ARTICLES

The architecture of respiratory complex I

Rouslan G. Efremov*, Rozbeh Baradaran* & Leonid A. Sazanov

Complex I is the first enzyme of the respiratory chain and has a central role in cellular energy production, coupling electron transfer between NADH and quinone to proton translocation by an unknown mechanism. Dysfunction of complex I has been implicated in many human neurodegenerative diseases. We have determined the structure of its hydrophilic domain previously. Here, we report the α-helical structure of the membrane domain of complex I from *Escherichia coli* at 3.9 Å resolution. The antiporter-like subunits NuoL/M/N each contain 14 conserved transmembrane (TM) helices. Two of them are discontinuous, as in some transporters. Unexpectedly, subunit NuoL also contains a 110-Å long amphipathic α-helix, spanning almost the entire length of the domain. Furthermore, we have determined the structure of the entire complex I from *Thermus thermophilus* at 4.5 Å resolution. The L-shaped assembly consists of the α-helical model for the membrane domain, with 63 TM helices, and the known structure of the hydrophilic domain. The architecture of the complex provides strong clues about the coupling mechanism: the conformational changes at the interface of the two main domains may drive the long amphipathic α-helix of NuoL in a piston-like motion, tilting nearby discontinuous TM helices, resulting in proton translocation.

Complex I (NADH:ubiquinone oxidoreductase) is one of the largest membrane protein assemblies known and has a central role in energy production by the mitochondrial respiratory chain, providing about 40% of the proton-motive force required for the synthesis of ATP[1–5]. Many mutations in complex I subunits have been associated with human neurodegenerative diseases[5,6]. Complex I has also been suggested to be a major source of reactive oxygen species in mitochondria, which can damage mitochondrial DNA and may be one of the causes of Parkinson's disease[7] and ageing[8]. Mitochondrial complex I consists of 45 subunits with a combined mass of about 980 kDa[9]. The prokaryotic enzyme is simpler, consisting of 14 'core' subunits conserved from bacteria to humans, with a total mass of about 550 kDa[1,2,5]. The mitochondrial and bacterial enzymes contain equivalent redox components and have a similar L-shaped structure[2,10,11]. The hydrophobic arm is embedded in the membrane and the hydrophilic peripheral arm protrudes into the mitochondrial matrix or the bacterial cytoplasm[2,5]. The high degree of sequence conservation of 'core' subunits indicates that complex I mechanism is likely to be the same throughout the species, and so the bacterial enzyme represents a 'minimal' model of human complex I.

The complete structure and the coupling mechanism of this large molecular machine are currently unknown. Our crystal structure of the hydrophilic domain (eight subunits, 280 kDa) of complex I from *T. thermophilus* established the electron transfer pathway from NADH to the primary electron acceptor flavin mononucleotide (FMN) and then through seven conserved iron-sulphur (Fe-S) clusters to the likely quinone binding site (Q-site) at the interface with the membrane domain[12,13]. The membrane spanning part of the enzyme lacks covalently bound prosthetic groups, but it must contain the proton translocating machinery. Its atomic structure is currently unknown. The three largest hydrophobic subunits of complex I, NuoL/Nqo12, NuoM/Nqo13 and NuoN/Nqo14 (*E. coli*/*T. thermophilus* nomenclature) are homologous to each other and to the Na^+ or K^+/H^+ antiporter family (Mrp)[14,15], and so are likely to participate in proton translocation. The transfer of two electrons from NADH to quinone is coupled to the translocation of four (current consensus value) protons across the membrane[1–3,5]. Although *T. thermophilus* and some other bacteria use lower potential menaquinone instead of ubiquinone, it is likely that H^+/e^- stoichiometry of the enzyme is the same across the species, as discussed previously[5]. Two possible mechanisms of coupling have been proposed: direct (redox-driven) and indirect (conformation-driven)[2,3,5,16].

Structure of the membrane domain

Crystallization of the intact complex I from *E. coli* has not yet produced high quality crystals, possibly owing to the flexibility of the connection between the two arms of the enzyme[17,18]. However, we have isolated and crystallized the membrane domain of complex I from *E. coli* (see Methods). The crystals contain the six subunits, NuoL, M, N, A, J and K (total mass 222 kDa), and lack NuoH (Supplementary Fig. 1a), which dissociates readily from the complex[19]. After extensive optimization, including post-crystallization dehydration, the crystals diffract to 3.9 Å resolution (Supplementary Table 1). Phasing using Ta_6Br_{12} cluster derivatives and density modification with multi-crystal averaging led to good quality maps showing clear electron density for transmembrane (TM) helices (Supplementary Fig. 2). An α-helical model (C_α trace) was built into this density. It consists of 55 TM helices (Fig. 1), in agreement with most predictions for the NuoL/M/N/A/J/K subunits[14,15,19–22]. The extent and the arrangement of the helices in this model are unambiguous, but the directionality and the connectivity are not established because of limited resolution. The domain is curved, if viewed both in plane and perpendicular to the membrane surface (Fig. 1 and Supplementary Fig. 3), consistent with previous electron microscopy studies[19,23]. It is about 160 Å long and about 40 Å thick, matching the average thickness of a lipid bilayer.

A pattern of 14 TM helices repeats three times in the structure, which allowed us to assign the helices to homologous subunits NuoL, M and N (Supplementary Fig. 4). Not only do all the helices align very well (Fig. 2), but also each subunit contains, in the same positions, two characteristic 'broken' helices (shown in red and orange in Fig. 1 and indicated in Fig. 2). Such discontinuous TM helices, interrupted by a loop (extended peptide) in the middle of the lipid bilayer, are thought to be important for the function of many transporters and channels[24] (see below). Four helices form a core of each subunit, surrounded by a ring of ten helices (including two discontinuous), tilted up to about 25° in the same direction relative to the centre of

Medical Research Council Mitochondrial Biology Unit, Wellcome Trust/MRC Building, Hills Road, Cambridge CB2 0XY, UK.
*These authors contributed equally to this work.

