement. This led to consensus refined complex I structures. Afterwards, particles were further 3D classified without searches to remove damaged particles and then focus-reverse-classified to sort out the heterogeneity as described above. All focus-reverse classifications were performed in the same manner, using 6 classes, T=16 and 35 iterations. In case of pH 7.4 and pH 5.5 conditions, this resulted in four good classes (one closed and three open) with 164501 and 130530 particles in total and two partially broken classes. The good classes were refined and post-processed separately and the open classes were joined, refined and post-processed to give final maps for model building. In case of pH 9 condition this resulted only in two good classes, one open and one closed, which were refined and post-processed separately for subsequent model building. The final density maps for each class of each condition were formed from focus-refined peripheral arm (PA) and membrane arm (MA) maps filtered by local resolution. The PA and MA maps were then aligned on the corresponding consensus map and merged using "vop max" command in Chimera software.

All resolutions are based on the gold-standard (two halves of data refined independently) FSC = 0.143 criterion. Local masks used for focused refinement correspond to regions shown as local resolution maps of PA and MA in Supplementary Figures S1-S13. Masks were created in RELION with extend 7 and soft-edge 10 pixels command. All maps were post-processed, B-factor sharpened and filtered by local resolution in RELION. The density for the weaker features, such as some of bound quinones, is better defined in non-sharpened maps – these can be reproduced, if needed, using B-factor "Blur" feature in Coot.

Model building and analysis

Initial models for the subunits of the peripheral arm, NuoH and NuoA were generated using homology modeling implemented in Phyre2 server[40], using *Tt*CI (PDB 4HEA) as a template. Coordinates for the rest of the subunits were from the X-Ray structure of *Ec*CI membrane arm (PDB ID 3RKO[9]). All the subunits were fitted as rigid bodies with UCSF Chimera into the best resolved focus-refined maps of the peripheral and membrane arms. One round of five cycles of a real space refinement workflow in PHENIX[41] software was applied to resolve atom clashes and create the

initial structure. Further, the initial structure was manually corrected in Coot[42] with *de-novo* rebuilding of incorrectly homology-modelled parts (mainly in NuoG, NuoCD and NuoB). Secondary structure identification and modelling of poor density regions were assisted by data from PredictProtein server[43]. The initial structure was further improved by iterating manual adjustment in Coot and automated real space refinement in PHENIX, using our script which performs two rounds of a single cycle of ADP refinement with subsequent three cycles of global energy minimization to optimize B-factors so that electron radiation-damaged carboxylate side-chains acquire high B-factors and do not lead to main-chain distortions[38]. This way we built template structures of the peripheral and membrane arms against the best resolved focus-refined maps. These template structures had been fit as rigid bodies into each class, followed by adjustment and correction with the iterative use of Coot and PHENIX. Topologies for lipids and ligands for refinement in Coot and PHENIX were generated using the grade web-server (http://grade.globalphasing.org). In order to reliably build experimental water molecules, the maps were filtered by local resolution and resampled at 0.5 Å per pixel using relion_image_handler. After this procedure, water molecules displayed strong signal (>2σ), had nearly spherical densities, were not clashing with other atoms, and participated in hydrogen bonds, which are all strongly indicative of real water molecules. This allowed automatic placement of water molecules in Coot, which were then all checked and corrected manually guided by "undowse" feature of Molprobity web-server[44]. Using this protocol, we could reliably build water molecules in the maps of higher than 2.8 Å resolution. The same resampled maps were used for depiction purposes (ED Fig. 1).

*Oa*CI models are based on the structures of ovine complex I determined previously[3] (PDB ID 6ZKC for closed and 6ZKE for open state). The initial structures were improved by iterating manual adjustment in Coot and automated real space refinement in PHENIX as described above. The models contain NADH, since the density for bound nucleotide, although relatively weak, was still observed due to pre-turnover on membranes before purification.

Cavities and channels inside the protein were predicted using MOLE web-server[45]. The overall quality of the models was assessed using Molprobity[46], Q-scores[47] and EMRinger[48]. Visualization and analysis of protein density and structure,

as well as Figures preparation were done using PyMol, UCSF Chimera and ChimeraX[39].

Site-directed mutagenesis studies

E. coli MC4100 (F^-, araD139, $^-$(arg F^-lac)U169, ptsF25, relA1, flb5301, rpsL 150.λ^-)[49] was used to generate site-specific mutations. The strategies used for generating knock-out mutants (ΔM, ΔN and ΔH) and mutagenesis of the *E. coli* *nuoM, nuoN* and *nuoH* genes were similar to those reported previously[50,51] with minor modifications. The knock-out mutants were generated by employing the pKOV system and the pKOV vector was purchased from Addgene (Addgene plasmid 25769). In brief, the *kat* gene was inserted into the *nuoN* gene using a NdhI restriction site to disrupt the *nuoN* gene, leading to the construction of the *E. coli* ΔN. In parallel, the *nuoN* gene together with a 500-pase pair DNA segment, both upstream and downstream, was cloned into the pCR-TOPO XL system to generate a template for the site-specific *nuoN* mutations. The mutated *nuoN* fragments were inserted into pKOV using the restriction sites NotI and SalI to construct pKOV (*nuoN* mutants). Then, the above pKOV plasmids were used to replace the *kat* gene in *E. coli* ΔN by recombination. The mutagenesis of *nuoM* and *nuoH* genes was done in a similar manner. The point mutations in the chromosome were confirmed by DNA sequencing (Source BioScience, Nottingham, UK) of amplified gDNA fragments containing the mutation.

Growth tests were carried out under aerobic conditions at 37 °C with shaking at 250 rpm in 25mL of minimal M9 media (22.5mM $Na_2HPO_4.2H_2O$, 22 mM KH_2PO_4, 19 mM NH_4Cl and 8.5 mM NaCl) supplemented with malate (2 g/L), 10 μM $FeSO_4.7H_2O$, 1 mM $MgSO_4.7H_2O$ and 20 μM $CaCl_2$ (M9+malate media). The M9+malate cultures were inoculated with overnight aerobic pre-cultures of *E. coli* strains grown in rich LB (Lysogeny Broth) media at 37ºC with shaking at 250 rpm. The cells were diluted to the OD_{600} of 5.0 before inoculating the M9+malate cultures to the starting OD_{600} of 0.1. OD_{600} was then measured every 60-90 minutes. In rich LB media complex I-deletion mutants grow as WT, since *E. coli* employs alternative pathways. However, in M9+malate media in mutants with impaired complex I activity the lag phase before entering the exponential growth phase is extended in comparison

to WT. The extent of the lag is roughly proportional to the degree of complex I impairment.

