








ARTICLE

<https://doi.org/10.1038/s41467-020-14347-4>

OPEN

# Redox-coupled proton pumping drives carbon concentration in the photosynthetic complex I

Jan M. Schuller <sup>1\*</sup>, Patricia Saura <sup>2,3,7</sup>, Jacqueline Thiemann <sup>4,7</sup>, Sandra K. Schuller <sup>1</sup>, Ana P. Gamiz-Hernandez<sup>2,3</sup>, Genji Kurisu <sup>5,6</sup>, Marc M. Nowaczyk <sup>4\*</sup> & Ville R.I. Kaila <sup>2,3\*</sup>

Photosynthetic organisms capture light energy to drive their energy metabolism, and employ the chemical reducing power to convert carbon dioxide (CO<sub>2</sub>) into organic molecules. Photorespiration, however, significantly reduces the photosynthetic yields. To survive under low CO<sub>2</sub> concentrations, cyanobacteria evolved unique carbon-concentration mechanisms that enhance the efficiency of photosynthetic CO<sub>2</sub> fixation, for which the molecular principles have remained unknown. We show here how modular adaptations enabled the cyanobacterial photosynthetic complex I to concentrate CO<sub>2</sub> using a redox-driven proton-pumping machinery. Our cryo-electron microscopy structure at 3.2 Å resolution shows a catalytic carbonic anhydrase module that harbours a Zn<sup>2+</sup> active site, with connectivity to proton-pumping subunits that are activated by electron transfer from photosystem I. Our findings illustrate molecular principles in the photosynthetic complex I machinery that enabled cyanobacteria to survive in drastically changing CO<sub>2</sub> conditions.

<sup>1</sup>Department of Structural Cell Biology, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany. <sup>2</sup>Department of Biochemistry and Biophysics, Stockholm University, SE-106 91 Stockholm, Sweden. <sup>3</sup>Center of Integrated Protein Science at the Department of Chemistry, Technical University of Munich, Garching, Germany. <sup>4</sup>Plant Biochemistry, Faculty of Biology and Biotechnology, Ruhr University Bochum, 44780 Bochum, Germany. <sup>5</sup>Institute for Protein Research, Osaka University, Suita, Osaka 565-0871, Japan. <sup>6</sup>Department of Macromolecular Science, Graduate School of Science, Osaka University, Toyonaka 560-0043, Japan. <sup>7</sup>These authors contributed equally: Patricia Saura, Jacqueline Thiemann \*email: [jansch@biochem.mpg.de](mailto:jansch@biochem.mpg.de); [marc.m.nowaczyk@rub.de](mailto:marc.m.nowaczyk@rub.de); [ville.kaila@dbb.su.se](mailto:ville.kaila@dbb.su.se)

Cyanobacteria evolved around 2.7 billion years ago with the ability to oxidise water into dioxygen ( $O_2$ ) using the energy captured from sunlight<sup>1</sup>. The released  $O_2$  powers respiratory chains in aerobic life, and the electrons extracted from the water oxidation are used for synthesis of organic molecules<sup>2</sup>. The light-driven water splitting catalysed by photosystem II (PSII), reduces plastoquinone (PQ) and establishes an electrochemical proton gradient across the thylakoid membrane that subsequently drives synthesis of adenosine triphosphate (ATP)<sup>3,4</sup>. During linear electron flow (LEF), the electrons are transferred to photosystem I (PSI), providing the reducing power for photosynthetic  $CO_2$  fixation that consumes nicotinamide adenine dinucleotide phosphate (NADPH) and drives the synthesis of complex organic compounds from inorganic carbon ( $C_i$ )<sup>2</sup>. Photosynthetic organisms also employ cyclic electron flow (CEF) around PSI to increase the ATP/NADPH ratio that powers the  $CO_2$  fixation.