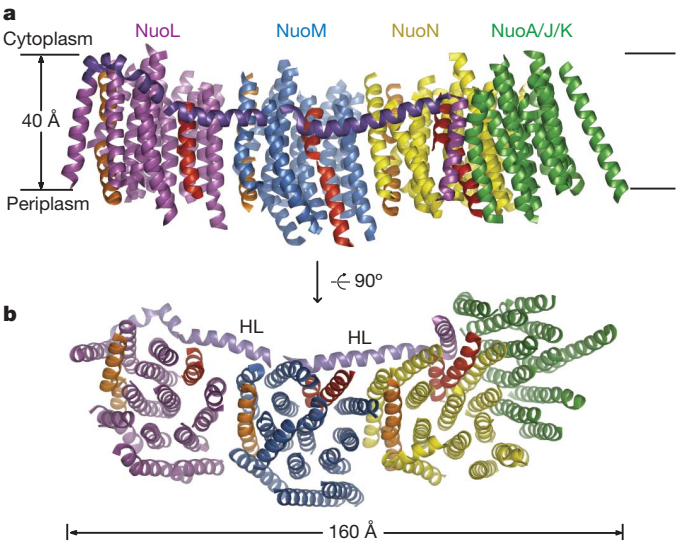

Figure 1 | The α-helical model of the membrane domain of *E. coli* complex I. **a**, Side view, in the membrane plane. **b**, View from the periplasm into the membrane. Subunit NuoL is magenta, NuoM blue, NuoN yellow and subunits NuoA/J/K green. Helix HL from NuoL is in a darker colour as indicated. The expected position of the lipid bilayer is shown. The cytoplasmic and periplasmic sides are identified in accordance with the structure of the entire complex (Fig. 3). Two discontinuous helices present in each of subunits NuoL, M and N are shown in red for those in contact with helix HL, and in orange for those farther away.

the subunit (Fig. 2c). Thus, all the antiporter-like subunits in complex I contain 14 conserved TM helices. This is in contrast to labelling studies, which suggested only 12 conserved TM helices[15,25] (probably, the introduction of the label in a sensitive position can disrupt the folding of these subunits). The assignment of the individual subunits within the domain (Fig. 1) is based on fragmentation and electron microscopy studies[26,27]. Because subunit NuoL (or subunits NuoL and M jointly) can be removed from the otherwise intact complex[26,27], NuoL is located at the distal (from the hydrophilic domain) end of the membrane domain, followed by NuoM.

Subunit NuoL contains a carboxy-terminal extension, absent in NuoM and N, with two TM helices predicted at the beginning and the end of the extension (Supplementary Fig. 4). This is consistent with the position of NuoL at the end of the domain, where there is one additional TM helix at the very tip. Thus, this helix can be identified as the fifteenth TM helix of NuoL. In the electron density (Supplementary Fig. 2b), this helix is continued by very strong α-helical density (helix HL in Fig. 1), running along almost the entire length of the domain and ending with another TM helix, which is likely to be the sixteenth and C-terminal TM helix of NuoL. There is a kink in the middle of helix HL (near subunit NuoM) and so it is modelled as discontinuous in that area. However, the electron density does connect two halves of the helix (Supplementary Figs 2 and 6), so this connection is likely to be rigid and we refer to a single helix HL. Consistently, secondary structure predictions strongly indicate (in all species examined by us) a near-continuous α-helix in this part of NuoL (residues 515–590, Supplementary Fig. 4). It is amphipathic, consistent with its position near the expected surface of lipid bilayer (Fig. 1). The number of residues (75) is just sufficient to form an α-helix of the observed length (110 Å). This is an unexpected and unusual feature of subunit NuoL, which is likely to be important for the mechanism, as discussed below. Additionally, such an arrangement anchors NuoL by its last TM helix in the vicinity of the interface with the hydrophilic domain, which may be beneficial for the stability of the complex.

Subunits NuoA/J/K must occupy the TM helices of the model unaccounted for by NuoL/M/N. There are 11 such helices, consistent with topology studies (NuoA[22] and NuoK[20] are predicted to contain three helices each and NuoJ five[21]). On the basis of mutagenesis studies it was suggested that subunits NuoA/J/K may form a multi-helix bundle involved in the coupling mechanism[28]. This is consistent with their joint position here at the interface with the hydrophilic domain.

Structure of the entire complex I

It is clearly important to know exactly how membrane domain structure fits within the entire complex. We purified intact complex I from *T. thermophilus* (Methods). The preparation shows high activity with decyl-ubiquinone (19 μmoles NADH oxidized per minute per mg of protein), which is more than 90% inhibited by rotenone or piericidin A, indicating that the complex is fully active and intact. We crystallized it in two different crystal forms ($P2_1$ and $P2_12_12_1$, Methods), containing all 15 known subunits (14 'core' and the frataxin-like subunit Nqo15, specific to *T. thermophilus*[12]; Supplementary Fig. 1b). The $P2_1$ crystals were of better quality (although pseudo-merohedrally twinned), and after extensive optimization diffracted to 4.5 Å resolution (Supplementary Table 2). The *T. thermophilus* hydrophilic domain model (PDB 3I9V, ref. 13) and the *E. coli* membrane domain model were used for molecular replacement. Similar solutions, yielding an L-shaped molecule, were obtained for both crystal forms (Supplementary Fig. 5). The electron density in the $P2_1$ crystal form was of sufficient quality to allow us to build an α-helical model for subunit Nqo8/NuoH (Supplementary Fig. 6), which was lacking in the *E. coli* crystals. It contains eight TM helices (one strongly bent), consistent with predictions and topology studies for most species[29]. In *T. thermophilus*, sequence analysis indicates one additional TM helix at the C terminus, but there is no clear evidence for it in the current

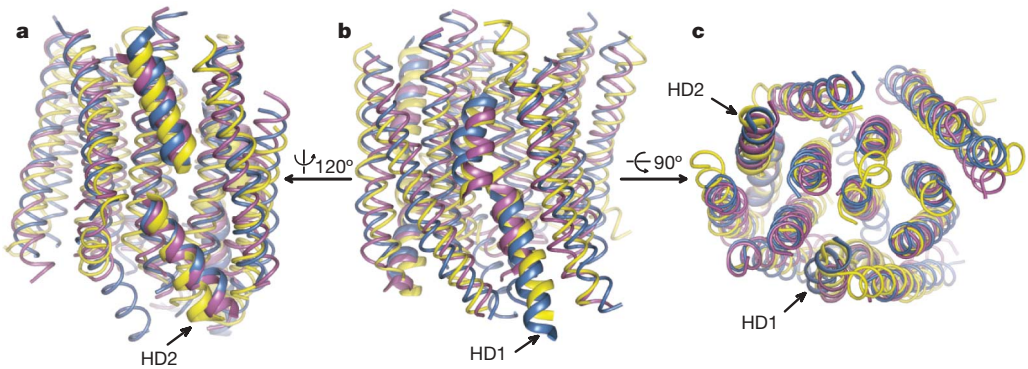

Figure 2 | Structural alignment of the antiporter-like subunits NuoL, NuoM and NuoN of *E. coli* complex I. **a**, **b**, Side views in the membrane plane, with the cytoplasmic side top. **c**, View from the cytoplasm into the membrane. The discontinuous helices are highlighted and indicated as HD1 for those in contact with helix HL and as HD2 for those farther away. Lipid-exposed helix next to HD2 (seen to the left of it in **a**) resembles the discontinuous one, but its main part approaches normal TM length and the nearby short helix is nearly parallel to the bilayer. Helices HD1/HD2 do not form an interacting pair, as is usual[24], but are separated by a layer of helices (**c**). Helices from the C-terminal extension of subunit NuoL, absent in NuoM and NuoN, are omitted. The colour scheme is the same as in Fig. 1.

electron density. The individual subunits of *T. thermophilus* membrane domain were refined as rigid bodies. No further rebuilding was done in the absence of independent phase information. Thus, the model reflects the overall arrangement of membrane subunits. Compared to the *E. coli* membrane domain, the *T. thermophilus* membrane domain is less curved when viewed both in the lipid plane and perpendicular to it (Supplementary Fig. 6).