For preparation of inverted vesicles, a single colony was inoculated into 1L of LB media and grown until the optical density at A_{600} of 3. Cells were then harvested at 6,800 *g* for 10 min and re-suspended in buffer containing 50 mM Bis-Tris (pH 6.5), 1mM PMSF and 10% glycerol. Then, the cell suspensions were sonicated five times for 5 s, and centrifuged at 6,800 *g* for 10 min. The supernatant was ultracentrifuged at 260,000 *g* for 45 min. The pellet was re-suspended in the same buffer as described above. The resulting membrane suspension was stored in small aliquots at -80 °C until use.

The activity assays were conducted using a Shimadzu UV-1601 spectrophotometer (Shimadzu, Milton Keynes, UK). dNADH oxidase activity of membrane samples was assayed at 340 nm (ε_{340} (d)NADH = 6220 M^{-1} cm^{-1}) in 50 mM Bis-Tris (pH 6.0), 2 mM $CaCl_2$, 100 mM NaCl buffer started by the addition of 0.1 mM deamino-NADH (dNADH). The dNADH:DQ reductase activity measurements were conducted in a similar manner, except that 20 mM KCN and 0.1 mM DQ were also included in the assay mixture. The dNADH:ferricyanide reductase activity was measured in the presence of 0.1 mM dNADH and 1 mM ferricyanide in the same buffer. When using purified complex I, 5-10 μg of purified complex I and 0.25 mg mL^{-1} *E. coli* total lipids in 2% (w/v) CHAPS) were added to the buffer, KCN excluded from the assay and NADH used in place of dNADH. For all assays, at least 3 measurements were made, and the mean ± S.D. calculated. The H^+-pumping activity was followed by ACMA (9-amino-6-chloro-2-methoxyacridine) fluorescence quenching. Membrane vesicles equivalent to complex I dNADH: ferricyanide activity of 0.5 μmol dNADH min^{-1} mg^{-1} were added to buffers described in ED Figs 8-9. The reaction was started by addition of 0.1 mM dNADH. Uncouplers FCCP or CCCP were added to dissipate the potential. Fluorescence was monitored in a Shimadzu RF-5301 PC dual wavelength spectrophotometer.

Additional references

34 Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, (2018).

35 Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332, (2017).

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

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

38 Letts, J. A., Fiedorczuk, K., Degliesposti, G., Skehel, M. & Sazanov, L. A. Structures of Respiratory Supercomplex I+III2 Reveal Functional and Conformational Crosstalk. *Mol Cell* **75**, 1131-1146.e1136, (2019).

39 Pettersen, E. F. *et al.* UCSF chimera - A visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612, (2004).

40 Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**, 845-858, (2015).

41 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221, (2010).

42 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126-2132, (2004).

43 Yachdav, G. *et al.* PredictProtein--an open resource for online prediction of protein structural and functional features. *Nucleic Acids Res* **42**, W337-343, (2014).

44 Prisant, M. G., Williams, C. J., Chen, V. B., Richardson, J. S. & Richardson, D. C. New tools in MolProbity validation: CaBLAM for CryoEM backbone, UnDowser to rethink "waters," and NGL Viewer to recapture online 3D graphics. *Protein Sci* **29**, 315-329, (2020).

45 Pravda, L. *et al.* MOLEonline: a web-based tool for analyzing channels, tunnels and pores (2018 update). *Nucleic Acids Res* **46**, W368-W373, (2018).

46 Williams, C. J. *et al.* MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci* **27**, 293-315, (2018).

47 Pintilie, G. *et al.* Measurement of atom resolvability in cryo-EM maps with Q-scores. *Nat Methods* **17**, 328-334, (2020).

48 Barad, B. A. *et al.* EMRinger: side chain-directed model and map validation for 3D cryo-electron microscopy. *Nat Methods* **12**, 943-946, (2015).

49 Peters, J. E., Thate Te Fau - Craig, N. L. & Craig, N. L. Definition of the Escherichia coli MC4100 genome by use of a DNA array. *J. Bacteriology* **185**, 2017-2021.

50 Link, A. J., Phillips, D. & Church, G. M. Methods for generating precise deletions and insertions in the genome of wild-type Escherichia coli: application to open reading frame characterization. *J Bacteriol* **179**, 6228-6237, (1997).

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

52 Schulte, M. *et al.* A mechanism to prevent production of reactive oxygen species by Escherichia coli respiratory complex I. *Nature communications* **10**, 2551, (2019).

53 Holt, P. J., Efremov, R. G., Nakamaru-Ogiso, E. & Sazanov, L. A. Reversible FMN dissociation from Escherichia coli respiratory complex I. *Biochim. Biophys. Acta* **1857**, 1777-1785, (2016).

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

55 Roessler, M. M. *et al.* Direct assignment of EPR spectra to structurally defined iron-sulfur clusters in complex I by double electron-electron resonance. *Proc Natl Acad Sci U S A* **107**, 1930-1935, (2010).

56 Reed, J. L. *et al.* Systems approach to refining genome annotation. *Proc Natl Acad Sci U S A* **103**, 17480-17484, (2006).

Acknowledgments

This research was supported by the Scientific Service Units (SSU) of IST Austria through resources provided by the Electron Microscopy Facility (EMF), the Life Science Facility (LSF) and the IST high-performance computing cluster. We thank Victor-Valentin Hodirnau from IST Austria, Michal Babiak from CEITEC for assistance with collecting cryo-EM data and Alexej Charnagalov for the assistance with protein purification. VK was a recipient of a DOC Fellowship of the Austrian Academy of Sciences at the Institute of Science and Technology, Austria. VK and OP are funded by the ERC Advanced Grant 101020697 RESPICHAIN to LAS.

Author contributions

VK performed biochemical procedures with *Ec*CI, prepared cryo-EM grids, acquired and processed cryo-EM data, built and analyzed atomic models and analyzed data. DK prepared cryo-EM grids of *Ec*CI, acquired DDM datasets and analyzed data. OP purified *Oa*CI, prepared cryo-EM grids with *Oa*CI, acquired and processed cryo-EM data, built and analyzed atomic models and analyzed data. AWB and ZB performed mutagenesis studies. LS designed and supervised the project, acquired funding, analyzed data and models and wrote the manuscript with contributions from all authors.

Author Information

Data availability statement: Cryo-EM maps and corresponding atomic models are deposited in EMDB/PDB with the following accession IDs.

*Ec*CI DDM datasets: Apo Resting (PDB ID 7P62, EMD-13215) and NADH Resting (PDB ID 7P61, EMD-13214).