$CO_2$  concentrations drastically changed during the last 2.7 billion years, from >0.5% (5000 ppm) to today's levels around 0.041% (410 ppm)<sup>5</sup>. To survive, cyanobacteria evolved carbon-concentrating mechanisms (CCMs) that enhance the efficiency of the photosynthetic  $CO_2$  fixation process<sup>6,7</sup>. In many organisms, hydration of the gaseous  $CO_2$  to the soluble bicarbonate ( $HCO_3^-$ ) is catalysed by carbonic anhydrase (CA), one of the fastest known enzymes with a  $k_{cat} \sim 10^6 s^{-1}$ <sup>8</sup>. However, cyanobacteria lack genes for canonical cytoplasmic CAs, and if artificially expressed in the cytoplasm, the bacteria do not survive in low  $CO_2$  concentrations < 20 ppm<sup>7</sup>. Cyanobacteria express instead the inducible NDH-1MS (NDH-1<sub>3</sub>) and constitutive NDH-1MS' (NDH-1<sub>4</sub>) photosynthetic complex I isoforms that convert  $CO_2$  into  $HCO_3^-$  by kinetically shifting the reaction equilibrium towards bicarbonate, against high cytoplasmic  $HCO_3^-$  concentrations<sup>7,9</sup>. The  $HCO_3^-$  subsequently diffuses into the carboxysome micro-compartments, where it is converted by a carboxysomal CA to  $CO_2$ <sup>7</sup>, which further carboxylates ribulose-1,5-bisphosphate (RuBP) into carbohydrates by the action of RuBisCO (Ribulose-1,5-bisphosphate carboxylase/oxygenase)<sup>10</sup>. This CCM of the photosynthetic complex I, prevents  $CO_2$  to diffuse out of the cell by concentrating the  $C_i$  for RuBisCO, providing a basis for the efficient carbon fixation that is hampered during photorespiration<sup>7</sup>.

To determine the molecular architecture of NDH-1MS (NDH-1<sub>3</sub>), we isolate the enzyme from the cyanobacterium *Thermosynechococcus elongatus*, solved its molecular structure at 3.2 Å resolution using cryo-EM (Fig. 1, Supplementary Fig. 1, Supplementary Table 1, and Supplementary Movie 1), and probe its molecular mechanism by classical and quantum mechanical simulations.

## Results

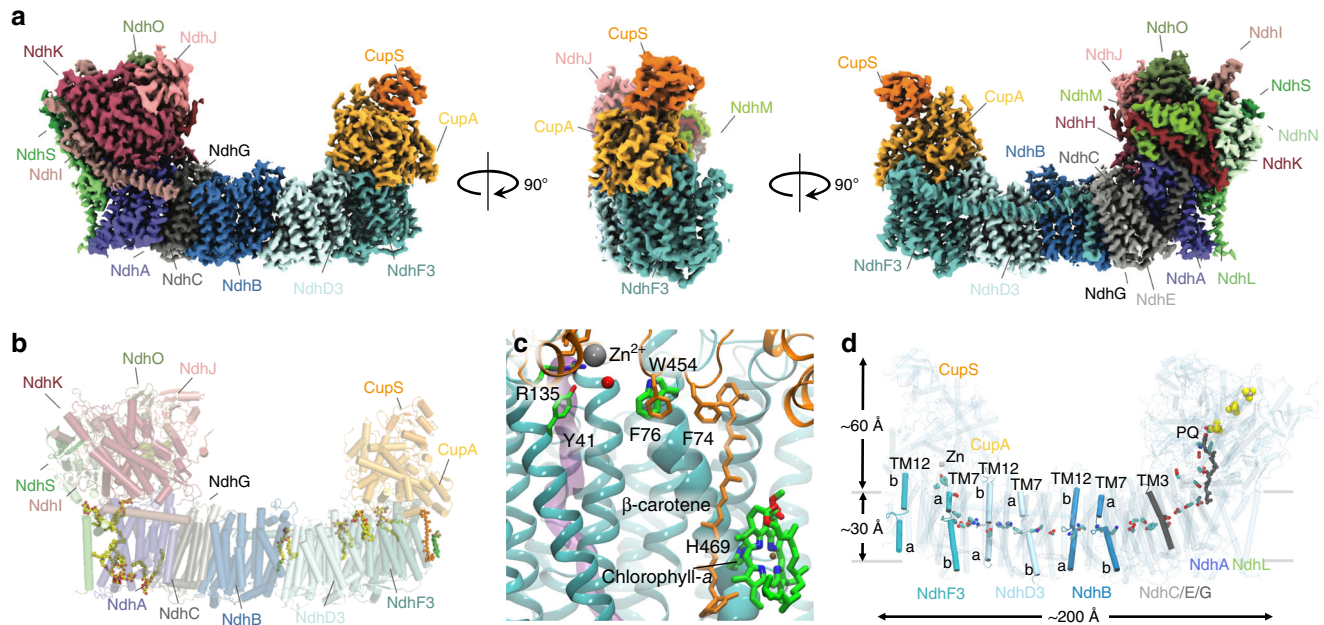
**Architecture of the photosynthetic complex I.** The 0.5 MDa complex has an overall U-shape with 19 isolated subunits (Supplementary Fig. 2 and Supplementary Table 2). The structure of NdhV could not be resolved. The core structure is highly conserved across the complex I superfamily<sup>11–14</sup>, but in contrast to the respiratory enzyme, the photosynthetic complex lacks the N-module that accepts electrons from nicotinamide adenine dinucleotide (NADH) (Fig. 1a–d). Instead, the electrons directly enter a chain of three iron-sulphur (FeS) centres in the ferredoxin (Fd)-binding domain (Fig. 1d), similar to the recently characterised NDH-1L type photosynthetic complex I and membrane-bound hydrogenases<sup>15–18</sup>. The PQ-binding site is located ca. 20 Å above the membrane plane (Fig. 1d and Supplementary Fig. 3a–c, f), and the modelled PQ tail extends into the lipid membrane in the vicinity of the NdhL subunit, with experimentally resolved lipids stabilising the PQ entry gate (Supplementary Fig. 3b, f).

