The prototypical hydrogen-producing enzyme, the membrane-bound formate hydrogenlyase (FHL) complex from *Escherichia coli*, links formate oxidation at a molybdopterin-containing formate dehydrogenase to proton reduction at a [NiFe] hydrogenase. It is of intense interest due to its ability to efficiently produce H2 during fermentation, its reversibility, allowing H2-dependent CO2 reduction, and its evolutionary link to respiratory complex I. FHL has been studied for over a century, but its atomic structure remains unknown. Here we report cryo-EM structures of FHL in its aerobically and anaerobically isolated forms at resolutions reaching 2.6 Å. This includes well-resolved density for conserved loops linking the soluble and membrane arms believed to be essential in coupling enzymatic turnover to ion translocation across the membrane in the complex I superfamily. We evaluate possible structural determinants of the bias toward hydrogen production over its oxidation and describe an unpredicted metal-binding site near the interface of FdhF and HycF subunits that may play a role in redox-dependent regulation of FdhF interaction with the complex.
The conserved architecture of the complex I superfamily consists of a mostly soluble arm catalyzing both redox half-reactions and a membrane-bound arm, through which protons can be pumped. This architecture appears to have emerged by interaction of a soluble [NiFe] hydrogenase with a multiple resistance and pH (MRP) sodium-proton antiporter. In the course of evolution, the pocket originally containing the [NiFe] hydrogenase active site has been modified to bind and reduce quinone in complex I, and turnover at this site is widely believed to be coupled to ion translocation across the membrane.

Here, we present full atomic models of the complex in the oxidized and reduced states, providing important insight into the function and regulation of FHL and related complexes.

**Results**

**Aerobic and anaerobic purification of FHL and single-particle cryo-EM studies**

The FHL complex was purified from a strain bearing an internal 10x His affinity tag in Hyc, according to a modified protocol (see Methods). The structure of aerobically isolated FHL, determined to 3.4 Å resolution (focused refinements 3.0–3.4 Å) by cryo-EM single-particle analysis provided important detail on the organization of the heptameric complex (Fig. 1a, b). However, the map density for the hydrogenase [NiFe] active site and proximal iron-sulfur cluster G1 was weak and poorly resolved (Fig. 2a), which we attribute to oxidative damage. We developed an anaerobic isolation protocol that allowed us to purify a complex exhibiting 20-fold higher hydrogen-uptake activity (620 ± 1445 s⁻¹) compared to aerobically purified complex from this study (253 ± 164 s⁻¹) (Fig. 2b) and previous studies (2.44 ± 0.62 U mg⁻¹ = 13 s⁻¹). The size exclusion chromatography (SEC) elution profile of the aerobically isolated complex shows two major peaks, the slower-eluting of which corresponds to a soluble complex lacking the membrane-associated subunits HycC and HycD (Supplementary Fig. 2a). The formate dehydrogenase subunit FdhF, which shows partial occupancy in the aerobically purified complex, was entirely absent in the anaerobically isolated complex, was entirely absent in the anaerobically purified complex (Supplementary Fig. 1). In contrast, the SEC profile of anaerobically purified protein shows a single predominant peak, with a much smaller peak corresponding to the complex without HycC/HycD (Supplementary Fig. 2a).

The formate dehydrogenase subunit FdhF, which shows partial occupancy in the aerobically purified complex, was entirely absent in the anaerobically isolated complex, was entirely absent in the anaerobically purified complex (Supplementary Fig. 2a). In contrast, the SEC profile of anaerobically purified protein shows a single predominant peak, with a much smaller peak corresponding to the complex without HycC/HycD (Supplementary Fig. 2a).

**Evolutionary relationship of FdhF and complex I subunit NDUF51**

According to a theory of modular evolution, complex I evolved from the association of a soluble [NiFe] hydrogenase with an MRP-type sodium-proton antiporter. An early form of this respiratory system is represented by the energy-conserving group 4 membrane-bound hydrogenase (MBH) from the hyperthermophilic archaeon *P. furiosus*, a complex that accepts electrons from oxidative reactions in the cell via the electron transfer protein ferredoxin. In contrast, like complex I, FHL catalyzes both the oxidative and reductive half-reactions directly. Molybdopterin-containing enzymes including FdhF exhibit sequence homology to complex I subunit NDUF51 (Supplementary Table 3). Sequence comparison suggests that NDUF51 forms part of an alternate electron-input module, which may be required in bacterial 12-subunit complex I analogs as they lack an NADH-oxidizing module. In other bacterial species, the pathway appears to have become a relic, leaving the NDUF51-ligated [4Fe4S] cluster N7 too remote to play a role in electron transfer between active sites. Sequence similarity of FdhF and HycB to NDUF51 raises the possibility that an FHL-type complex may be the most recent MBH ancestor of complex I.
Superposing our structure with the C-terminal region of *Thermus thermophilus* Nqo3 (NDUFS1) gives an overlap of the [4Fe4S] cluster A1 of FdhF with the off-pathway [4Fe4S] cluster N7 of Nqo3 (NDUFS1) (Supplementary Fig. 6a). However, the way in which FdhF interfaces with the FHL complex is different to that seen for Nqo3 (NDUFS1) in complex I (Supplementary Fig. 6b, c). The significant difference in architecture suggests that in spite of the sequence homology between these subunits, their association to members of the complex I superfamily is likely to have arisen from independent evolutionary events.