*Ec*CI DDM/LMNG datasets: Apo Open (PDB ID 7P7C, EMD-13235), Apo Resting (PDB ID 7P7E, EMD-13236), DQ Open (PDB ID 7P7J, EMD-13237), DQ Resting (PDB ID 7P7K, EMD-13238), NADH+FMN Open (PDB ID 7P7L, EMD-13239), Piericidin A+NADH Open (PDB ID 7P7M, EMD-13240), Turnover_pH6 Closed (PDB ID 7P63, EMD-13216), Turnover_pH6 Open (PDB ID 7P64, EMD-

13217), Turnover_pH6 Resting (PDB ID 7P69, EMD-13222), Turnover_pH8 Closed (PDB ID 7Z80, EMD-14540), Turnover_pH8 Open (PDB ID 7Z83, EMD-14541), Turnover_pH8 Open-ready (PDB ID 7Z84, EMD-14542) and Turnover_pH8 Resting (PDB ID 7ZC5, EMD-14620).

*Ec*CI LMNG datasets: Apo Open-ready (PDB ID 7Z7R, EMD-14535), Turnover_pH6 Open (PDB ID 7Z7T, EMD-14537), Turnover_pH6 Open-ready (PDB ID 7Z7V, EMD-14538), Turnover_pH6 Closed (PDB ID 7Z7S, EMD-14536) and Turnover_pH6 Resting (PDB ID 7ZCI, EMD-14632).

*Oa*CI datasets: pH 5.5 open (PDB ID 7ZDJ, EMD-14651), pH 5.5 closed (PDB ID 7ZDM, EMD-14658), pH 7.4 open (PDB ID 7ZD6, EMD-14637), pH 7.4 closed (PDB ID 7ZDH, EMD-14648), pH 9 open (PDB ID 7ZDP, EMD-14664) and pH 9 closed (PDB ID 7ZEB, EMD-14688).

Competing interest declaration: The authors declare no competing financial interests.

Corresponding author: LAS, sazanov@ist.ac.at

Supplementary Information

Supplementary Discussion

Supplementary Figures S1-S13

Supplementary Tables S1-S8

Supplementary Videos S1-S3

Main Figures legends

Fig. 1. The overall structure of the *E. coli* complex I in different states.

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

Fig. 2. Conformational changes induced upon *Ec*CI closing.

A) Global conformational changes upon open (grey) to closed (coloured) state transition. Key areas with changing conformation are indicated. NuoCD LHL element radically changes conformation, guiding NuoA loop to order and close the Q cavity. This process is assisted by the re-arrangement of NuoJ TM3-4 loop (note different W87 positions) and the rotation of $_J$TM3 and $_H$TM4 (note different $_J$F67 and $_H$Y156 positions). **B)** Conformational changes inside the Q cavity upon open to closed transition. NuoCD loop adopts retracted conformation, allowing quinone (magenta sticks) to bind in the Q_d site. Here quinone would clash with the extended NuoCD loop (grey), so the retraction is necessary. NuoH and NuoA loops get ordered, enclosing the Q cavity. **C)** Comparison of key NuoCD and NuoH loops in different

states. In resting state NuoCD loop is disordered and NuoH is ordered in the "up" conformation resembling, but distinct from that in open-ready state. In open-ready state both loops are ordered such that conserved $_{H}$E220 can compensate for $_{C}$H224 charge, allowing key $_{CD}$H224 and $_{CD}$H228 to interact strongly, stabilizing extended conformation of NuoCD loop. In open state NuoCD loop keeps this conformation whilst NuoH loop is disordered, facilitating quinone movement. In closed state NuoH loop changes its conformation to "down" such that $_{H}$E220 flips to form a salt bridge with conserved $_{CD}$R407. This helps $_{CD}$H224 and $_{CD}$H228 to separate, allowing retracted conformation of NuoCD loop and quinone binding in Q_d site.

Fig. 3. **Proton translocation pathways.**

A) Membrane arm contains the central axis of charged residues, essential for the proton transfer and the coupling. Structure of the LMNG turnover open-ready state is shown colored by subunit, with essential residues shown as sticks. Key ALS residues are also identified by their TM helix. Experimentally observed waters are shown as red spheres (waters beyond 5 Å from essential residues are omitted for clarity). Putative proton pathways through Grotthus-competent residues (shown as lines unless key residue) and waters are shown as black dashes. The activity of mutant variants of *Ec*CI is shown as NADH:DQ oxidoreduction in % of WT activity (Supplementary Table S8). **B)** Left: In all open, open-ready and resting states $_{J}$TM3 I63-A62 residues impose a hydrophobic block between $_{K}$E36 and $_{A}$D79. Right: In the closed state $_{J}$TM3 rotates anticlockwise, which removes the hydrophobic block and allows waters to come in and connect $_{K}$E36 and $_{A}$D79. $_{H}$Y156 also comes in helping to establish robust connection further in the E-channel. Structures are of LMNG_Tunover_pH6 open-ready (left) and closed (right) states with experimentally observed waters (red spheres).

Fig. 4. A "domino effect" coupling mechanism of complex I.

An overview of the proposed complex I mechanism. Individual steps involve conformational changes around Q cavity / E-channel and electrostatic interactions in antiporters NuoL/M/N, as described in the text. Complex I cycles between the open state (Steps 1, 4 and 5), where the Q cavity is widened and opened both to the lipid bilayer (Q) and to the cytosol (W), and the closed state (Steps 2 and 3), where the Q cavity is enclosed and tightly engulfs the bound quinone. NADH oxidation and

electron transfer in PA in Step 2 are fast and not rate-limiting. Charged quinone intermediate is indicated by the red headgroup, and quinol by the grey-filled headgroup. The charged residues on the MA central axis are indicated in blue for lysines and in red for glutamates/aspartates. For clarity, the protonated forms are shown with a + sign and un-protonated are empty (although the actual charge would be +/0 for lysine and 0/- for glutamate/aspartate). The key helices in antiporters are indicated by their numbers. In the open state the water wire between the Q cavity and the central axis in the E-channel is broken at ⱼTM3 (indicated as J3). The connection is established in the closed state due to ⱼTM3 rotation. Black arrows indicate proton transfer, including re-distribution along the central axis. Access from the cytosol happens only via NuoL/M and the exit into the periplasm only via NuoL. Electrostatic interactions, resulting in the ejection of four protons into periplasm in Step 5, are indicated as red dashes in Step 4.