The modular membrane domain extends up to 200 Å away from the PQ reduction site, and it comprises the antiporter-like subunits NdhA, NdhB, NdhD3, and NdhF3, as well as the smaller transmembrane (TM) subunits NdhC/E/G/L. These subunits contain a chain of buried charged residues that establish central elements of the proton-pumping machinery<sup>11–17,19,20</sup>, with the isoform-specific NdhD3 and NdhF3 subunits at the terminal end of the enzyme (Fig. 1d, see below).

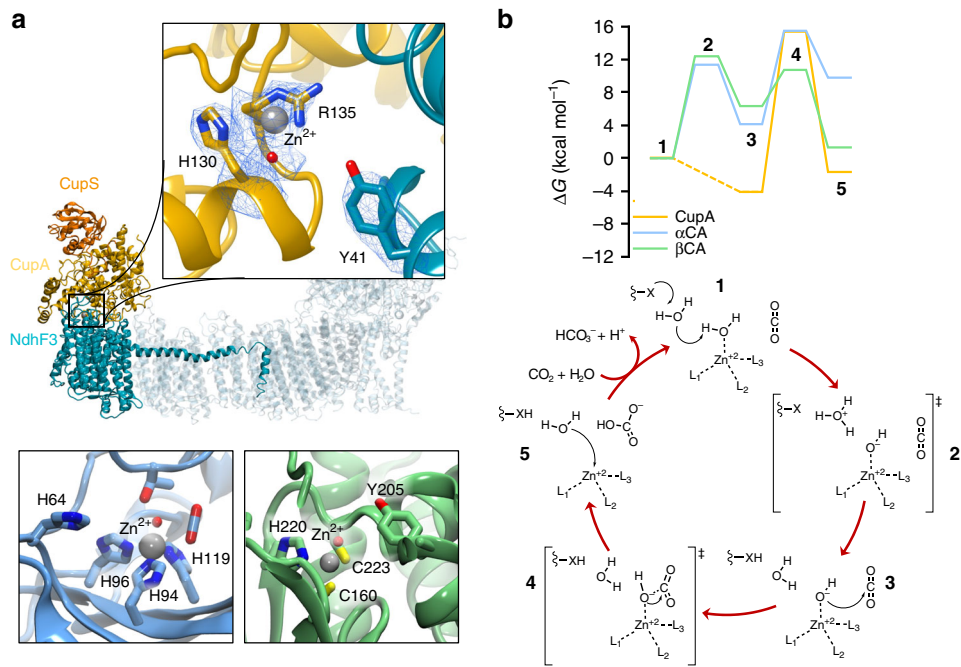
**Structure and function of the  $CO_2$  concentrating module.** The  $CO_2$ -concentrating CupA/S module ( $CO_2$  uptake, Cup/ $CO_2$  hydration protein, ChpY) is located on the cytoplasmic side of the membrane, extending up to 60 Å above the membrane plane (Figs. 1d and 2a and Supplementary Fig. 3d, e). The catalytic site responsible for the  $CO_2$  chemistry in the 50 kDa CupA protein electrostatically binds NdhF3 (Fig. 1 and Supplementary Fig. 4a). CupA has a fold, comprising  $\alpha$ -helices, which drastically differs from known carbonic anhydrases (CAs) in which the catalytic zinc ( $Zn^{2+}$ ) ion binds to histidine ( $\alpha$ CA family) or histidine/cysteine residues ( $\beta$ CA family) at the interface of  $\beta$ -sheets (Fig. 2a)<sup>8</sup>. CupA is capped by the 16 kDa  $\alpha$ -helical CupS subunit that binds two helical lobes of the former, and could help in stabilising the CupA-NdhF3 interaction (Supplementary Fig. 4a). CupS undergoes a conformational change from its solution structure upon binding to CupA in which an  $\alpha$ -helical lid ( $\alpha_2$ ,  $\alpha_3$ ) closes upon a  $\beta$ -sheet (Supplementary Fig. 4b)<sup>21</sup>. The active site of CupA shows a strong electron density that we assign to a  $Zn^{2+}$  ion (Fig. 2a), as also supported by inductively-coupled plasma optical emission spectrometry (ICP-OES), which indicates a protein-to-zinc ratio of  $\sim 0.7$  in the NDH-1MS sample (Supplementary Fig. 2e). The  $Zn^{2+}$  ion is coordinated by His130, a putative water/hydroxide ligand or carbonate, and Arg135 (Fig. 2a). Our mass spectrometric (MS) measurements reveal no indication for citrullination of the arginine (Supplementary Fig. 2d, e), and our electrostatic calculations suggest that the residue is neutral with  $pK_a \ll 7$ , due to its proximity to the  $Zn^{2+}$  ion (Fig. 2a and Supplementary Fig. 5). Such  $Zn^{2+}$ -Arg coordination is rare, but similar structures are also found in a human genetic variant of carbonic anhydrase I ( $\alpha$ CA type I)<sup>22</sup>.