A recently published structure of the soluble *Rhodobacter capsulatus* FDH complex\(^{38}\) revealed a subunit architecture and cofactor arrangement much more similar to that seen in Nqo3 (NDUFS1), in which the equivalent of the off-pathway cluster N7 connects to the electron relay via an additional [4Fe4S] cluster. In contrast, the way in which FdhF interacts with HycB is more similar to the interaction of the α- and β-subunits of the *E. coli* formate dehydrogenase-N\(^{39}\), a homolog of FdhF that plays a role in nitrate respiration (Supplementary Fig. 6d).
FdhF interacts with HycB and an unpredicted metal ion of HycF

Electron transfer between the electron-input module (FdhF) and the primary hydrogenase modules (HycG and HycE) is mediated by the ferredoxin-like proteins HycB and HycF, as predicted from genetic experiments. A loop region between the B3 and B2 [4Fe4S]-cluster-ligating cysteine motifs of HycB interacts extensively with HycF and HycE. HycF has an N-terminal extension that comprises a long loop, with a short amphipathic helix near the membrane surface at the extreme N-terminus. This structural feature is present in all structures of the complex I superfamily solved to date and has been shown to bind cardiolipin, with a suggested influence upon global structural dynamics and quinone access to the active site. The C-terminal extension of HycF interacts with HycB and contains a cysteine quartet (Cys117, Cys120, Cys155 and Cys158) that coordinates an unpredicted cofactor at the periphery of the protein, consistent with a single metal ion that we have tentatively assigned as iron. This is consistent with previous inductively coupled plasma mass spectrometry (ICP-MS) analysis that found a higher Fe:Ni ratio than expected given the predicted cofactors, and that did not report additional metals. The coordination sequence (CxxC-x34-CpxC) bears similarity to ligand-binding sequences for rubredoxin-type Fe2+/3+ centers, or Zn2+. Neither the fold nor the detailed sequence closely matches characterized rubredoxins. The metal ion is about 17 Å from the B3 [4Fe4S] cluster of HycB (Fig. 1c). The position of the ion argues against a role in conducting electrons between active sites, however the ion could be in redox equilibrium with the iron-sulfur cluster relay: Marcus theory predicts electron transfer on the order of $10^2$ s−1 over a distance of 17 Å for $\Delta G = 0$.

The metal-binding site forms part of a positively charged patch of HycF that is directly opposite a patch of acidic residues of FdhF (Fig. 4a). We suggest that a redox-active ion in this binding site could play a role in regulating affinity of the complex for FdhF as reduction of the ion would decrease the strength of charge-charge interaction between the two sites. Such a mechanism would be expected to lower the rate of hydrogen-dependent CO2 reduction by the complex, since this reaction requires interaction of a reduced Hyc complex with FdhF. Arg122 of HycF, which is near to the ion, exhibits two alternate conformations in the aerobically isolated structure: one facing toward the ion and the other toward the acidic patch of FdhF (Fig. 4b). The structure of anaerobically isolated FHL shows the arginine in a single conformation facing the metal ion (Fig. 4b). The ion-binding motif and Arg122 are conserved in FHL complexes from several other bacteria and the HyfH subunit of the related E. coli FHL-2 complex but not in the homologous complex I subunit NDUFS8 (Supplementary Fig. 10a).

Fig. 2 | Comparison of aerobically and anaerobically prepared FHL. a Solid-view representation of unsharpened maps of aerobically prepared FHL shows strong intensity for the [4Fe4S] clusters in FdhF (A1), HycB (B1-B4) and HycF (F1, F2). Density for the proximal [4Fe4S] cluster G1 in HycG and the [NiFe] cofactor in HycE is weak (the approximate position is indicated with a dashed ellipse). The anaerobically prepared FHL shows strong density for the proximal [4Fe4S] cluster (G1) and the [NiFe] cofactor. b Hydrogen-uptake activity of aerobically (253 ± 164 s−1) and anaerobically (6201 ± 1445 s−1) prepared FHL. Values indicate mean ± standard deviation. Initial slopes were determined for 4 independent assays from n = 1 biologically independent samples for the aerobic preparation and from 13 independent experiments from n = 2 biologically independent samples for the anaerobic preparation. Source data are provided as a Source Data file.

Fig. 3 | Geometry of the Mo(MGD)2 cofactor. a Atomic model and cryo-EM density of the Mo(MGD)2 cofactor with the proximal MGD (MGDP), distal MGD (MGDO), and the metal-coordinating selenocysteine (Sec) in subunit FdhF of the aerobically isolated FHL. b Superposition of our aerobically isolated FHL model with crystal structures of oxidized (blue; PDB 1FDO) and reduced (red; PDB 2IV2) FdhF. The geometry of MGDP is similar to the crystal structure of FdhF in the oxidized [Mo(VI)] state.
The catalytic mechanism of [NiFe] hydrogenases has been most extensively studied for soluble hydrogenases, and is summarized in recent reviews. Similarity in sequence and structure support a conserved heterolytic mechanism for the reversible interconversion of H₂/2H⁺ across different types of [NiFe] hydrogenases.

The active site is buried deep within the polypeptide, so that protons, electrons and hydrogen require pathways to exchange with the environment. The electron pathway provided by the iron-sulfur clusters is well established and molecular hydrogen appears to be able to enter or leave the active site via several hydrophobic routes as shown by MD simulations and crystallographic studies. Highly adapted proton transfers are central to efficiency of electrocatalysts including hydrogenases, and several studies have traced the pathway taken by protons to/from the [NiFe] active site. At least one of the protons is transferred to a terminal cysteine ligand of nickel, (Cys546 in D. vulgaris HydB, Cys531 in E. coli FHL), and from there to a strictly conserved carboxylate (Glu34 in D. vulgaris HydB) that is essential for function and occupies a position equivalent to a highly conserved histidine (His95 in Y. lipolytica NDUFS2) of complex I. This residue exchanges protons with Cys531, in spite of the fact that a sequence alignment of MBHs shows Asp192 of HycE closer to Cys531-S suggesting that this residue exchanges protons on the β₂ loop glutamate, several proton pathways have been proposed, based on MD simulations and high-resolution X-ray structures.