Extended Data Legends

ED Fig. 1. Cryo-EM density examples. A) Composite LMNG_Turnover dataset open-ready state map combined from focus-refined PA and MA maps filtered to local resolution. **B)** Density examples of various regions of *Ec*CI, including helices and β-strands from PA and MA. **C)** Densities for the key NuoCD loop. **D)** Density for ligands, including waters. **E)** Density for the NuoJ TM3 helix in open and closed states of LMNG_Turnover dataset. **F)** Density for the key NuoA loop in open and closed states. **G)** Density for the key NuoH loop in resting ("up" conformation), Open-ready and closed ("down" conformation) states. **H)** Density of the Ca^{2+} binding site. **I-K)** Densities in the Fe-S clusters' environment.

ED Fig. 2. Features of the Peripheral Arm structure. A) Novel elements stabilizing the PA. The NuoG insertion loop (magenta) interacts with NuoG and the rest of the complex, increasing the interaction surface area. Ca^{2+} binds at the interface of the insertion loop and core NuoG structure. Loop-helix-loop connecting element of NuoCD subunit ($_{CD}$LHL, yellow) is located on the surface and interacts with the NuoG insertion loop. **B)** Unique C-terminal extensions in *Ec*CI increase surface interacting area and stabilize the minimal CI version. **C)** Position of Fe-S clusters along the peripheral arm. Electrons are transferred as a hydride from NADH to FMN

then one by one via eight Fe-S clusters to quinone. The edge-to-edge distance (Å) between the clusters is indicated. N1a and N7 are off-path clusters. The density of the water molecules within 10 Å from Fe-S clusters is shown in blue, and all experimentally identified waters are shown as red spheres. **D)** Multiple salt bridges stabilize the Q cavity and PA-MA interface in open and closed states. NuoI helix H1 is surrounded on all sides by tightly bound lipid molecules, which strengthen its binding to the rest of the complex. This interaction is likely essential for the overall stability of the Q cavity. **E, F)** The conformation of the backbone between $_{F}$D92 and $_{F}$P96 is unchanged in *Ec*CI regardless of NADH presence. Red arrows point to the backbone oxygen atom, which was suggested to change conformation (peptide bond flip) in *Aquifex aeolicus* studies[52]. **G)** Comparison of NuoF subunits from different CI species. *Ec*CI contains unique $_{F}$R320, which points into the active site and interacts with NADH. **H)** At the PA-MA interface, in the closed state the Q cavity is sealed by the NuoA loop (in orange), stabilised by the indicated conserved salt bridges.

ED Fig. 3. A-D) Cryo-EM density of the FMN binding site at different conditions in the presence of NADH. NuoFE subunits are highlighted with a dashed circle, FMN and NADH are indicated by the orange and yellow arrows, respectively. The concentration of CI applied to EM grids was ~10 mg/ml in condition A and ~0.2 mg/ml in conditions B-D. **A)** When the concentration of the holoenzyme is high (above the K_d for FMN dissociation[53]) NuoFE, FMN and NADH all have clear densities. **B)** The same is true when the protein concentration is low but a constant electron flow from FMN to DQ occurs during turnover. **C)** When the protein concentration is low and the complex is reduced without electron acceptor present, NuoFE subunits get disordered and most of FMN completely dissociates from the active site. **D)** However, FMN remains bound when external excess of FMN is added, even though NuoEF subunits still get disordered. Inserts show zoom-in into FMN/NADH (A, B) and FMN (C, D) density. Taken together, this data provides additional confirmation of the true turnover condition in B. **E)** Nomenclature of core subunits of complex I in some reference species. [1]The traditional nomenclature for Fe-S clusters (Nx, derived from initially described electron paramagnetic resonance (EPR) signatures[54], as well as the nomenclature proposed[55] on the basis of re-assignment of EPR signals to structurally observed clusters, is shown. In the new nomenclature, clusters are named according to their nuclearity (2Fe or 4Fe), their

subunit location (using bovine nomenclature) and when necessary, as ligated by four Cys (C) or three Cys and one His (H). [2]Cluster N7 is present only in some bacteria (for example, *E. coli* and *T. thermophilus*). [3]Subunits NuoC and NuoD are fused in *E. coli* and some other bacteria. [4]Number of transmembrane helices.

ED Fig. 4. A) Lipids' binding sites. Modelled molecules of phosphatidylethanolamine (PE, cyan) and eicosane (LFA, green). PDB ligand LFA was used when lipid headgroup was not resolved. Most lipids bind in hydrophobic crevices between subunits. The approximate boundaries of the lipid bilayer are indicated by the blue lines. **B)** Membrane-exposed hydrophobic belt. Top: the surface electrostatic potential of *Ec*CI. Middle: *Ec*CI coloured by the hydrophobicity of residues (white – hydrophobic, red - polar). Bottom: 10 Å low-pass filtered densities of *Ec*CI. Lipid-detergent belt density is shown in grey, DDM/LMNG_Turnover_pH6 Open in light cyan and DDM Resting in dark cyan. All datasets in DDM/LMNG (except Apo) or LMNG with external lipids contain CI in an expanded lipid-detergent belt, while the datasets in DDM only contain CI in a shrunk lipid-detergent belt. **C)** Global conformational changes upon open (grey) to closed (coloured) state transition viewed from the cytosol. This process is assisted by the rotation of $_{J}$TM3 and $_{H}$TM4 helices as well as the tilt of the entire NuoH subunit. **D)** Key NuoH TM5-6 loop in different conformations. Structures in different states (indicated by NuoH colour as labelled) from LMNG_Turnover_pH6 dataset are aligned by subunit NuoH. Conserved E220, which forms different interactions in different states, is shown as sticks. Also shown are F212 and Y225, which indicate the borders of the variable region of the loop. NuoCD subunit is in grey. **E, F)** NADH:DQ oxidoreduction activity assays. Results are represented in μmol NADH min^{-1} mg^{-1} protein, as the mean ± SEM with values from three individual measurements shown as circles. **E)** *Ec*CI. When present, lipids were added as 0.25 mg/ml ETL, and piericidin A (pA) inhibitor was added to 30 μmol. **F)** *Oa*CI. Assays were performed in the presence of 0.25 mg/mL DOPC:cardiolipin (4:1) lipid mixture. Further details provided in Methods. **G)** The architecture of ALS subunits. The N- and C-terminal 5TM repeats with inverted symmetry are coloured. TM helices are numbered, with key residues indicated by circles in blue for lysines and in red for glutamates. Beta-hairpin (β–h) and C-terminal amphipathic helix (CH), forming contacts between subunits, are also indicated.