The  $Zn^{2+}$ -site is located at the interface of two 12 Å helices, and CupA forms further contacts with NdhF3 and several lipid molecules (Fig. 3a, b and Supplementary Fig. 3d, e). The putative  $CO_2$  binding pocket is completed by residues from NdhF3: the  $Zn^{2+}$ - $H_2O/OH^-$  is stabilised by Arg37 that forms an ion-pair with Glu114 and Tyr41 of NdhF3. A non-polar tunnel starting from NdhF3 that leads to the  $Zn^{2+}$ -site via this Arg37/Glu114/Tyr41 network could function as a  $CO_2$  conduction channel from the luminal side across the membrane (Fig. 3c). Remarkably, NdhF3 binds a chlorophyll *a*/ $\beta$ -carotene motif that connects via His469 of NdhF3 in a TM-helix to the CupA site (Fig. 1c and Supplementary Fig. 3d). Despite lacking experimental evidence of its functionality, we speculate that the motif could be involved in light-triggered regulation of the CA activity and/or to provide structural stability. A similar Chl *a*/ $\beta$ -carotene motif with unknown function is also found in cytochrome *b\_6/f*<sup>23</sup>.

**Mechanism of the  $CO_2$  hydration reaction.** To probe the catalytic properties of the CupA module, we performed quantum chemical density functional theory (DFT) calculations on the  $CO_2$  hydration reaction and compared this to the reaction energetics in  $\alpha$ CA and  $\beta$ CA (Fig. 2b and Supplementary Fig. 5, see Methods). The  $CO_2$  hydration is initiated by proton transfer (pT) from the Zn-bound water molecule followed by a nucleophilic attack of the hydroxide on  $CO_2$ . In  $\alpha$ CA ( $\beta$ CA), the rate-limiting reaction barrier of 11 (12) kcal mol<sup>-1</sup> is connected with pT to His64