Thus, the current model of complex I from *T. thermophilus* (total mass 523 kDa) consists of an atomic model for the hydrophilic domain and an α-helical model for the membrane domain (Fig. 3). In total, the membrane domain contains 63 TM helices, one of the largest hydrophobic protein complexes yet described. As expected from electron microscopy studies[10,11], the complex is L-shaped, with the membrane arm about 180-Å long and the peripheral arm extending about 130 Å over the lipid bilayer. Comparison to recent three dimensional electron microscopy reconstructions[10] shows that only the mirror image of the electron microscopy density fits well with the crystal structure (Supplementary Fig. 7), indicating that the hand of these electron microscopy reconstructions needs to be reversed.

The overall orientation of the hydrophilic domain relative to the membrane domain is unexpected. Helix H1 from subunit Nqo6, rather than extending over the surface of the membrane domain as suggested previously[5], points sideways at about 90° (Fig. 3). The contacts between the two arms of the complex are mediated mostly through subunit Nqo8 of the membrane domain and subunits Nqo4/6 of the hydrophilic domain. One of subunits in the Nqo7, 10, 11 (NuoA/J/K) bundle also contributes by contacting the amino-terminal part of Nqo4. Consistently, Nqo7 can be cross-linked to Nqo4 and Nqo6 (ref. 30). Subunit Nqo8 contains very large cytoplasmic loops with many conserved charged residues, which may form part of the interface with the hydrophilic domain. These loops are not resolved and so the contacts between the two domains do not seem to be extensive in the current model.

Quinone-binding site

The last in the chain of Fe-S clusters in the hydrophilic domain is thought to be the donor of electrons to the quinone[5,12]. This high-potential cluster N2 is near the cavity between subunits Nqo4 and Nqo6, which forms part of a larger cavity at the interface of these subunits with hydrophobic subunits Nqo8 and the Nqo7, 10, 11 bundle (Fig. 3d). The N-terminal β-sheet of Nqo4 (residues 26–60) forms part of this cavity, consistent with labelling of the corresponding region of bovine 49 kDa subunit with azidoquinazoline[31]. Subunit Nqo8 was also labelled with quinone-like inhibitors[32] and is thought to form a part of quinone-binding pocket along with subunits Nqo4 and Nqo6 (refs 2, 3, 5, 12). Thus, cumulative evidence indicates that this cavity at the interface of subunits Nqo4, Nqo6, Nqo8 and the Nqo7, 10, 11 bundle provides a quinone-binding site. It is a large cavity, up to about 30 Å across, consistent with it providing a large common binding domain, with partially overlapping sites, for various quinone-like inhibitors[3,32]. A quinone head group, bound in the cavity, can easily reach to within about 14 Å (required for physiological electron transfer[33]) or to about 12 Å (suggested by N2-semiquinone interactions[34]) from the edge of cluster N2. This cluster is about 20–25 Å from the expected surface of the lipid bilayer (Fig. 3), thus some movement of the quinone out of the membrane is likely to be required to approach N2, and its extent will depend on the exact location of the bound quinone in the membrane domain. There are several possible sites for accommodating the isoprenoid tail, formed either by subunit Nqo8 alone or at its interface with the Nqo7, 10, 11 bundle. In the latter case quinone will need to move about 20 Å to approach cluster N2, whereas in the former the movement does not need to exceed about 10 Å, which means that the tail of