ED Fig. 5. Quinone-binding site comparisons. A) Comparison of key NuoCD loop in different CIs. Ovine CI (*Oa*CI) is from PDBs 6ZKC for closed and 6ZKE for open state. *Te*NDH (*Thermosynechococcus elongatus* NDH complex) is from PDB 6NBY. Key conserved histidines are shown as sticks. **B)** Q binding sites. Left: in the resting state Q binds at the entrance to the Q cavity, consistent with the mammalian Q_s binding site. Middle: in the open state Q binds in the Q_m site, in between Q_d and Q_s sites (Q_d and Q_s quinones from the aligned structures are shown for comparison). Right: in the closed state DQ binds deep inside of the cavity, consistent with the mammalian Q_d binding site, and also in the Q_s site (LMNG datasets). Key residues interacting with quinone headgroup in each site are indicated. Quinone molecules from the aligned *Oa*CI structures are shown as grey sticks. Q_d site is narrow with a tight Q coordination, while Q_s is looser, with some variability in the mode of binding. **C)** Extended NuoCD loop and side-chains of NuoB Helix3 block access to the Q_d binding site in *Ec*CI open state (left) and in *Te*NDH (PDB ID 6KHJ) (right), with plastoquinone (PQ) bound in the same site as Q_m in *E. coli*. **D)** Q hydrophobic tail seals the Q entrance. Top: *Oa*CI, DQ bound in the Q_s site (PDB 6ZKE) is depicted as magenta spheres and protein atoms within 8 Å as transparent gray spheres. Bottom: *Ec*CI, model of UQ8 fitted into Q_d site of the closed state structure is depicted as magenta spheres and protein atoms within 8 Å as transparent grey spheres, except for $_H$M64 and $_H$M67 (yellow), framing the entry. **E)** Q cavity in mammalian open complex I (PDB 6ZKE) is exposed to the matrix via W site, consistent with *Ec*CI. **F)** In the open-ready state of *Ec*CI, although NuoA loop is partly ordered, the Q cavity is still exposed to the matrix via W site.

ED Fig. 6. Density and the environment of bound native quinone, externally added DQ and inhibitor piericidin A. The dataset and state of the enzyme are indicated. Cryo-EM density is carved within 2 Å of the model of the ligand. In the open state (**A, C, D, E and G**) the Q_m site is occupied either by the native quinone (**C and E**), or DQ (**D**) or piericidin A (**G**). The headgroup interacts mainly with $_C$Q328 and $_B$V85-L86 residues in this position. In the DDM resting state (**F**) native quinone binds in the Q_s site, stacking against $_H$F238 and $_B$W55. In the DDM/LMNG_Turnover_pH6 closed state (**B**) DQ binds in the Q_d site, interacting with the key $_{CD}$Y277 and $_{CD}$H228. In the DDM/LMNG_Turnover_pH8 (**I**) and LMNG_Turnover_pH6 (**H**) closed states the density for both Q_d- and Q_s-bound DQ is visible.

ED Fig. 7. Waters and proton translocation pathways. A) Cryo-EM densities for the experimental waters in the MA-focus-refined maps of *Ec*CI LMNG_Turnover_pH6 open-ready (top) and closed (bottom) states. To allow clear visualization, the density is carved around modelled waters (red spheres) and is shown in light blue. The model is coloured by subunit as in Fig. 1a. Key residues from the central hydrophilic axis of *Ec*CI are shown as sticks. **B)** A putative proton transfer pathway between the E-channel and the key $_{C}$H228/$_{C}$D329 residues, likely proton donors for quinone. Key protonatable residues, experimentally resolved waters and quinones from LMNG_Turnover_pH6 closed state are shown. Potential H-bonds are indicated by black dashes. **C)** Detailed analysis of cryo-EM density reveals charge of Glu and Asp residues in MA. Carboxyl side-chain densities of some key residues are absent (circled) in the closed state, suggesting their negative charge. In contrast, the same residues in the open state preserve densities suggesting their neutral charge. **D)** Comparison of NuoL TM8 helices from different CI species. Structures were aligned on *Ec*CI NuoL subunit. Key $_{L}$H254 residue and $_{L}$S150 with which it can interact are shown as sticks. Due to flexibility of TMH8 key histidine can be preferentially linked either to key TM12 residue as in *Ec*CI, *Tt*CI and *Yl*CI, or to key TM7 residue and the rest of the central axis as in *Oa*CI, *Te*NDH (PDB 6KHJ) and *Ai*CI (*Arabidopsis italiana* mitochondrial CI, PDB 7AR8). **E)** Comparison of NuoM TM8 helices from different CI species. In *Ec*CI $_{M}$TM8 is flexible and adopts different conformations. It is "linked" (green) in DDM/LMNG datasets PieA, Apo, Turnover in open states, and in resting states in Apo, Turnover and DDM_NADH. In DDM/LMNG datasets NADH+FMN and DQ in the open states, and resting states in DQ and DDM_Apo it is "flipped" (grey). However, both of these conformations are consistent with other CI structures as shown.

ED Fig. 8. Characterization of mutations in NuoM and NuoN subunits of *E. coli* complex I. A) Growth curves of the cultures grown aerobically at 37^{o}C in M9+malate minimal media [56]. In these conditions in mutants with impaired complex I activity the lag phase before entering the exponential growth phase is extended in comparison to the wild type *E. coli*. The extent of the lag is roughly proportional to the degree of complex I impairment. **B)** Activities of the inverted membrane vesicles of WT and mutant *E. coli* membranes. The activities were measured with deamino-NADH (dNADH), which is used exclusively by complex I in *E. coli*. dNADH:FeCy activity

involves only the peripheral FMN site and so it reflects the assembly and the overall content of complex I in the membranes, which is similar to WT in all mutants except for subunit deletion strains. WT activity was 1.33 μmol dNADH min^{-1} mg^{-1} total membrane protein. dNADH:O_2 activity reflects the activity of the entire respiratory chain, with complex I using native *E. coli* quinone. WT activity was 0.62 μmol dNADH min^{-1} mg^{-1} total membrane protein. dNADH:DQ reflects activity of complex I using decyl-ubiquinone, with complex IV inhibited. WT activity was 0.77 μmol dNADH min^{-1} mg^{-1} total membrane protein. Results are represented in % of WT activity, as the mean ± SEM. **C)** H^+ translocation activities of *E. coli* inverted membrane vesicles, measured by the quenching of ACMA fluorescence at room temperature with an excitation wavelength of 434 nm and an emission wavelength of 477 nm. The buffer contained 2 μM ACMA, 50 mM Bis-Tris at pH 6.0, 2 mM $CaCl_2$, 10mM $MgCl_2$, 10 μM valinomycin and 50mM KCl. The addition of 0.1 mM dNADH or 2 μM of uncoupler FCCP (which dissipates ΔpH) is indicated.