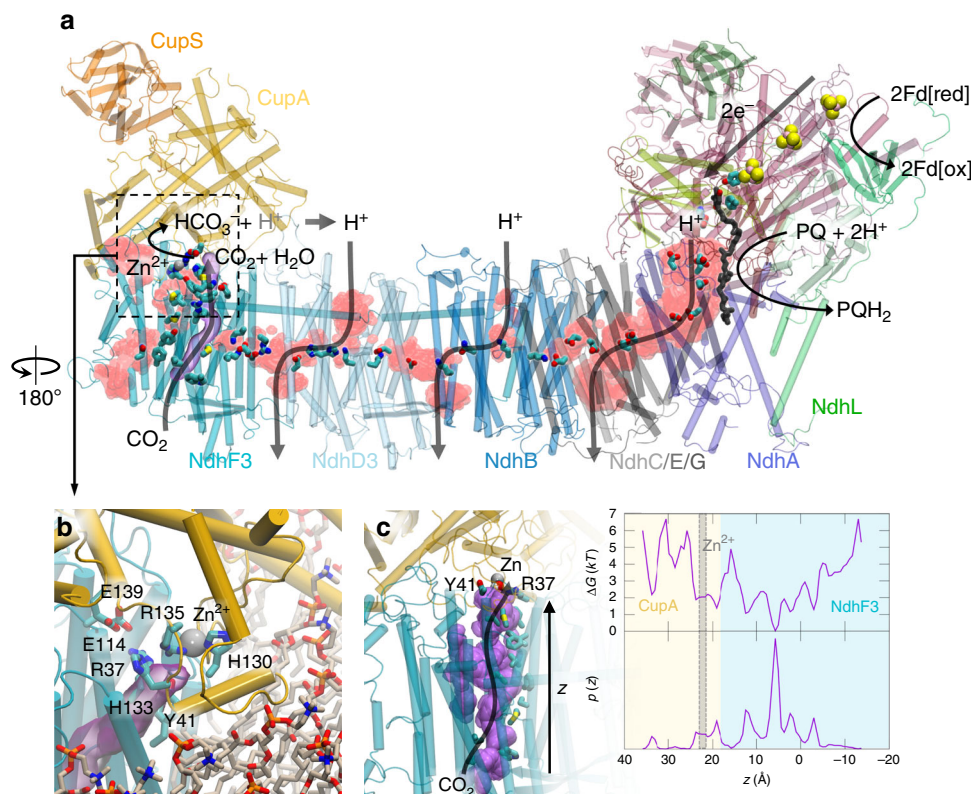


**Fig. 1 Structure of the carbon-concentrating photosynthetic complex I.** **a** The electron density map of NDH-1MS (NDH-1<sub>3</sub>) shown from a back view (left), side view (middle), and front view (right). **b** Experimentally refined lipid molecules and cofactors in the hydrophobic domain of NDH-1MS. **c** The structure of the Chl *a*/ $\beta$ -carotene motif of NdhF3 that connects to the active site of CupA. See Supplementary Fig. 3 for example densities. **d** Chain of charged elements transmitting the redox-signal into proton pumping and CO<sub>2</sub> uptake. The PQ was computationally modelled into the experimental structure.



**Fig. 2 Structural comparison and mechanism of carbon concentration in CupA and canonical carbonic anhydrases.** **a** The active site structure of the CO<sub>2</sub> concentrating CupA subunit, showing the density of Zn-coordinating residues (5.5 sigma value, contour level 2, top),  $\alpha$ -carbonic anhydrase ( $\alpha$ CA, PDB ID: 5YUI, bottom left), and  $\beta$ -carbonic anhydrase ( $\beta$ CA, PDB ID: 1EKJ, bottom right). The Zn-bound density has been modelled as a water ligand, although the character of the ligand cannot be unambiguously assigned based on the map. **b** Reaction mechanism and free energy profiles for the CO<sub>2</sub> hydration process based on quantum chemical DFT models in CupA,  $\alpha$ CA, and  $\beta$ CA with Tyr41, His64, and Tyr205 as proton acceptors, respectively. Free energies are reported at the B3LYP-D3/def2-TZVP/def2-SVP/ $\epsilon = 4$  theory level (see Methods). Ligands L<sub>1/2/3</sub> = His130/Arg135/H<sub>2</sub>O, L<sub>4</sub> = H<sub>2</sub>O/OH<sup>-</sup>, X = Tyr41 in CupA; L<sub>1/2/3</sub> = His94/96/119, L<sub>4</sub> = H<sub>2</sub>O/OH<sup>-</sup>, X = His64 in  $\alpha$ CA; L<sub>1/2/3</sub> = Cys160/223/His220, L<sub>4</sub> = H<sub>2</sub>O/OH<sup>-</sup>, X = Tyr205 in  $\beta$ CA. The CupA reaction takes place via two water molecules, modelled based on MD simulations (Supplementary Fig. 5). See Supplementary Fig. 5 for further details.