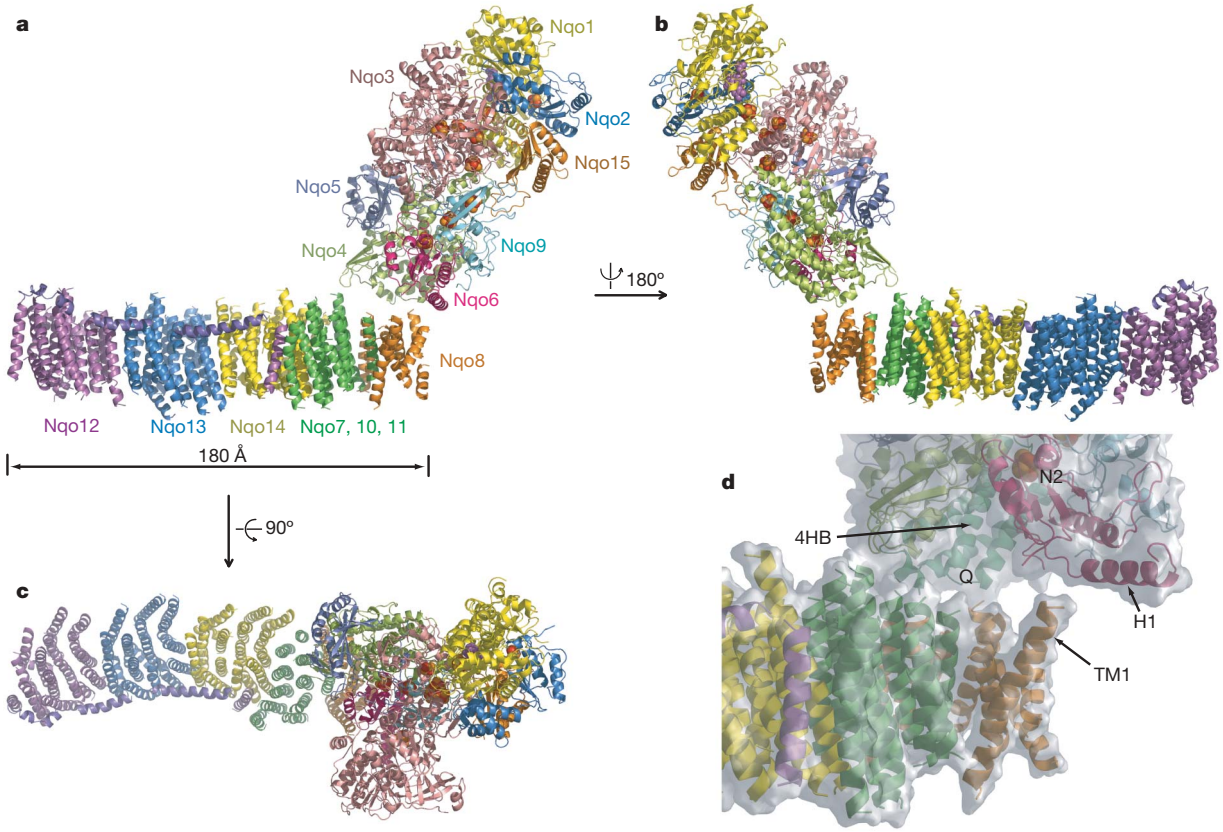

Figure 3 | The structure of the entire complex I from *T. thermophilus*. The structure consists of the atomic model for the hydrophilic domain, determined previously[12,13] (PDB 3I9V), and the α-helical model for the membrane domain. **a**, **b**, Side views, in the membrane plane. **c**, Top view, from the cytoplasm into the membrane. **d**, The likely quinone-binding cavity (Q) at the interface of the two main domains. Helix H1 from subunit Nqo6, the 4-helix bundle from subunit Nqo4 (4HB, in darker colour), helix TM1 from subunit Nqo8 and cluster N2 are indicated. Fe-S clusters are shown as red and yellow spheres, and FMN as magenta spheres. Each subunit is coloured differently and indicated.

the quinone can still reside mostly within the membrane, as expected for a highly hydrophobic moiety. This is in contrast to suggestions that cluster N2 is 30–60 Å away from the membrane and the quinone moves to it along a hydrophobic "ramp"[10]. Additional quinone-binding sites in subunits NuoL[35] and NuoM[36] have been proposed. There is no clear evidence for them in the structure, although it cannot be excluded that the quinone binds at the surface of these subunits.

Functional implications

The overall architecture of complex I has strong implications for the coupling mechanism. All antiporter-like subunits, including Nqo14, are separated from the quinone-binding site and the interface with the hydrophilic domain by a belt of two layers of α-helices in the Nuo7, 10, 11 bundle. This arrangement precludes a direct link between electron transfer and proton translocation via the antiporter-like subunits. Many cross-linking studies have indicated conformational changes in complex I upon reduction[23,37–39]. In our 'combined' mechanism proposed previously, in the course of the catalytic cycle, one proton is translocated by 'direct' coupling via cluster N2 and its unusual coordinating tandem cysteines and the rest are driven by conformational changes involving the four-helix bundle of Nqo4 and helix H1 of Nqo6, which shift upon reduction[13]. We can see now that the four-helix bundle is in direct contact with the Nqo8 subunit, and that helix H1 is also likely to contact the longest helix in Nqo8 (TM1, Fig. 3d). Thus, these Nqo4/6 helices are well positioned to transmit conformational changes to Nqo8 and the Nqo7, 10, 11 bundle (cross-links between Nqo7 and Nqo10 change upon reduction[37]). The six outermost helices of Nqo8 (including TM1) are highly tilted (by up to about 40°) relative to the lipid bilayer normal, which may reflect the role of Nqo8 in conformational coupling. Furthermore, conformational changes in Nqo8 and the Nqo7, 10, 11 bundle may be aided by quinone movement to and from cluster N2 on binding, reduction and release. This would agree with the suggestion that most of the energy from NADH is released only on delivery of its second electron to quinone[40].

The unusual, conserved, long amphipathic helix HL of NuoL/Nqo12 (Figs 1 and 3) is well positioned to act as a mechanical link transmitting conformational changes from near the Q-site to the antiporter-like subunits. Helix HL is in direct contact with one of the discontinuous helices in each of the three antiporter-like subunits (red in Fig. 1 and HD1 in Fig. 2). Such helices are thought to be involved in ion translocation by introducing a charge (and some flexibility) into the core of the membrane[24]. The other three discontinuous helices (orange in Fig. 1 and HD2 in Fig. 2) may not be involved directly in the translocation as they are farther away from helix HL. A possible mechanism (Fig. 4) is that conformational changes in Nqo8 and the Nqo7, 10, 11 bundle affect the nearby C-terminal helix of Nqo12 (Fig. 3) and a tilt of this helix leads to a piston-like motion of helix HL along the membrane domain. This movement synchronously tilts the three nearby discontinuous helices, resulting in proton translocation. From charged residues predicted to be located deep within the membrane, one glutamate and one lysine are conserved in all three antiporter-like subunits and they are essential for energy transduction (Glu 144 and Lys 234 for *E. coli* NuoM; Supplementary Fig. 4)[25,41,42]. Glu 144 is located in TM helix 5, for which there are weak α-helical prediction scores in the middle and it may therefore be discontinuous. This glutamate may thus be located in the discontinuous helices contacting helix HL. During turnover, an ionizable residue in such a position, if exposed in the proton channel with suitable environment, could bind and release protons on the opposite sides of the membrane via conformational changes driven by helix HL. Without such a mechanical link, it would be difficult to imagine how conformational changes might propagate over the considerable length of the membrane domain.

A net result of such conformationally driven catalytic cycle would be three translocated protons (one per antiporter-like subunit). The fourth proton can be translocated via cluster N2 and its tandem

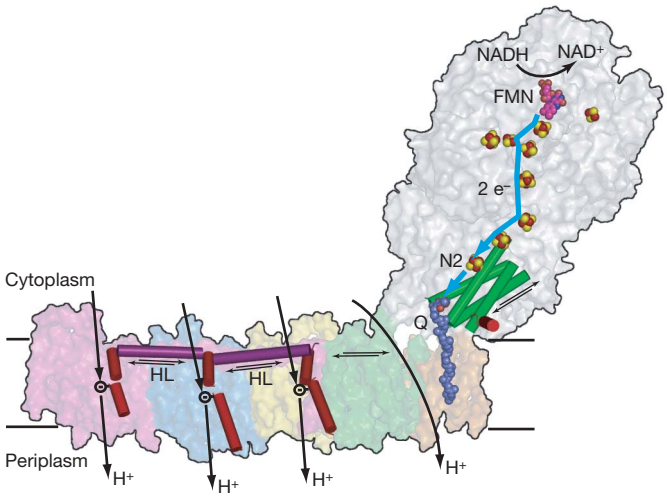

Figure 4 | Proposed model of proton translocation by complex I. NADH, via FMN (magenta), donates two electrons to the chain of Fe-S clusters (red and yellow spheres), which are passed on (blue line), via terminal cluster N2, to the quinone (dark blue, moved out of the membrane by about 10 Å). Electron transfer is coupled to conformational changes (indicated by arrows) in the hydrophilic domain, observed[13] for Nqo4 four-helix bundle (green cylinders) and Nqo6 helix H1 (red). These changes are transmitted to the amphipathic helix HL (magenta), which tilts three discontinuous helices (red) in antiporter-like subunits, changing the conformation of ionizable residue inside respective proton channels, resulting in translocation of three protons. The fourth proton is translocated at the interface of the two main domains. The hydrophilic domain surface is shown in grey, whereas the membrane domain surface is coloured as in Fig. 3.

cysteines[13], if an appropriate network of charged residues is formed in the loops between the two main domains (not resolved in the structure). Alternatively, in view of the elevated position of cluster N2 over the membrane (Figs 3 and 4), the translocation of this proton may also be conformation-driven, via subunits Nqo7, 10, 11 (ref. 28) (or Nqo8). Subunit Nqo11 also contains a conserved and essential for energy-transducing activities glutamate within predicted TM helix[28,43].