ED Fig. 9. Characterization of mutations in NuoH subunit of *E. coli* complex I. **A)** Growth curves of the cultures grown aerobically at 37°C in M9+malate minimal media [56]. In these conditions in mutants with impaired complex I activity the lag phase before entering the exponential growth phase is extended in comparison to the wild type *E. coli*. The extent of the lag is roughly proportional to the degree of complex I impairment. **B)** Activities of the inverted membrane vesicles of WT and mutant *E. coli* membranes, measured with deamino-NADH (dNADH) as described in ED Fig. 8b. **C)** H^+ translocation activities of *E. coli* inverted membrane vesicles, measured by the quenching of ACMA fluorescence at room temperature with an excitation wavelength of 410 nm and an emission wavelength of 480 nm. The buffer contained 1.6 μM ACMA, 20mM Bis-Tris at pH 6.0, 2mM $CaCl_2$, 20 mM KCN, 200 μM DQ, 1 μM valinomycin and 120 mM KCl. The addition of 0.1 mM dNADH or 1 μM of uncoupler CCCP (which dissipates ΔpH) is indicated. **D)** Inhibition of the NADH:DQ activity of the purified *Ec*CI by rotenone. Mutant $_H$M67A was purified similarly to WT [26]. Non-inhibited activity of the purified mutant enzyme is somewhat higher than in membrane vesicles (C), indicating that DQ access is facilitated in detergent.

ED Table 1. Cryo-EM datasets. Summary of collected datasets showing different states of *E. coli* and ovine complex I.

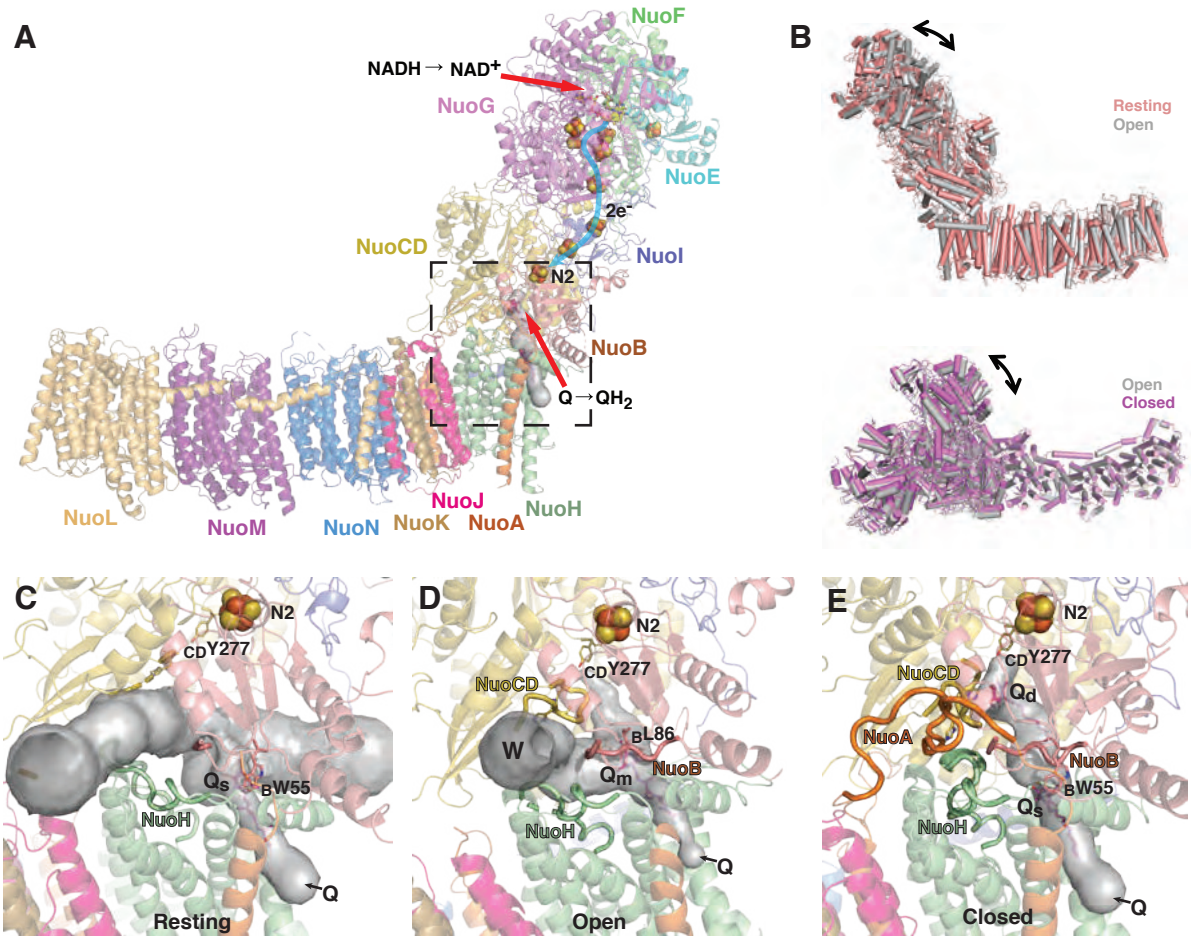
A
NADH → NAD+
NuoF
NuoG
NuoE
2e-
NuoCD
NuoI
N2
NuoB
Q → QH2
NuoL
NuoM
NuoN
NuoK
NuoA
NuoJ
NuoH
B
Resting
Open
Open
Closed
C
N2
CDY277
Qs
BW55
NuoH
Q
Resting
D
N2
CDY277
NuoCD
W
BL86
Qm
NuoB
NuoH
Q
Open
E
N2
CDY277
NuoCD
Qd
NuoA
NuoB
BW55
Qs
NuoH
Q
Closed