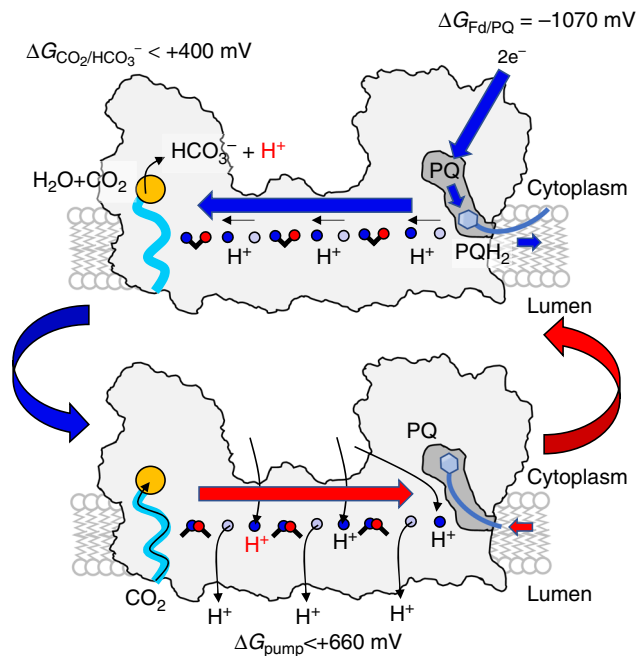


**Fig. 3** Proton pathways and putative CO<sub>2</sub> channel in the photosynthetic complex I. **a** Location of proton pathways formed by water molecules (in red) during 250 ns atomistic molecular dynamics simulations. The figure shows the overall water density during MD. **b** Closeup of the active site of CupA at the NdhF3 interface (in cyan) and the lipid membrane from MD simulations. The putative CO<sub>2</sub> channel (in purple), visualised based on CAVER analysis (see Methods), is surrounded by hydrophobic and bulky residues, and connects to the Zn<sup>2+</sup>-binding site by conserved Tyr41, Arg37, and Glu114 of the NdhF3 subunit (see Supplementary Fig. 6g for details). **c** Diffusion pathway of CO<sub>2</sub> molecules from the luminal side to CupA during MD simulations. Left: Dynamics of CO<sub>2</sub> in its putative gas channel in NdhF3. The average of the CO<sub>2</sub> position sampled during dynamics is represented as a purple surface. Right: Probability distribution and resulting potential of mean force ( $pmf$  in units of  $kT$ ),  $G(z) = -RT \log(p(z))$ , of CO<sub>2</sub> inside the channel along the membrane axis ( $z$ -axis) obtained from 200 ns MD simulation. Owing to sampling gaps, the  $pmf$  is subjected to large errors.

(Tyr205), which compares well with the experimentally-observed barriers of ca. 10–12 kcal mol<sup>-1</sup> (Fig. 2b)<sup>8</sup>. In CupA, proton transfer from Zn-bound water to Tyr41 is slightly exergonic in our DFT models, whereas the nucleophilic attack of the Zn-bound OH<sup>-</sup> to the CO<sub>2</sub> has a barrier of ca. 15 kcal mol<sup>-1</sup>, predicting that catalysis takes place in the millisecond timescale. The involvement of Tyr41, similarly to Tyr205 in  $\beta$ CA<sup>24</sup>, is also supported by QM/MM models (Supplementary Fig. 5e), as well as by MD simulations and  $pK_a$  calculations (Supplementary Fig. 5f). These findings suggest that our resolved structure is catalytically efficient, with kinetic barriers similar to the canonical CAs, despite its unique molecular architecture. CupA becomes well-hydrated during the MD simulations, establishing protonic connectivity with the NdhD3 subunit (Fig. 3a and Supplementary Figs. 5d and 6c), and with global dynamics inferred from the local resolution of the cryo-EM map, closely resembling the motion extracted from the simulations (Supplementary Fig. 7). The proton channels are established across the membrane around charged residues in the broken helices TM7 and TM12 of the antiporter-like subunits NdhB, NdhD3<sup>19,20</sup>, and also in NdhA/C/E/G (Figs. 1d and 3, and Supplementary Fig. 6a–d, f) at locations where conformational changes were recently observed in the mammalian enzyme (Supplementary Fig. 6e)<sup>12</sup>. The structural architecture thus supports that the long-range protonation signal could be triggered by dissociation of conserved ion-pairs in the antiporter-like subunits that leads to lateral proton transfer in the proton channels by coupled conformational and hydration changes (Fig. 4)<sup>16,19,20</sup>.

## Discussion

The NdhF3 comprises conserved ion-pairs at the interface to the NdhD3 subunit, similar as in the other antiporter-like subunits. However, charged elements in the lateral proton channel in NdhF3 are replaced by non-polar residues that could be employed to channel CO<sub>2</sub> into CupA. Indeed, we observe that CO<sub>2</sub> effectively diffuses along this channel to the active site of the CupA subunit in MD simulations (Fig. 3c and Supplementary Fig. 6g). Although the solubility of CO<sub>2</sub> is similar in lipid membranes and water<sup>25</sup>, the gas tunnel could channel CO<sub>2</sub> into the active site of the CupA subunit, and prevent its escape under low CO<sub>2</sub> and high bicarbonate concentrations, arising from the substrate channelling to RuBisCO and bicarbonate pumps in the plasma membrane<sup>7,26</sup>. The thermodynamically unfavourable CO<sub>2</sub> hydration reaction is thus kinetically gated in the bicarbonate direction when the protons released in the CO<sub>2</sub> hydration process are pumped across the membrane by the antiporter-like subunits (Fig. 4). The driving force for this is achieved by the ca. 1070 mV redox span transduced from the PQ oxido-reduction chemistry ( $E_m(\text{Fd}) \approx -420$  mV<sup>27</sup>;  $E_m(\text{PQ}) \approx +117$  mV<sup>28</sup>), that is transmitted to the terminal edge of the membrane domain (Fig. 4)<sup>19,20</sup>. The carbon concentration thus couples to generation of proton motive force ( $pmf$ ) across the membrane, that in turn, is transduced in ATP synthesis and active transport<sup>3</sup>. With a high  $pmf$  of up to 220 mV<sup>29</sup>, the NDH-1MS is expected to thermodynamically drive the hydration of CO<sub>2</sub> (at 410 ppm; [CO<sub>2</sub>] = 14  $\mu$ M; [HCO<sub>3</sub><sup>-</sup>] = 30 mM<sup>26</sup>; pH = 7), and couple this to pumping of at least three