Multi-subunit Mrp antiporters are thought to represent a common ancestor of complex I and membrane-bound hydrogenases[15], and the essential Glu 144 is also conserved in these enzymes[25,44]. The antiporter subunit MrpA, hydrogenase-4 subunit HyfB and subunit NdhF from the chloroplast Ndh complex, related to NuoL, all contain, according to secondary structure predictions, a long cytoplasmic helix between their fifteenth and sixteenth TM helices. Therefore, ion-translocating membrane protein complexes related to complex I may operate using a mechanical link similar to helix HL.

The overall architecture of this large molecular machine is now clear. F-ATPase has been compared to a turbine. In a similar vein, complex I seems to resemble a steam engine, where the energy of the electron transfer is used to move a piston, which then drives, instead of wheels, a set of discontinuous helices. The full mechanistic details remain to be clarified by atomic structures of the membrane domain and the entire complex.

METHODS SUMMARY

Complex I was purified from *E. coli* strain BL21 as described previously[17]. Membrane and hydrophilic domains were separated by treatment with high concentration of Mg^{2+}, and the membrane domain was purified further on a Mono-S HR5/5 column. For crystallization, a protein solution (10–11 mg ml^{-1} in 10 mM sodium acetate pH 4.8, 50 mM NaCl and 0.05% *n*-dodecyl-β-maltoside (DDM)) was mixed first with 0.5% *n*-heptyl-β-glucoside and 2 mg ml^{-1} lipids and then, in sitting drop or microbatch plates, with 0.1 M sodium acetate pH 4.8, 0.8–1.0 M sodium or potassium formate and 9% PEG4000. The diffraction properties were improved by dehydration of crystals in 40% PEG4000 before cryo-cooling. Phasing was performed with Ta_6Br_{12} cluster derivatives. Electron density was improved by multi-crystal averaging and density modification, using three phased and one unphased non-isomorphous data sets. The

α-helical model of the membrane domain was built automatically initially and completed manually.

Intact complex I from *T. thermophilus* was purified by a series of three ion-exchange columns and a gel-filtration column. For crystallization in $P2_1$ space group, complex I (20 mg ml^{-1} in 20 mM Bis-Tris pH 6.0, 0.002% PMSF, 2 mM $CaCl_2$, 25% glycerol, 100 mM NaCl and about 1% *n*-tridecyl-β-maltoside) was mixed, in sitting drop plates, with 100 mM Bis-Tris pH 6.0, 100 mM KCl, 100 mM glutaric acid pH 6.0, 16% (w/v) PEG 4000 and 2.04 mM *n*-octyl-β-maltoside fluorinated. $P2_12_12_1$ crystals were grown under broadly similar conditions, except that protein solution was mixed with *E. coli* polar lipids at 9:1 (w/w) ratio before set-up. Phasing was performed by molecular replacement, using the previously determined structure of the hydrophilic domain of *T. thermophilus* complex I (refs 12, 13) and the model of the membrane domain of *E. coli* complex I described here.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 21 January 2009; accepted 31 March 2010.