From the β₁-β₂ loop glutamate, several proton pathways have been proposed, based on MD simulations and high-resolution X-ray structures. The β₁-β₂ loop appears less restricted in FHL than in soluble hydrogenases, having fewer stabilizing H-bonds (Fig. 5c). In FHL the Glu193 is too far (4 Å) from Cys531-S to allow direct proton transfer (Fig. 5a), suggesting that this residue exchanges protons with Cys531, in spite of the fact that a sequence alignment of MBHs shows Asp192 of HycE closer to Cys531-S suggesting that this residue exchanges protons on the β₂ loop glutamate, several proton pathways have been proposed, based on MD simulations and high-resolution X-ray structures.
crystallography structures of soluble [NiFe] hydrogenases, but the residues involved in many of these pathways are not strictly conserved in MBHs (Supplementary Fig. 7). Mutational evidence and comparison to [FeFe] hydrogenases and synthetic catalysts suggests a role for the strictly conserved canopy residues (HycE Arg478/Asp529) in the first proton abstraction from H₂ (or, in the direction of H₂ production, protonation of a metal-bound hydride). The canopy residues lie at the end of a putative proton transfer pathway including HycE His284, Glu437 and Arg395, which appears to be conserved in FHL, MBH², soluble hydrogenases as well as in complex I (Fig. 5b). The canopy arginine (HycE Arg478) is present as either an arginine or more commonly a lysine in complex I, while the canopy aspartate (HycE Asp529) appears to be strictly conserved (Fig. 5d). Mutation of either residue to non-protonatable amino acids substantially impairs quinone reductase activity of complex I⁶⁰. Studies of complex I have suggested the β₁-β₂ loop may be reprotonated through the TMH1-2 loop of subunit ND3 (HycC homolog, Supplementary Table 3) or by the so-called E-channel, a glutamate-rich pathway connecting the quinone reductase active site to the bulk (Fig. 5c).
reduction site to the membrane arm (Fig. 1d), although the proton pathway to the E-channel remains under debate\(^{42,43}\).

**Structural basis of hydrogen production**

Hyd-3, the hydrogenase of the FHL complex, exhibits a bias toward hydrogen production and a very low degree of product (H\(_2\)) inhibition for H\(_2\) production, more similar to the unrelated [FeFe] hydrogenases than to other [NiFe] hydrogenases. The structural basis of this difference is unknown and is of critical relevance to the design of biological and synthetic catalysts. Our structures suggest two possibilities for this observed difference. Ser283 of HycE (Fig. 5a), one of three so-called canopy residues at the [NiFe] site, is variably an Asp, Asn, or Ser in other [NiFe] hydrogenases (Fig. 5d). This position is frequently occupied by Asp in H\(_2\)-oxidizing enzymes, while both the MBHs and the relatively distantly related [NiFeSe] hydrogenases contain a Ser at this position\(^{42}\), and both groups exhibit enhanced H\(_2\) production activity\(^{43,44}\). Mutation of this residue (HybD D118S) to alanine in the H\(_2\)-oxidizing hydrogenase-1 (Hyd-1) from *E. coli* resulted in an enzyme exhibiting an extremely high K\(_H2\)(H\(_2\)); however, this mutant did not produce H\(_2\)\(^{45,42}\). As noted by the authors, this could be due to other limitations on H\(_2\) production in Hyd-1\(^{10}\); a mutation of D118 to serine was not reported to our knowledge. A second possibility is that the conformational flexibility of the β1-β2 loop may play a role in tuning the properties of the active site. Recent MD simulations of the P. furiosus MBH showed that the β1-β2 loop comes nearer to the [NiFe] site in its protonated state than in its deprotonated state\(^{44}\). Such a dependence would favor the protonated state of the [NiFe] cluster, and may thereby disfavor binding of H\(_2\). In this way, the enhanced flexibility of the β1-β2 loop could underlie a tendency toward H\(_2\) production, even in the presence of high p(H\(_2\)).

Conserved loops form the interface between catalytic and membrane subunits

The β1-β2 loop lies at the interface of the soluble and membrane arms along with three additional peptide loops, which are proposed to play a key role in coupling turnover with proton translocation across the membrane in complex I\(^{10}\); all four loops have homologs in FHL (Supplementary Fig. 3). For simplicity, the complex I nomenclature for these loops will be used throughout. The ND3 loop (HycC TMIHS-THM16 loop), β1-β2 loop (HycE β1-β9 loop), ND5β5 loop (HycG residues 69-80) and the ND1 loop (HycD TMI5-TMH6 loop) (Supplementary Fig. 3b- e) are all well-resolved in our anaerobic cryo-EM map and could be modeled with confidence. In the map obtained from aerobically prepared sample only weak density was observed for the ND1 and β1-β2 loops (Supplementary Fig. 4a, d). The ND3 loop, which is frequently disordered in complex I structures and the structure of *P. furiosus* MBH, is in a similar conformation to that observed for active mouse complex I\(^{43}\), with the conserved Cys537 of HycC in a binding pocket formed by Phe122 of HycD and His188 of the β1-β2 loop of HycE.