**Fig. 4 Putative mechanism of redox-driven CO<sub>2</sub> concentration in the photosynthetic complex I.** Top: PQ reduction to PQH<sub>2</sub> leads to dissociation of the quinol to a membrane-bound PQ-binding site<sup>71</sup> that triggers sequential opening of ion-pairs in NdhB, NdhD3, and NdhF3, by coupled hydration, conformational, and protonation changes, protonating terminal proton loading sites (blue forward arrow)<sup>19, 20</sup>. Bottom: Closing of ion-pairs in a sequential electrostatic backward pulse (red backward arrow) releases the protons to the luminal side of the membrane, and allows for re-protonation of the antiporter-like subunits by protons released from the CO<sub>2</sub> hydration reaction in the active site of CupA (orange circle). Horizontal proton transfer reactions within each antiporter-like subunit are shown by small horizontal black arrows, and PQH<sub>2</sub> (PQ) diffusion out (in) is indicated by small thick blue (red) arrows. CO<sub>2</sub> is taken up by the putative gas channel (in light blue) that is expected to be open depending on the ion-pair conformation in NdhF3 (arrow along light blue channel).

protons across the membrane, e.g., by NdhD3, NdhB, and NdhA/C/E/G<sup>11–14, 19, 20</sup>. However, the free energy available could allow the pump to operate also under considerably lower CO<sub>2</sub> concentrations and pH. Under high *pmf* and a reduced quinone pool, the respiratory complex I operates in reverse direction by consuming the proton gradient to oxidize quinol<sup>30</sup>, and due to conserved functional elements, a similar operation mode is also expected for NDH-1MS. During such putative backward operation mode, coupled protonation and/or conformational changes at the NdhF3/NdhD3 interface could close the gas channel similar to conformational changes observed in the bacterial complex I<sup>19, 31</sup>. Such changes might decouple the pump to avoid the back-reaction of HCO<sub>3</sub><sup>-</sup> to CO<sub>2</sub>, and the diffusion of the latter out of the cell.

We have demonstrated here how unique molecular adaptation enabled the photosynthetic complex I to concentrate CO<sub>2</sub> by redox-driven proton-pumping modules and powered by cyclic electron flow around photosystem I. Our findings show how combination of functional modules from different energy transduction machineries<sup>32</sup> allowed primordial organisms to generate functionality and harness energy under changing environmental conditions.

## Methods

**Construction of the *T. elongatus* mutant and purification of the NDH-1MS complex.** The C-terminal tagged CupS-TwinStrep-tag mutant of *T. elongatus* was generated by homologous recombination with plasmid DNA that contains the

modified cupS gene and a kanamycin resistance cassette for selection of the mutant allele (see Supplementary Table 4). Cells of the fully segregated mutant were grown in liquid BG-11 medium and expression of the NDH-1MS complex was induced by low CO<sub>2</sub> growth conditions. Preparation of the thylakoid membranes and isolation of the tagged complex were performed as described earlier<sup>15, 33</sup> with minor modifications. See SI Materials and Methods for details.

**Cryo-EM grid preparation and data collection.** To prepare cryo-EM grids, four microliter of NDH-1MS at 5 mg mL<sup>-1</sup> were applied to R2/1 grids (Quantifoil) that were glow-discharged for 20 s immediately before use. The sample was incubated 30 s at 100% humidity and 4 °C before blotting for 3.5 s with blotforce 5 and then plunge-frozen into liquid ethane/propane mix cooled by liquid nitrogen using a Vitrobot Mark IV (FEI). Data was acquired on a Titan Krios electron microscope (ThermoFisher, FEI) operated at 300 kV, equipped with a K2 Summit direct electron detector (Gatan) and a GIF quantum energy filter (20 eV) (Gatan). Movies were recorded in counting mode at a pixel size of 1.35 Å per pixel using a cumulative dose of 40.20 e<sup>-</sup>/Å<sup>2</sup> and 50 frames. Data acquisition was performed using SerialEM<sup>34</sup> with four exposures per hole with a target defocus range of 0.5 to 3.0 μm.