1. Walker, J. E. The NADH:ubiquinone oxidoreductase (complex I) of respiratory chains. *Q. Rev. Biophys.* **25,** 253–324 (1992).
2. Yagi, T. & Matsuno-Yagi, A. The proton-translocating NADH-quinone oxidoreductase in the respiratory chain: the secret unlocked. *Biochemistry* **42,** 2266–2274 (2003).
3. Brandt, U. Energy converting NADH:quinone oxidoreductase (complex I). *Annu. Rev. Biochem.* **75,** 69–92 (2006).
4. Ohnishi, T. Iron-sulfur clusters/semiquinones in complex I. *Biochim. Biophys. Acta* **1364,** 186–206 (1998).
5. Sazanov, L. A. Respiratory complex I: mechanistic and structural insights provided by the crystal structure of the hydrophilic domain. *Biochemistry* **46,** 2275–2288 (2007).
6. Schapira, A. H. Human complex I defects in neurodegenerative diseases. *Biochim. Biophys. Acta* **1364,** 261–270 (1998).
7. Dawson, T. M. & Dawson, V. L. Molecular pathways of neurodegeneration in Parkinson's disease. *Science* **302,** 819–822 (2003).
8. Balaban, R. S., Nemoto, S. & Finkel, T. Mitochondria, oxidants, and aging. *Cell* **120,** 483–495 (2005).
9. Carroll, J. *et al.* Bovine complex I is a complex of 45 different subunits. *J. Biol. Chem.* **281,** 32724–32727 (2006).
10. Clason, T. *et al.* The structure of eukaryotic and prokaryotic complex I. *J. Struct. Biol.* **169,** 81–88 (2010).
11. Morgan, D. J. & Sazanov, L. A. Three-dimensional structure of respiratory complex I from *Escherichia coli* in ice in the presence of nucleotides. *Biochim. Biophys. Acta* **1777,** 711–718 (2008).
12. Sazanov, L. A. & Hinchliffe, P. Structure of the hydrophilic domain of respiratory complex I from *Thermus thermophilus*. *Science* **311,** 1430–1436 (2006).
13. Berrisford, J. M. & Sazanov, L. A. Structural basis for the mechanism of respiratory complex I. *J. Biol. Chem.* **284,** 29773–29783 (2009).
14. Fearnley, I. M. & Walker, J. E. Conservation of sequences of subunits of mitochondrial complex I and their relationships with other proteins. *Biochim. Biophys. Acta* **1140,** 105–134 (1992).
15. Mathiesen, C. & Hagerhall, C. Transmembrane topology of the NuoL, M and N subunits of NADH:quinone oxidoreductase and their homologues among membrane-bound hydrogenases and bona fide antiporters. *Biochim. Biophys. Acta* **1556,** 121–132 (2002).
16. Friedrich, T. Complex I: a chimaera of a redox and conformation-driven proton pump? *J. Bioenerg. Biomembr.* **33,** 169–177 (2001).
17. 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).
18. Guénebaut, V., Schlitt, A., Weiss, H., Leonard, K. & Friedrich, T. Consistent structure between bacterial and mitochondrial NADH:ubiquinone oxidoreductase (complex I). *J. Mol. Biol.* **276,** 105–112 (1998).
19. Baranova, E. A., Holt, P. J. & Sazanov, L. A. Projection structure of the membrane domain of *Escherichia coli* respiratory complex I at 8 Å resolution. *J. Mol. Biol.* **366,** 140–154 (2007).
20. Kao, M. C., Di Bernardo, S., Matsuno-Yagi, A. & Yagi, T. Characterization of the membrane domain Nqo11 subunit of the proton-translocating NADH-quinone oxidoreductase of *Paracoccus denitrificans*. *Biochemistry* **41,** 4377–4384 (2002).
21. Kao, M. C., Di Bernardo, S., Matsuno-Yagi, A. & Yagi, T. Characterization and topology of the membrane domain Nqo10 subunit of the proton-translocating NADH-quinone oxidoreductase of *Paracoccus denitrificans*. *Biochemistry* **42,** 4534–4543 (2003).
22. Bernardo, S. D., Yano, T. & Yagi, T. Exploring the membrane domain of the reduced nicotinamide adenine dinucleotide-quinone oxidoreductase of *Paracoccus denitrificans*: characterization of the NQO7 subunit. *Biochemistry* **39,** 9411–9418 (2000).
23. Mamedova, A. A., Holt, P. J., Carroll, J. & Sazanov, L. A. Substrate-induced conformational change in bacterial complex I. *J. Biol. Chem.* **279,** 23830–23836 (2004).
24. Screpanti, E. & Hunte, C. Discontinuous membrane helices in transport proteins and their correlation with function. *J. Struct. Biol.* **159,** 261–267 (2007).
25. Torres-Bacete, J., Sinha, P. K., Castro-Guerrero, N., Matsuno-Yagi, A. & Yagi, T. Features of subunit NuoM (ND4) in *Escherichia coli* NDH-1: topology and implication of conserved Glu^{144} for coupling site 1. *J. Biol. Chem.* **284,** 33062–33069 (2009).
26. Holt, P. J., Morgan, D. J. & Sazanov, L. A. The location of NuoL and NuoM subunits in the membrane domain of the *Escherichia coli* complex I: implications for the mechanism of proton pumping. *J. Biol. Chem.* **278,** 43114–43120 (2003).
27. Baranova, E. A., Morgan, D. J. & Sazanov, L. A. Single particle analysis confirms distal location of subunits NuoL and NuoM in *Escherichia coli* complex I. *J. Struct. Biol.* **159,** 238–242 (2007).
28. 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).
29. Roth, R. & Hagerhall, C. Transmembrane orientation and topology of the NADH:quinone oxidoreductase putative quinone binding subunit NuoH. *Biochim. Biophys. Acta* **1504,** 352–362 (2001).
30. Kao, M. C., Matsuno-Yagi, A. & Yagi, T. Subunit proximity in the H+-translocating NADH-quinone oxidoreductase probed by zero-length cross-linking. *Biochemistry* **43,** 3750–3755 (2004).
31. Murai, M., Sekiguchi, K., Nishioka, T. & Miyoshi, H. Characterization of the inhibitor binding site in mitochondrial NADH-ubiquinone oxidoreductase by photoaffinity labeling using a quinazoline-type inhibitor. *Biochemistry* **48,** 688–698 (2009).
32. Sekiguchi, K., Murai, M. & Miyoshi, H. Exploring the binding site of acetogenin in the ND1 subunit of bovine mitochondrial complex I. *Biochim. Biophys. Acta* **1787,** 1106–1111 (2009).
33. Page, C. C., Moser, C. C., Chen, X. & Dutton, P. L. Natural engineering principles of electron tunnelling in biological oxidation–reduction. *Nature* **402,** 47–52 (1999).
34. Yano, T., Dunham, W. R. & Ohnishi, T. Characterization of the $\Delta\mu_{H+}$-sensitive ubisemiquinone species (SQ_{Nf}) and the interaction with cluster N2: new insight into the energy-coupled electron transfer in complex I. *Biochemistry* **44,** 1744–1754 (2005).
35. Nakamaru-Ogiso, E., Sakamoto, K., Matsuno-Yagi, A., Miyoshi, H. & Yagi, T. The ND5 subunit was labeled by a photoaffinity analogue of fenpyroximate in bovine mitochondrial complex I. *Biochemistry* **42,** 746–754 (2003).
36. Gong, X. *et al.* The ubiquinone-binding site in NADH:ubiquinone oxidoreductase from *Escherichia coli*. *J. Biol. Chem.* **278,** 25731–25737 (2003).
37. Berrisford, J. M., Thompson, C. J. & Sazanov, L. A. Chemical and NADH-induced, ROS-dependent, cross-linking between subunits of complex I from *Escherichia coli* and *Thermus thermophilus*. *Biochemistry* **47,** 10262–10270 (2008).
38. Belogrudov, G. & Hatefi, Y. Catalytic sector of complex I (NADH:ubiquinone oxidoreductase): subunit stoichiometry and substrate-induced conformation changes. *Biochemistry* **33,** 4571–4576 (1994).
39. Gondal, J. A. & Anderson, W. M. The molecular morphology of bovine heart mitochondrial NADH-ubiquinone reductase. Native disulfide-linked subunits and rotenone-induced conformational changes. *J. Biol. Chem.* **260,** 12690–12694 (1985).
40. Verkhovskaya, M. L., Belevich, N., Euro, L., Wikstrom, M. & Verkhovsky, M. I. Real-time electron transfer in respiratory complex I. *Proc. Natl Acad. Sci. USA* **105,** 3763–3767 (2008).
41. Euro, L., Belevich, G., Verkhovsky, M. I., Wikstrom, M. & Verkhovskaya, M. Conserved lysine residues of the membrane subunit NuoM are involved in energy conversion by the proton-pumping NADH:ubiquinone oxidoreductase (Complex I). *Biochim. Biophys. Acta* **1777,** 1166–1172 (2008).
42. Torres-Bacete, J., Nakamaru-Ogiso, E., Matsuno-Yagi, A. & Yagi, T. Characterization of the NuoM (ND4) subunit in *Escherichia coli* NDH-1: conserved charged residues essential for energy-coupled activities. *J. Biol. Chem.* **282,** 36914–36922 (2007).
43. Kervinen, M., Patsi, J., Finel, M. & Hassinen, I. E. A pair of membrane-embedded acidic residues in the NuoK subunit of *Escherichia coli* NDH-1, a counterpart of the ND4L subunit of the mitochondrial complex I, are required for high ubiquinone reductase activity. *Biochemistry* **43,** 773–781 (2004).
44. Kajiyama, Y., Otagiri, M., Sekiguchi, J., Kudo, T. & Kosono, S. The MrpA, MrpB and MrpD subunits of the Mrp antiporter complex in *Bacillus subtilis* contain membrane-embedded and essential acidic residues. *Microbiology* **155,** 2137–2147 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was funded by the Medical Research Council. We thank the European Synchrotron Radiation Facility and the Swiss Light Source for provision of synchrotron radiation facilities. We are grateful to the staff of beamlines ID23, ID29 (ESRF, Grenoble) and X06SA (Swiss Light Source, Villigen) for assistance. We thank J. E. Walker and A. Leslie for discussions of the manuscript.