The membrane arm harbors a putative substrate proton pathway

The membrane arm of complex I homologs includes the highly conserved membrane-anchor subunit (ND1 in complex I nomenclature) together with a variable number of antiporter-like subunits. FHL has a minimal membrane arm consisting only of HycD (ND1 homolog) and the antiporter-like subunit HycC (ND2, ND4, ND5 homolog; Supplementary Table 3). In contrast to *P. furiosus* MBH, which is believed to establish a sodium gradient\(^{39,40}\), FHL lacks homologs to MbhA-G. HycD has eight highly tilted transmembrane helices (Fig. 6a) that build up a hydrophilic cavity connecting the hydrogenase active site with the center of the membrane arm (Fig. 6c). The cavity is lined with a number of conserved charged residues, similar to other members of the complex I superfamily. Compared to complex I architecture, HycC is rotated by 180° in the plane of the membrane relative to HycD in FHL (Fig. 1d), consistent with structures of the MBH and MBS complexes of *P. furiosus*. Unlike the *P. furiosus* MBH structure, in which the membrane subunits MbhM and MbhH are separated by lipids\(^{45}\), HycC and HycD are tightly packed against each other, allowing close contact between residues of the proposed proton-conducting E-channel. HycC comprises 16 TMHs (Fig. 6a) exhibiting strong homology to antiporter-like subunits of complex I\(^{16}\), MBH\(^{27}\), MBS\(^{28}\), and MBP\(^{46}\), with characteristic broken helices in TMH7 and TMH12 that provide flexibility and are thought to play a role in proton pumping in the complex I superfamily.

In FHL the E-channel extends from Glu201, Glu206 and Glu199 on the conserved NDI loop along Glu30, Glu59, Arg108, Glu338 to Tyr102, which is located at the subunit boundary approximately 4 Å from the adjacent Glu391 in HycC TMH12 (Fig. 6c). Site-directed mutagenesis of HycD shows that substitutions of residues along this pathway, reaching as far as Glu338 at the interface of HycD/HycC (Fig. 6c), strongly impact hydrogen production in vivo\(^{40}\). This provides strong support for the idea that either a substrate proton travels through this pathway to the [NiFe] active site, or that turnover is tightly coupled to proton translocation via this site. Interestingly, HycC Glu391 is only conserved for complex I subunit ND4 and its homologs, differing from other group 4 membrane-bound hydrogenases, in which a highly conserved lysine takes up this position (Supplementary Fig. 14a). Densities consistent with water suggest well-ordered water molecules form H-bonds with Glu391. We examined the structure for proton pathways from the solvent to Glu391 and found that FHL differs from previously characterized homologs in this regard. A path from the N-side conserved in complex I subunit ND5, MrAP, MbhI and MbsH/MbsH', is interrupted in HycG by Glu285 and Ala230 (Fig. 6c and Supplementary Fig. 14a). Instead, a pathway including residues Glu294, His291, Cys343 and Lys433, which is not conserved in MBH, MBS or complex I, leads from Glu391 to the N-side. This observation is consistent with a previous mutational study showing that HycC mutants E391A and D354A reduced hydrogen production by 45%, whereas E135A, H222A, K239A, T292A, H328A and K336A mutations did not cause a significant decrease in hydrogen production in vivo\(^{40}\). It was previously noted that the lysine of a conserved Glu/Lys pair in antiporter-like subunits is substituted by Leu208 in HycC\(^{46}\), while the strictly conserved Glu338 on TMH5 is present. These residues are located toward the membrane-exposed end of the HycC subunit. We did not observe a clear proton pathway connecting the E-channel to the P-side of the membrane.

While the majority of HycC (N-terminal residues 1-485) is homologous to antiporter-like subunits, the C-terminal extension (residues 500-608) is homologous to subunit ND3 (Supplementary Fig. 14b). This region includes two transmembrane helices, separated by a long loop, equivalent to the ND3 loop. A similar fusion is seen in MbsH of MBS\(^{46}\). A lateral helix (HL) between TMH15 and TMH16 is a conserved element of antiporter-like subunits in complex I-related enzymes. The helix length varies between different members of the superfamily to accommodate interaction with between one (FHL) and three antiporter-like subunits (complex I)\(^{16}\). Additional density between the helix HL and TMH7 of HycC could be confidently assigned as cardiolipin (CDL). Arg567, Tyr290, His215, Gln220 and Asn571 of HycC interact with the headgroup of CDL (Fig. 6b), whereas non-polar residues in the broken helix in TMH7 interact with the alkyl chains. Cardiolipin plays an important role in the stability and activity of respiratory complexes in prokaryotes\(^{48,49}\) and eukaryotes\(^{50}\) including complex I.

**Discussion**

Our structures of the FHL complex reveal the basis of hydrogen production in *E. coli*. We show that anaerobic isolation results in vastly higher hydrogenase activity, and differences between the aerobic and anaerobic structures indicate that irreversible damage occurs at the [NiFe] site and proximal iron-sulfur cluster when the protein is purified...
aerobically. Our results suggest that the FdhF subunit is less stably attached to the complex under reducing conditions. This observation could be related to a redox-state change of an unprotonated ion coordinated by a cysteine quartet in HycF near to the interface with FdhF. Further biochemical and mutational studies will be required to evaluate this proposal.

Structural similarity to respiratory complex I is obvious for six of the seven subunits of our atomic model. In spite of sequence and structural homology of FdhF to Nqo3, the interaction of FdhF with the membrane and soluble arms, throughout several billions of years of evolution, strongly support the idea that a conserved core mechanism underlies coupling of redox chemistry to ion translocation across the superfamily. An obvious but critical difference between MBHs and complex I is that the product of the MBH reaction (H2) is a small, non-polar gas molecule bearing no discernible similarity to quinol. Therefore, an evolutionary perspective speaks against any mechanistic proposal that fundamentally depends on quinone/quinol dynamics, favoring instead mechanisms that rely on movement of charges (electrons and/or substrate protons) through the complex.