Figure 1

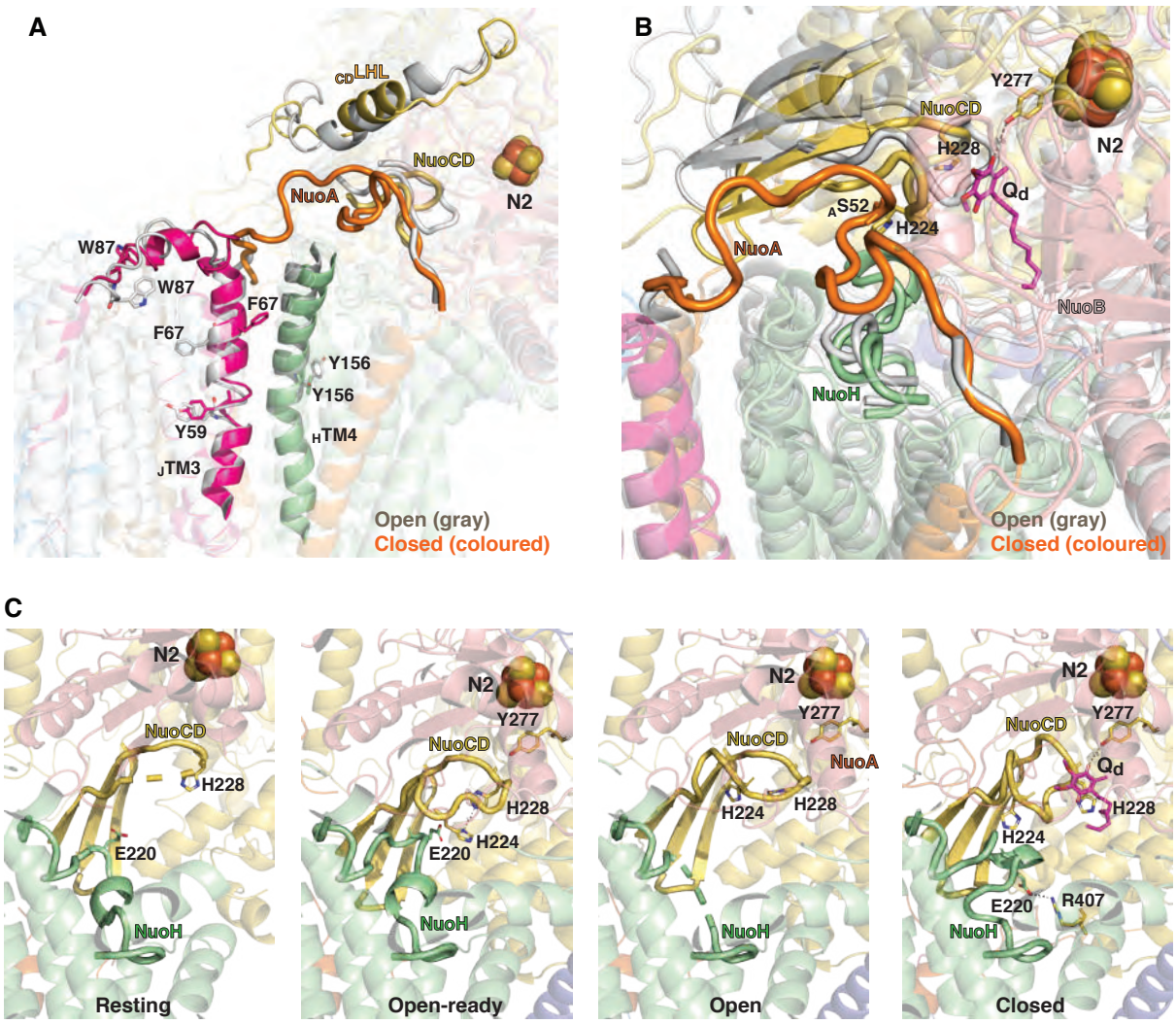
A
CDLHL
NuoCD
NuoA
N2
W87
W87
F67
F67
Y156
Y156
Y59
HTM4
JTM3
Open (gray)
Closed (coloured)
B
Y277
NuoCD
H228
N2
Qd
ATS52
H224
NuoA
NuoB
NuoH
Open (gray)
Closed (coloured)
C
N2
NuoCD
H228
E220
NuoH
Resting
N2
Y277
NuoCD
H228
H224
E220
NuoH
Open-ready
N2
Y277
NuoCD
NuoA
H224
H228
NuoH
Open
N2
Y277
NuoCD
Qd
H228
H224
E220
R407
NuoH
Closed

Figure 2

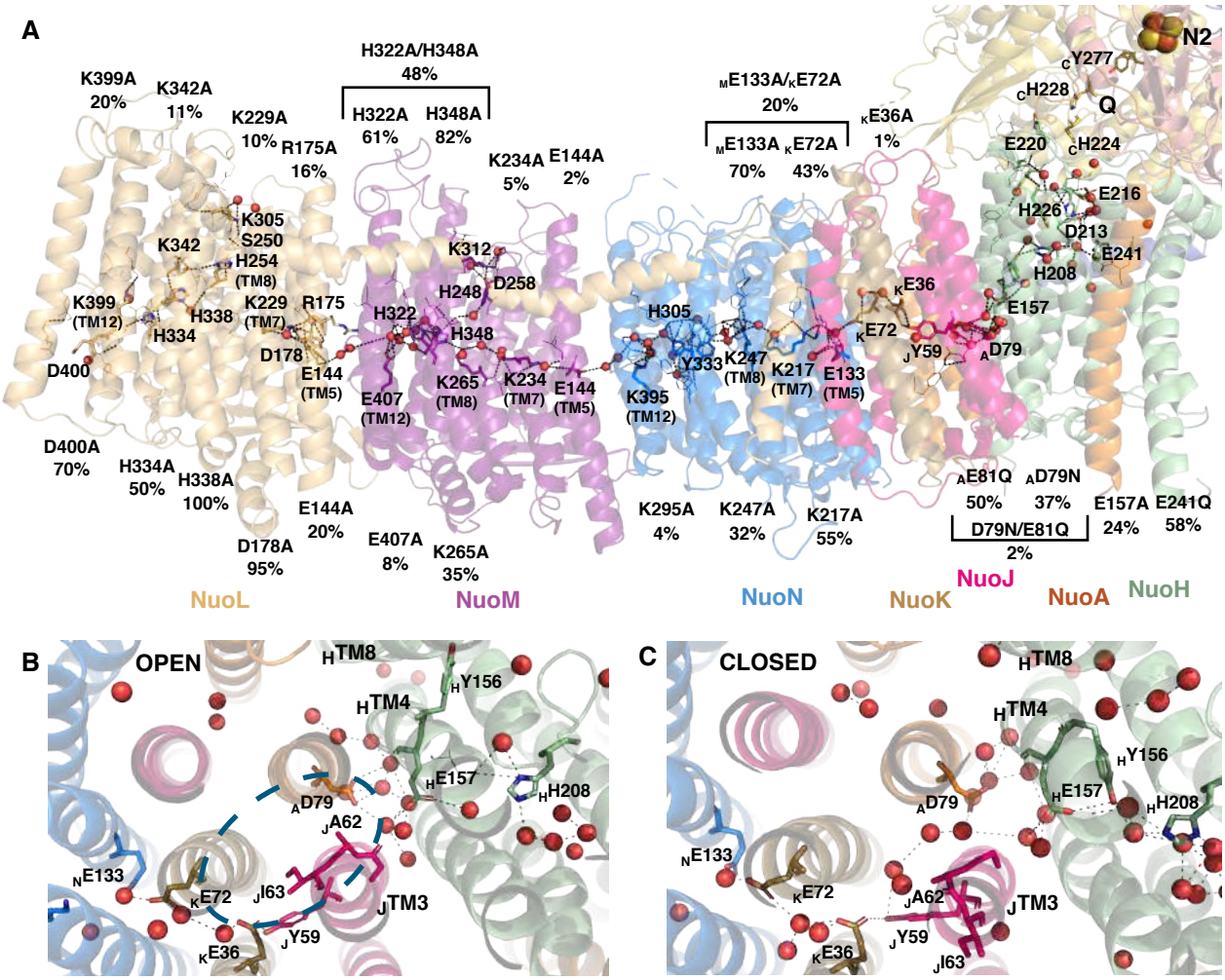
A
K399A 20%
K342A 11%
K229A 10%
R175A 16%
H322A/H348A 48%
H322A 61%
H348A 82%
K234A 5%
E144A 2%
ME133A/KE72A 20%
ME133A 70%
KE72A 43%
KE36A 1%
N2
CY277
CH228
Q
E220
CH224
E216
H226
D213
E241
H208
E157
K305
S250
H254 (TM8)
K342
K399 (TM12)
H334
H338
K229 (TM7)
R175
D178
D400
E144 (TM5)
H322
K312
H248
D258
H348
E407 (TM12)
K265 (TM8)
K234 (TM7)
E144 (TM5)
H305
Y333
K247 (TM8)
K217 (TM7)
E133 (TM5)
K395 (TM12)
KE36
KE72
JY59
AD79
D400A 70%
H334A 50%
H338A 100%
D178A 95%
E144A 20%
E407A 8%
K265A 35%
K295A 4%
K247A 32%
K217A 55%
AE81Q 50%
AD79N 37%
D79N/E81Q 2%
E157A 24%
E241Q 58%
NuoL
NuoM
NuoN
NuoK
NuoJ
NuoA
NuoH
B
OPEN
HTM8
HY156
HTM4
HE157
HH208
AD79
JA62
NE133
KE72
JI63
JTM3
JY59
KE36
C
CLOSED
HTM8
HTM4
HY156
HE157
HH208
AD79
NE133
KE72
JA62
JTM3
JY59
KE36
JI63