Author Contributions R.G.E. purified and crystallized the membrane domain of *E. coli* complex I; R.B. purified and crystallized intact *T. thermophilus* complex I; R.G.E. and R.B. collected and analysed X-ray data; L.A.S. designed the project, analysed data and wrote the manuscript, with contributions from R.G.E. and R.B.

Author Information The coordinates and structure factors have been deposited in the RCSB Protein Data Bank under accession codes 3M9C (membrane domain of complex I from *E. coli*) and 3M9S (intact complex I from *T. thermophilus*). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to L.A.S. (sazanov@mrc-mbu.cam.ac.uk).

METHODS

Protein purification. *E. coli* complex I was purified from strain BL21 as described[45]. The membrane and hydrophilic domains were separated by treatment with a high concentration of Mg^{2+}: purified complex I (2.5 mg ml^{-1}) was incubated in 20 mM Bis-Tris pH 6.0, 5% (v/v) glycerol, 400 mM $MgCl_2$, 12.5 mM NaCl, 0.5 mM $CaCl_2$ and 0.25% (w/v) *n*-dodecyl-β-maltoside (DDM, Glycon) for 2 h on ice. The protein solution was diluted with buffer EA (20 mM Bis-Tris pH 6.0, 0.05% DDM), loaded onto a Mono-S HR 5/5 column and eluted with linear gradient of buffer EB (1 M NaCl in buffer A). Fractions containing the membrane arm were pooled, diluted fivefold with 10 mM sodium acetate pH 4.8, 50 mM NaCl, 0.05% DDM, concentrated to 10–11 mg ml^{-1} using 100 kDa MWCO concentrators (Ultracel-100K, Amicon) and used immediately for crystallization trials.

Intact complex I from *T. thermophilus* was purified by a novel procedure. All steps were carried out at room temperature. Membranes were prepared from 150 g of cells[46], resuspended in 50 mM Bis-Tris pH 6.0, 2 mM $CaCl_2$, 100 mM NaCl, 10% (v/v) glycerol, 0.002% (w/v) PMSF, and protease inhibitor cocktail (EDTA-free, Roche), and solubilized for 1 h with 1% (w/v) *n*-tridecyl-β-maltoside (TDM, Glycon). Solubilized material was clarified by centrifugation (150,000*g*, 1 h) and filtration (0.45 μm) and passed through a series of three ion-exchange columns (HiLoad 26/10 Q-Sepharose, HiPrep 16/10 ANX FF and Bio-Scale DEAE 20). All ion-exchange columns were equilibrated with buffer TA (20 mM Bis-Tris pH 6.0, 0.002% PMSF, 2 mM $CaCl_2$, 10% glycerol and 0.05% TDM) and the protein was eluted with a linear gradient of buffer TB (1 M NaCl in buffer A). Fractions containing complex I were identified by NADH:FeCy activity, pooled and diluted in buffer TA before applying to the next column. Pooled fractions from the third ion-exchange column were concentrated to about 1 ml using 100 kDa MWCO concentrators and applied to a HiLoad 16/60 Superdex 200 gel-filtration column, equilibrated with GF buffer (buffer TA with 100 mM NaCl). Fractions were pooled on the basis of purity (assessed by SDS–PAGE), concentrated to about 15 mg ml^{-1} and an additional 15% glycerol was added for storage in liquid nitrogen. NADH:decyl-ubiquinone activity was measured at 50 °C and in the presence of asolectin.

Crystallization. Crystals of the membrane domain of *E. coli* complex I were grown using sitting drop or microbatch crystallization under paraffin oil. First, 0.5% *n*-heptyl-β-glucoside and 2 mg ml^{-1} lipids (1,2-dimyristoleoyl-*sn*-glycerol-3-phosphocholine:*E. coli* polar lipids 3:1 (w/w), Avanti Polar Lipids) were added to the protein solution, then it was mixed 1:1 (v/v) with crystallization reagent, comprising 0.1 M sodium acetate pH 4.8, 0.8–1.0 M sodium or potassium formate and 9% (w/v) PEG4000. Crystals grew at 10 °C for between 3 weeks and 2 months and were rod-like (about 30 × 50 × 200–400 μm). Initially the crystals diffracted to about 7 Å resolution. Diffraction was improved by slow exchange of mother liquor (using microdialysis buttons at 4 °C) to 0.1 M sodium acetate, 0.1 M sodium or potassium formate and 40% PEG4000 over a period of 2–4 days, increasing PEG4000 concentration in 5% steps. The simultaneous decrease in salt concentration together with increase in PEG concentration was essential for the success of this dehydration treatment. Crystals were plunge-frozen in liquid nitrogen without additional cryo-protection. For heavy atom derivatization, mother liquor was exchanged to 0.1 M sodium citrate pH 5.2, 15% PEG4000 and crystals were soaked in Ta_6Br_{12} solution overnight at 4 °C. Subsequently PEG4000 concentration was raised to 40% and crystals were plunge-frozen in liquid nitrogen.

Crystals of intact *T. thermophilus* complex I were grown at 23 °C using sitting drop crystallization. The $P2_1$ crystals were grown by mixing complex I (20 mg ml^{-1}) with a crystallization reagent comprising 100 mM Bis-Tris pH 6.0, 16% (w/v) PEG4000, 100 mM KCl, 100 mM glutaric acid pH 6.0 and 2.04 mM *n*-octyl-β-maltoside (OM) fluorinated at 1:1 ratio. These conditions were obtained after extensive optimization that improved the resolution from 8 Å to about 4.5 Å through use of glutaric acid as a small molecule additive and OM fluorinated as a detergent additive. The $P2_12_12_1$ crystals were grown by first incubating complex I with *E. coli* polar lipids (prepared in 10 mM Bis-Tris pH 6.0 and 2% (w/v) DDM; Avanti Polar Lipids) at 9:1 (w/w) ratio, and then mixing this solution with a crystallization reagent comprising 100 mM Bis-Tris pH 6.0, 18% PEG4000, 100 mM KCl, 100 mM glutaric acid pH 6.0, 1.02 mM OM fluorinated and 10 mM spermidine tetra-hydrochloride at 1:1 ratio. These crystals were not twinned, but diffracted only to 8.5 Å resolution even after optimization. Further post-crystallization treatments did not improve resolution limits for both crystal forms, unlike for the membrane domain crystals. The $P2_1$ crystals were rod-like (about 20 × 20 × 200–300 μm) and the $P2_12_12_1$ crystals were plate-like (about 30 × 50 × 300–400 μm). Crystals were fully grown within 1 week and were cryo-cooled by plunging into liquid nitrogen, using 25% (v/v) glycerol as the cryoprotectant.