Methods

Protein expression and purification

E. coli strain MG059ei, which harbors a sequence encoding a His60-tag placed after Gly-83 in the hycE gene of the hycABCDEFGHI operon was expressed in LB media supplemented with 0.4% glucose by anaerobic fermentation in a 200 L reactor (Sartorius). The culture was inoculated...
with 400 mL of aerobically grown preculture and stirred at 50 rpm for 16 h at 37 °C. Cells were harvested anaerobically in a flow centrifuge (Carl Padberg Zentrifugenbau), flash frozen in liquid N₂ and stored at −80 °C. Both the anaerobically and anaerobically purified samples were prepared from these cells.

For the anaerobic isolation, 10-20 g of frozen cell pellet was resuspended in lysis buffer (20 mM Tris, 200 mM NaCl, 1 mM DTT, pH 7.5) containing 50 µg/mL lysozyme, 10 µg/mL DNase I and protease inhibitor cocktail (Roche Diagnostics) and passed through a cell disrupter (Constant Systems) two times at 1.8 kbar. After removal of unlysed cells by centrifugation (4000 x g for 15 min), the membrane fraction was prepared by ultracentrifugation at 100,000 x g for one hour, resuspended in lysis buffer and protein was solubilized by addition of 0.5% LMNG (Anatrace) for 1 h. The lysate was supplemented with 0.5% CMC–4 fluorinated fos-choline-8 (Anatrace) was added −4 mg/mL, and subsequently injected onto a Superdex 200 Increase 5/150 GL (GE Healthcare), washed with 6 column volumes (CV) wash buffer (20 mM Tris, 200 mM NaCl, 50 mM imidazole, 1 mM DTT, 0.005% LMNG, pH 7.5) and eluted with a gradient of 50 mM imidazole and loaded onto a 5 mL HisTrap FF column (GE Healthcare), washed with 6 column volumes (CV). Fractions containing FHL (70–300 mM imidazole) were pooled and concentrated to 2–4 mg/mL, and subsequently injected onto a Superdex 200 Increase 5/150 GL (GE Healthcare) size exclusion column pre-equilibrated with 25 mM Tris, 700 mM NaCl, 0.005% LMNG, 1 mM DTT). The SEC step revealed peaks due to the full FHL complex and the dissociated soluble arm (Supplementary Fig. 1a).

Anaerobic isolation of FHL was performed in a similar fashion, with the following modifications. Purification was conducted in an anaerobic chamber (<1 ppm O₂, Coy Laboratory Products) in an atmosphere of 3-4% hydrogen in nitrogen. All buffers were made anaerobic by extensive degassing and subsequent equilibration within the anaerobic chamber (>24 h). Cells were lysed on ice with an ultrasonicator (Bandelin) running 6 cycles of 1 min at 200 W. In contrast to the aerobic procedure, the reduced FHL complex could be eluted as a monodisperse peak from the size exclusion column, which only showed small amounts of a complex lacking the membrane subunits HycC and HycD (Supplementary Fig. 2a). FdhF was detached from the complex during purification under reducing conditions and only traces of an ~90 kDa band could be identified by SDS-PAGE (Supplementary Fig. 2a). Repeated efforts to obtain a complete, anaerobically purified complex were unsuccessful. The cause for this is unclear, but it is worth noting that the difference cannot be due to differential expression, as aliquots of the same cells were used for both purification protocols. Small-scale aerobic purifications of cells lysed by sonication also allowed isolation of complex containing FdhF, while anaerobic purification following cell lysis by bead-beating or B-PER (Thermo Fisher Scientific) resulted in complex lacking FdhF, suggesting that the cell lysis method is not a deciding factor in copurification of the FdhF subunit.

Cryo-EM sample preparation and data acquisition
Purified FHL was concentrated to 1–2 mg/mL using an Amicon centrifugal concentrator (100-kDa cutoff, Merck) and 3 µL of protein solution was applied to freshly glow-discharged C-flat 1.2/1.3 or 2/1 holey carbon grids (Protochips), automatically blotted and plunge frozen in liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific) at 4 °C with 100% humidity. Cryo-EM grids of anaerobically purified FHL were prepared similarly within the anaerobic chamber except that 0.5x CMC–fluorinated fos-choline-8 (Anatrace) was added immediately before the sample was applied to the grid. This resulted in a more even particle distribution and orientation within the ice (Supplementary Fig. 2b).

For the aerobically prepared sample a dataset of 1,207 dose-fractionated movies with 40 frames were recorded automatically with EPU (Thermo Fisher Scientific) on an FEI Titan Krios G2 transmission electron microscope operating at 300 kV equipped with a Bioquantum energy filter and K2 direct electron detector (Gatan) at a pixel size of 0.828 Å operated in counting mode. The total dose was ~72 e/Å² and the defocus range was set from −2.2 to −1.6 µm.

Cryo-EM data of anaerobically prepared FHL were acquired automatically on an FEI Titan Krios G2 transmission electron microscope operating at 300 kV equipped with a Bioquantum energy filter and K2 direct electron detector (Gatan) at a pixel size of 0.831 Å in counting mode. 7,338 dose-fractionated movies of anaerobically purified FHL were collected using fast acquisition by aberration-free image shift (AFIS) in EPU. Movies contained 61 frames at a total dose of ~61 e/Å² and a defocus range from −2.5 to −1.0 µm.