Figure 3

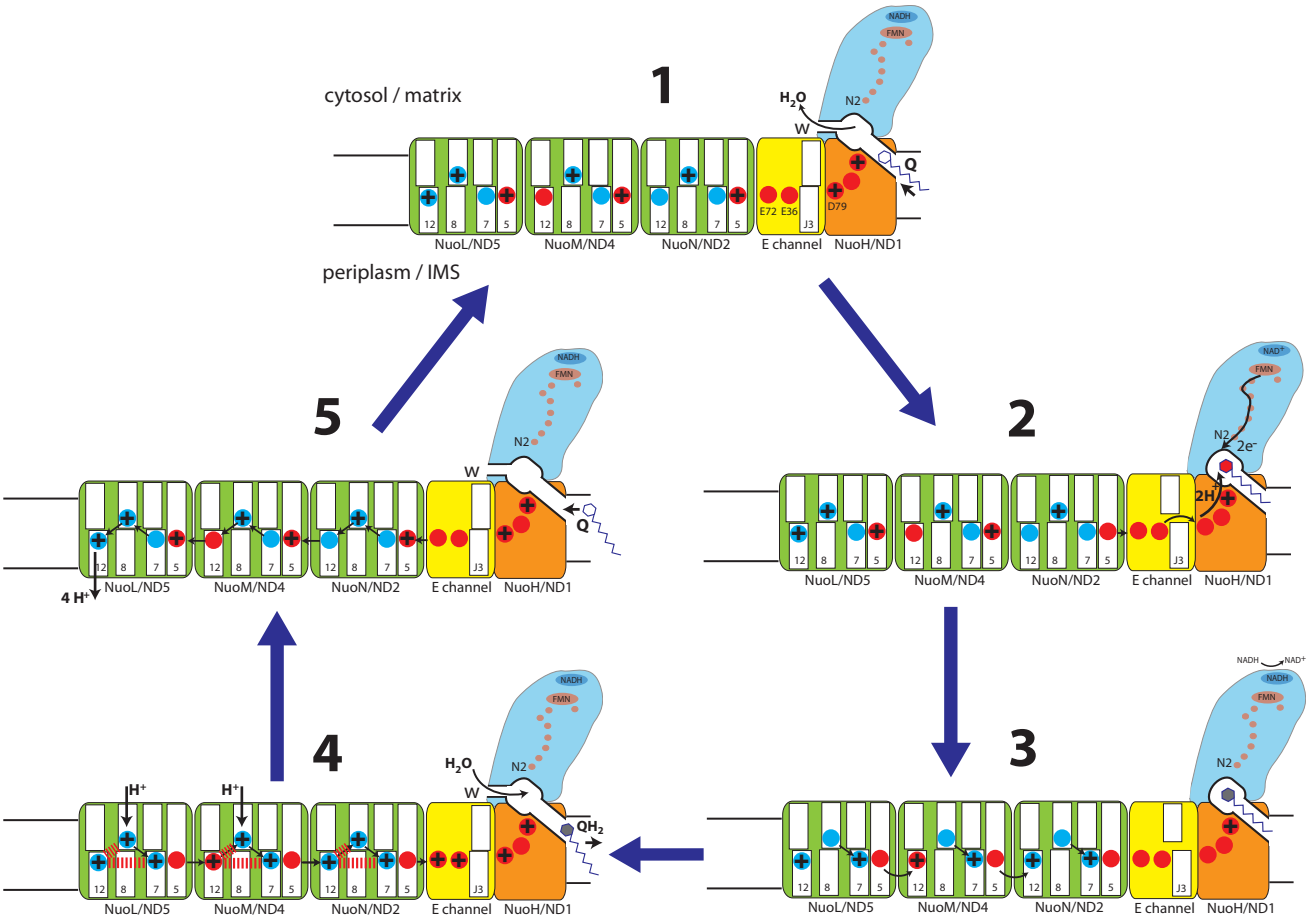
cytosol / matrix
periplasm / IMS
1
2
3
4
5
NADH
FMN
NAD+
N2
W
Q
QH2
H2O
2e-
2H+
4 H+
H+
E72 E36
D79
J3
NuoL/ND5
NuoM/ND4
NuoN/ND2
E channel
NuoH/ND1

Figure 4