Data collection and processing. Data were collected at 100 K with an ADSC Q210 detector at beamline ID29, with a Marmosaic 225 detector at beamline ID23-1, with a Marmosaic 225 detector at beamline ID23-2 (using helical strategy for collection along the crystal to minimize radiation damage) at the European Synchrotron Radiation Facility (ESRF, Grenoble) and with a Marmosaic 225 detector at the microdiffractometer MD2 of beamline X06SA (using several points along the crystal to minimize radiation damage) at the Swiss Light Source (Villigen). Image data were processed with MOSFLM and SCALA from the CCP4 suite[47].

Membrane domain of *E. coli* complex I crystallized in the orthorhombic space group $P2_12_12_1$ with one molecule per asymmetric unit (ASU) (Supplementary Table 1). The crystals were strongly non-isomorphous. The locations of TaBr clusters (six sites), used for phasing, were determined using SHELXD[48] on the anomalous data from the Ta absorbance peak. In total, seven different crystals were used for phasing: four native and three TaBr derivatives. First, the best TaBr crystal was used with the most isomorphous native crystal for phasing in SHARP[49] and density modification with SOLOMON, PIRATE and DM from the CCP4 suite. A preliminary α-helical structure was built (mostly automatically using the find_helices_strands module of PHENIX[50]) and used for molecular replacement with PHASER[51] on the other five data sets. This allowed us to determine the three most isomorphous native/TaBr pairs for three independent phasing runs in SHARP. The phasing pairs were not isomorphous to each other owing to differences in unit cell dimensions and the position of the molecule in the unit cell. Therefore, DMmulti program from the CCP4 suite was used, with the molecular masks and transformation matrices obtained from molecular replacement solutions. Multi-crystal electron density averaging and modification (with histogram matching and solvent flattening) was performed using three phased data sets and one additional not phased native data set (Supplementary Table 1). This resulted in marked improvement of the electron density (Supplementary Fig. 2), so that the model could be manually updated and completed in O[52] and COOT[53]. In the crystals, the membrane domains are stacked vertically, interacting via hydrophilic surfaces, with no visible 'side' contacts (Supplementary Fig. 3), leading to high solvent content (about 73%) and anisotropic diffraction. It is possible that some of side contacts are mediated by lipids, which are necessary for crystallization.

For the intact complex I from *T. thermophilus*, crystals in the monoclinic $P2_1$ space group were obtained under a range of conditions and all were subject to the rare case of pseudo-merohedral twinning[54], imitating space group $C222_1$. The twin fraction was high, in the range of 0.46 to 0.49 for several data sets, as estimated in PHENIX using twin refinement target. Two sets of clear molecular replacement solutions (related by the twinning operator) were obtained in PHASER with the *T. thermophilus* hydrophilic domain (PDB 3I9V, ref. 55) and the *E. coli* membrane domain model, resulting in two molecules of intact complex per ASU (68% solvent content; Supplementary Fig. 5b). A clear solution for one molecule per ASU (72% solvent content) was obtained in $P2_12_12_1$ crystal form (Supplementary Fig. 5c). When $P2_1$ and $P2_12_12_1$ solutions are aligned by the hydrophilic domains, the membrane domains align as well (Supplementary Fig. 5a), confirming the correctness of molecular replacement solutions.

The $P2_1$ solutions were then further processed in PHASER by allowing the membrane domain subunits to refine as independent rigid bodies. This resulted in slight shifts of subunits relative to the *E. coli* model (Supplementary Fig. 6bc). Consistently, the shifts were similar for both molecules in the ASU. The electron density for the membrane domain improved, including the density for the subunit NuoH/Nqo8, lacking in the *E. coli* model (Supplementary Fig. 6a). Despite twinning, the density was somewhat improved by solvent flattening and NCS averaging in DM. The α-helical model of subunit NuoH/Nqo8 was built in COOT using the DM modified and the original density (both from PHASER and from PHENIX, calculated allowing for twinning).

Bioinformatics. ClustalW2 was used for sequence alignment[56], TMHMM v2.0 (ref. 57) for prediction of transmembrane helices, PsiPred[58] and PredictProtein[59] for secondary structure prediction.

45. 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).
46. Hinchliffe, P., Carroll, J. & Sazanov, L. A. Identification of a novel subunit of respiratory complex I from *Thermus thermophilus*. *Biochemistry* **45**, 4413–4420 (2006).
47. CCP4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
48. Schneider, T. R. & Sheldrick, G. M. Substructure solution with *SHELXD*. *Acta Crystallogr. D* **58**, 1772–1779 (2002).
49. de La Fortelle, E. & Bricogne, G. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* **276**, 472–494 (1997).

50. Adams, P. D. *et al. PHENIX*: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **58,** 1948–1954 (2002).
51. McCoy, A. J. *et al. Phaser* crystallographic software. *J. Appl. Cryst.* **40,** 658–674 (2007).
52. Jones, T. A. & Kjeldgaard, M. Electron-density map interpretation. *Methods Enzymol.* **277,** 173–208 (1997).
53. Emsley, P. & Cowtan, K. *Coot*: model-building tools for molecular graphics. *Acta Crystallogr. D* **60,** 2126–2132 (2004).
54. Rudolph, M. G., Wingren, C., Crowley, M. P., Chien, Y. H. & Wilson, I. A. Combined pseudo-merohedral twinning, non-crystallographic symmetry and pseudo-translation in a monoclinic crystal form of the γδ T-cell ligand T10. *Acta Crystallogr. D* **60,** 656–664 (2004).
55. Berrisford, J. M. & Sazanov, L. A. Structural basis for the mechanism of respiratory complex I. *J. Biol. Chem.* **284,** 29773–29783 (2009).
56. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23,** 2947–2948 (2007).
57. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305,** 567–580 (2001).
58. McGuffin, L. J., Bryson, K. & Jones, D. T. The PSIPRED protein structure prediction server. *Bioinformatics* **16,** 404–405 (2000).
59. Rost, B., Yachdav, G. & Liu, J. The PredictProtein server. *Nucleic Acids Res.* **32,** W321–W326 (2004).