Image processing
Dose-fractionated movies of aerobically prepared FHL were subjected to motion correction and dose-weighting using MotionCor2 and the per-micrograph contrast transfer function (CTF) was determined by CTFFIND4.1 (major processing steps are visualized in Supplementary Fig. 1c). Initially, particles were picked manually from ten micrographs in order to train a model for neural-network picking in cryoYOLo, particle coordinates were imported into RELION-3.0 and subsequently extracted from RELION-3.0 with a box size of 480 by 480 pixels. 2D classification was performed in cryoSPARC, 150,117 particles were selected and subjected to homogenous refinement using a de novo reconstruction generated from the same dataset. Particle coordinates were converted into a star file using cryoSPARC and reimported to RELION for 3D classification. The predominant 3D class (91,127 particles) was selected and refined to a resolution of 3.86 Å. Two rounds of CTF refinement and Bayesian polishing with intermediate 3D refinements yielded a consensus map with a global resolution of 3.40 Å. Focused refinements with masks covering individual parts of the complex gave resolutions of 3.06 Å (FdhF), 2.99 Å (soluble arm without FdhF) and 3.43 Å (membrane arm) respectively. An initial model was built using the globally refined maps, and this model was used as input with the focussed maps and consensus map to generate a composite map using phenix.combine focused maps that was used for further model building and refinement (Supplementary Fig. 1d).

For the dataset of anaerobically purified FHL, initial image processing steps were performed in RELION-3.1 (major processing steps are visualized in Supplementary Fig. 2c). Movies were sorted into optics groups according to their EPU AFIS metadata before motion correction and CTF estimation. Particles were initially picked on a micrograph subset with RELION's AutoPick implementation; positive picks based on appearance of 2D class averages were selected and used to train a particle-detection model with Topaz78. The Topaz model was used for particle picking from all micrographs, yielding 686,602 particles, which were initially extracted with a box size of 480 by 480 pixels and 4-fold down-sampling. Initial 2D classes were selected based on interpretable features and used for de novo initial model generation. In a first, coarse 3D classification all particles were aligned to the initial model, lowpass filtered to 60 Å. The major 3D class contained 565,559 particles that were subjected to another round of 3D classification using the same parameters. The major class, containing 300,443 particles, was selected and extracted with a box size of 384 by 384 pixels and 2-fold down-sampling. 3D auto refinement gave a map with 3.42 Å resolution. CTF refinement was performed on unbinned particles (box size 300 by 300 pixels) to correct for aberrations, anisotropic magnification and per-particle defocus, with each AFIS position being refined as an independent optics group. Subsequent 3D refinement improved the resolution to 3.14 Å. Bayesian polishing was performed from trained parameters. Polished particles were imported to cryoSPARC and subjected to non-uniform refinement. All processing to this point was carried out with a pixel size of 0.837 Å/pix. The half-maps were combined and the resulting map sharpened using relion_postprocess with the mask from the non-
uniform refinement and an updated calibrated pixel size of 0.831 Å/pixel yielding a final reconstruction with an average resolution of 2.62 Å.

**Model building and refinement**

Templates for all subunits, except FdhF (PDB 1FD0), were generated from homologous structures using the SWISS-MODEL server. The cryo-EM structure of *P. furiosus* MBH (PDB 6CFW) chain M (coverage 179-537) served as a template for the Hyce C-terminus, chain A (coverage 11-488) and D (coverage 469-606) for Hycc N- and C-terminus, chain E (coverage 9-303) for HyccD and chain K (coverage 25-168) for HyccG. The complex 1 structures PDB 6GCS chain G (coverage 8-153) and PDB 5LDX chain I (coverage 14-132) were used as templates for the Hyce N-terminus and Hycc, respectively. For HyccB a template from the arsenate respiratory reductase PDB 6CZ7 chain B (coverage 1-171) was used.

Homology models were initially rigid body fitted into the cryo-EM density of aerobically isolated FHL, followed by manual building and refinement in Coot, from sequence information and secondary structure prediction. The final models contained 2690 residues (aerobic preparation) and 2026 residues (anaerobic preparation without FdhF). Cofactors were built in Coot and respective restraints were added using ReadySet within PHENIX. Phospholipids were added to the model obtained from anaerobically isolated sample, where the density allowed identification (1 CDL, 2 PE, 1 PG, 1 LMNG). In addition, water molecules were modeled in well-defined densities with reasonable hydrogen-bonding distances to the polypeptide. The models were refined against the map using phenix.real_space_refine and validated by MolProbity (for model statistics see Supplementary Table 1). All figures and models and maps were prepared with ChimeraX.

**Hydrogen-uptake assay**

Hydrogenase activity was assessed by hydrogen-dependent benzyl viologen reduction. Anaeroic gas-tight rubber-stoppered cuvettes were filled with anaerobic buffer (20 mM Tris, 200 mM NaCl, 1 mM DTT, pH 17.5) supplemented with 4 mM benzyl viologen (Sigma-Aldrich) and the headspace was saturated with H2. Enzyme was added and benzyl viologen reduction was monitored photometrically at a wavelength of 590 nm. The rate of benzyl viologen reduction was calculated from the initial slopes assuming a molar extinction coefficient of 7400 M⁻¹ cm⁻¹. This rate was converted to mol(H₂) · mol(FHL)⁻¹ · s⁻¹ by dividing by 2.

**Sequence alignments**

Alignments were generated with ClustalW and visualized with the ESPript server.

**Data availability**

Atomic models and maps for anaerobically isolated FHL were deposited in the Protein Data Bank with accession code 7ZOS and Electron Microscopy Data Bank with accession code EMD-14429; for aerobically isolated FHL models and maps are available as 7Z0T and EMD-14430. Composite map EMD-14430 is derived from focused re

**References**

1. Sawers, R. G. & Clark, D. P. Fermentative pyruvate and acetyl-coenzyme A metabolism. Ecosal Plus 1, https://doi.org/10.1128/ecosalplus.3.5.3 (2004).
2. Suppmann, B. & Sawers, G. Isolation and characterization of hypophosphite-resistant mutants of Escherichia coli: identification of the FocA protein, encoded by the pft operon, as a putative formate transporter. Mol. Microbiol. 11, 965–982 (1994).
3. Rossmann, R., Sawers, G. & Böck, A. Mechanism of regulation of the formate-hydrogenlyase pathway by oxygen, nitrate, and pH: definition of the formate regulon. Mol. Microbiol. 5, 2807–2814 (1991).
4. Redwood, M. D., Milkeenko, I. P., Sargent, F. & Macaskie, L. E. Dissecting the roles of Escherichia coli hydrogenases in biogydrogen production. FEMS Microbiol. Lett. 276, 48–55 (2008).
5. Pakes, W. C. C. & Jollyman, W. H. Xl.—The bacterial decomposition of formic acid into carbon dioxide and hydrogen. J. Chem. Soc. Trans. 79, 386–391 (1903).
6. McDowell, J. S. et al. Bacterial formate hydrogenlyase complex. Proc. Natl Acad. Sci. USA 111, E3948–E3956 (2014).
7. Kim, Y. J. et al. Formate-driven growth coupled with H2 production. Nature 467, 352–355 (2010).
8. Lim, J. K., Mayer, F., Kang, S. G. & Müller, V. Energy conservation by oxidation of formate to carbon dioxide and hydrogen via a sodium ion current in a hyperthermophilic archaean. Proc. Natl Acad. Sci. USA 111, 11497–11502 (2014).
9. Martins, M., Mourato, C. & Pereira, I. A. C. Desulfovibrio vulgaris growth coupled to formate-driven H2 production. Environ. Sci. Technol. 49, 14655–14662 (2015).
10. Pinske, C. & Sargent, F. Exploring the directionality of Escherichia coli formate hydrogenlyase: a membrane-bound enzyme capable of fixing carbon dioxide to organic acid. Microbiologyopen 5, 721–737 (2016).
11. Roger, M., Brown, F., Gabrielli, W. & Sargent, F. Efficient hydrogen-dependent carbon dioxide reduction by Escherichia coli. Curr. Biol. 28, 140–145.e2 (2018).
12. Roger, M., Reed, T. C. P. & Sargent, F. Harnessing Escherichia coli for bio-based production of formate under pressurized H2 and CO2 gases. Appl. Environ. Microbiol. 87, e00299–21 (2021).
13. Maeda, T., Sanchez-Torres, V. & Wood, T. K. Escherichia coli hydrogenase 3 is a reversible enzyme possessing hydrogen uptake and synthesis activities. Appl. Microbiol. Biotechnol. 76, 1035–1042 (2007).
14. Vignais, P. M. & Billoud, B. Occurrence, classification, and biological function of hydrogenases: an overview. Chem. Rev. 107, 4206–4272 (2007).
15. Marreiros, B. C., Batista, A. P., Duarte, A. M. S. & Pereira, M. M. A missing link between complex I and group 4 membrane-bound [NiFe] hydrogenases. Biochim. Biophys. Acta 1827, 198–209 (2013).
16. Batista, A. P., Marreiros, B. C. & Pereira, M. M. The antiporter-like subunit constituent of the universal adaptor of complex I, group 4 membrane-bound [NiFe] hydrogenases and related complexes. Biol. Chem. 394, 659–666 (2013).
17. Sapra, R., Bagramyan, K. & Adams, M. W. W. A simple energy-conserving system: proton reduction coupled to proton translocation. Proc. Natl Acad. Sci. USA 100, 7545–7550 (2003).
18. Welte, C., Krätzer, C. & Deppenmeier, U. Involvement of Ech hydrogenase in energy conservation of Methanosarcina mazei. FEBS J. 277, 3396–3403 (2010).
19. Efremov, R. G. & Sazanov, L. A. The coupling mechanism of respiratory complex I - a structural and evolutionary perspective. Biochim. Biophys. Acta 1817, 1785–1795 (2012).
20. Bagramyan, K. A. & Martirosov, S. M. Formation of an ion transport supercomplex in Escherichia coli. An experimental model of direct transduction of energy. FEBS Lett. 246, 149–152 (1989).
21. Hakobyan, M., Sargsyan, H. & Bagramyan, K. Proton translocation coupled to formate oxidation in anaerobically grown fermenting Escherichia coli. Biophys. Chem. 115, 55–61 (2005).
Telleria Marloth, J. & Pinske, C. Susceptibility of the formate hydrogenase reaction to the protonophore CCCP depends on the total hydrogenase composition. Inorg. Chem. 58, 38 (2020).

Friedrich, T., & Weiss, H. Modular evolution of the respiratory NAD-H-ubiquinone oxidoreductase and the origin of its modules. J. Theor. Biol. 187, 529–540 (1997).

Kashani-Poor, N., Zwicker, K., Kerscher, S., & Brandt, U. A central functional role for the 49-kDa subunit within the catalytic core of mitochondrial complex I. J. Biol. Chem. 276, 24082–24087 (2001).

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

Zickermann, V. et al. Mechanistic insight from the crystal structure of mitochondrial complex I. Nature 519, 460–465 (2015).

Yu, H. et al. Structure of an ancient respiratory system. Sci. Adv. 5, eaav1850 (2019).

Volbeda, A., et al. Crystal structure of the nickel-iron hydrogenase from Desulfovibrio gigas. Nature 373, 580–587 (1995).

Videira, A. & Azevedo, J. E. Two nuclear-coded subunits of mitochondrial complex I are similar to different domains of a bacterial formate hydrogenase subunit. Int. J. Biochem. 26, 1391–1393 (1994).

Peters, J. W. et al. [FeFe]- and [NiFe]-hydrogenase diversity, mechanism, and maturation. Biochim. Biophys. Acta 1853, 1350–1369 (2015).

Tai, H., Hirot, S., & Stripp, S. T. Proton transfer mechanisms in bimetallic hydrogenases. Acc. Chem. Res. 54, 232–241 (2021).

Ogata, H., Lubitz, W., & Higuchi, Y. Structure and function of [NiFe] hydrogenases. J. Biochem. 160, 251–258 (2016).

Jussupow, A., Di Luca, A., & Kaila, V. R. I. How cardiolipin modulates the dynamics of respiratory complex I. Sci. Adv. 5, eaav1850 (2019).
65. Agip, A.-N. 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).
66. Schutt, G. J., Boyd, E. S., Peters, J. W. & Adams, M. W. W. The modular respiratory complexes involved in hydrogen and sulfur metabolism by heterotrophic hyperthermophilic archaea and their evolutionary implications. FEMS Microbiol. Rev. 37, 182–203 (2013).
67. Steiner, J. & Szanzov, L. Structure and mechanism of the Mrp complex, an ancient cation/proton antiporter. Elife 9, e59407 (2020).
68. Arias-Cartin, R., Grimaldi, S., Arnow, P., Guigliarelli, B. & Magalon, A. Cardiolipin binding in bacterial respiratory complexes: structural and functional implications. Biochim. Biophys. Acta 1817, 1937–1949 (2012).
69. Paradies, G., Paradies, V., De Benedictis, V., Ruggiero, F. M. & Petrossil, G. Functional role of cardiolipin in mitochondrial bioenergetics. Biochim. Biophys. Acta 1837, 408–417 (2014).
70. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods 14, 331–332 (2017).
71. Rohou, A. & Grigorieff, N. CTFFind4: Fast and accurate defocus estimation from electron micrographs. J. Struct. Biol. 182, 216–221 (2015).
72. Wagner, T. et al. SPHIRE-cryOLO is a fast and accurate fully automated particle picker for cryo-EM. Commun. Biol. 2, 218 (2019).
73. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife 7, e42166 (2018).
74. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoS-PARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods 14, 290–296 (2017).
75. Asarnow, D., Palovcak, E. & Cheng, Y. Asarnow/pyem: UCSF pyem v0.5. Zenodo https://doi.org/10.5281/zenodo.3576630 (2019).
76. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr. D. Struct. Biol. 75, 861–877 (2019).
77. Morado, D. EPU_group_AFIS. GitHub https://github.com/DustinMorado/EPU_group_AFIS (2021).
78. Bepler, T. et al. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. Nat. Methods 16, 1153–1160 (2019).
79. Zivanov, J., Nakane, T. & Scheres, S. H. W. Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. IUCrJ 7, 253–267 (2020).
80. Waterhouse, A. et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 46, W296–W303 (2018).
81. Parey, K. et al. Cryo-EM structure of respiratory complex I at work. Elife 7, e39213 (2018).
82. Zhu, J., Vinothkumar, K. R. & Hirst, J. Structure of mammalian respiratory complex I. Nature 536, 354–358 (2016).
83. Glasser, N. R., Oyala, P. H., Osborne, T. H., Santini, J. M. & Newman, D. K. Structural and mechanistic analysis of the arsenate respiratory reductase provides insight into environmental arsenic transformations. Proc. Natl Acad. Sci. USA 115, E8614–E8623 (2018).
84. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. Sect. D., Biol. Crystallogr. 66, 486–501 (2010).
85. Buchan, D. W. A. & Jones, D. T. The PSIPRED Protein Analysis Workbench: 20 years on. Nucleic Acids Res. 47, W402–W407 (2019).
86. Williams, C. J. et al. MolProbity: more and better reference data for improved all-atom structure validation. Protein Sci. 27, 293–315 (2018).
87. Pettersen, E. F. et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. Protein Sci. 30, 70–82 (2021).
88. Sawers, R. G., Ballantine, S. P. & Boxer, D. H. Differential expression of hydrogenase isoenzymes in Escherichia coli K-12, evidence for a third isoenzyme. J. Bacteriol. 164, 1324–1331 (1985).
89. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7, 539 (2011).
90. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 42, W320–W324 (2014).
91. Tian, W., Chen, C., Lei, X., Zhao, J. & Liang, J. CASTp 3.0: computed atlas of surface topography of proteins. Nucleic Acids Res. 46, W363–W367 (2018).

Acknowledgements
We thank the Central Electron Microscopy Facility of the Max Planck Institute of Biophysics for providing cryo-EM infrastructure and technical support, and Werner Kühlerbrandt and the department of Structural Biology for support and helpful discussions. We are grateful to Frank Sargent for providing E. coli strain MG9596 and for helpful discussions. We thank Stella Vitt and Rita Zimmermann for sharing their valuable expertise in anaerobic work. We thank Volker Zickermann and Jan vonck for critical reading and comments on the manuscript. This work was funded by the Max Planck Society (to B.J.M.) and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project-ID 450648163—CRC 1507/1 2022 to B.J.M.

Author contributions
R.S. purified protein and carried out activity assays, cryo-EM sample preparation, data acquisition, data processing, data analysis and interpretation, atomic model building, preparation of figures, manuscript preparation and revision, G.H. carried out anaerobic fermentation and anaerobic cell harvesting, and revised the manuscript. J.H. provided fermentation infrastructure and revised the manuscript. B.J.M. supervised the project and carried out cryo-EM sample preparation, data acquisition, data analysis and interpretation, manuscript preparation and revision.

Funding
Open Access funding enabled and organized by Projekt DEAL.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-32831-x.

Correspondence and requests for materials should be addressed to Bonnie J. Murphy.

Peer review information Nature Communications thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