**Image processing.** The dose-fractionated movies were gain normalised, aligned and dose-weighted using the motion correction algorithm (Motion Cor2)<sup>35</sup> implemented in RELION 3.0<sup>34, 36</sup>. GCTF<sup>37</sup> was used to estimate the defocus values, and particles were selected using Gautomatch. All subsequent processing steps were carried out in RELION 3.0<sup>34</sup>. Following two-dimensional classification, an ab initio model was generated using stochastic gradient descent (SGD). The entire data set (526,925 particles) was subjected to three-dimensional (3D) classification into five classes using the previously determined 60 Å low-pass-filtered ab initio model. A single class consisting of the majority of particles (51.7%) resembled the complex structure. The particles were selected and subsequently sub-classified into 3 classes, using a finer angular sampling (3.75°). The best aligning class consisted of 170,151 particles (62.4%), which were subjected to 3D refinement, yielding an overall resolution of 3.9 Å. The particles were further polished using Bayesian-polishing and the CTF-values were refined on a per particle basis, improving the resolution of the reconstruction to 3.2 Å. The temperature factor (−89.9 Å<sup>2</sup>) and the resolution of the map were estimated by applying a soft mask around the protein density in the post-processing routine, using the gold standard Fourier shell correlation (FSC) = 0.143 criterion. Directional FSC curves and map anisotropy were assessed using the 3DFSC server<sup>38</sup>.

**Model building and validation.** The models for the conserved subunits were taken from the previously published NDH-1L complex (PDB ID: 6HUM)<sup>15</sup>. For the NdhD3 and NdhF3 paralog subunits homology models were generated with the Phyre2 server and fitted as rigid bodies with UCSF Chimera<sup>39</sup>. They were subsequently manually adjusted and rebuilt using Coot<sup>40</sup>. The carbonic anhydrase CupA and its binding partner CupS was built de novo using Coot. The model of NDH-1MS was refined against the cryo-EM map using the phenix.real\_space\_refine routine in the PHENIX software package<sup>41</sup>. The statistical quality of the final model was assessed using MOLPROBITY<sup>42</sup> and EMRinger<sup>43</sup>. Figures were prepared using PyMOL<sup>44</sup> or UCSF Chimera X<sup>45</sup>.

**DFT calculations.** DFT models of the CupA/NdhF3 active site were built based on the cryo-EM atomic coordinates of NDH-1MS. The model comprises the Zn<sup>2+</sup> ion and its water/OH<sup>-</sup> ligands, Phe70, Lys78, Tyr79, His130, Ile131, His133, Arg135, Glu139, and Glu142 of CupA subunit, and Tyr41, Arg37, and Glu114 of the NdhF3 subunit, in addition to two water molecules and the CO<sub>2</sub>/HCO<sub>3</sub><sup>-</sup> substrate. Water molecules between Zn and Tyr41 were modelled based on MD simulations (Supplementary Fig. 5). The model comprises 185 atoms. DFT models of α-carbonic anhydrase were built based on type II human αCA (PDB ID: 5YUI)<sup>46</sup>, comprising the Zn<sup>2+</sup> ion, and the water/OH<sup>-</sup> group, His94 His96, His119, His64, Asn67, Gln92, Glu106, Thr199, Thr200, Val121, Val143, Trp209, and eight water molecules and a CO<sub>2</sub> molecule solved in the crystal structure. The model comprises 155 atoms. DFT models of βCA were built based on the PDB ID: 1EKJ<sup>47</sup>. The model comprises the Zn<sup>2+</sup> ion, Cys160, Cys223, His220, the water/OH<sup>-</sup> ligand, Gln151, Asp162, Arg164, Val165, Val184, Asn186, the backbone of Gly224, Tyr205, in addition to four water molecules and the CO<sub>2</sub>. The model comprises 132 atoms. Protein residues were cut at the Ca-Cβ bond and saturated with hydrogen atoms. The Cβ atoms were kept fixed during geometry optimisations that were performed at the B3LYP-D3/def2-SVP/def2-TZVP(Zn<sup>2+</sup>) level, and single point energies were evaluated at the B3LYP-D3/def2-TZVP level<sup>48–51</sup>. Solvation effects were treated using a polarisable dielectric medium with ε = 4 to model the protein environment<sup>52</sup>. Free energies were computed using the free module of TURBOMOLE based on electronic and zero-point energies (ZPE), and enthalpic (ΔH) and entropic (TΔS) effects, estimated at the B3LYP-D3/def2-SVP level by calculating the molecular Hessian. The free energy estimates do not consider dynamical sampling effects. Reaction pathways and transition states were optimised using a chain-of-states method<sup>53, 54</sup>. All QM calculations were performed with TURBOMOLE versions 6.6–7.3<sup>55</sup>.

**QM/MM calculations.** Hybrid QM/MM calculations were performed based on our NDH-1MS structure. The models comprise the CupA and NdhF subunits, including

the Zn-binding site, and water/lipid/ions in their 12-Å surroundings. The QM region comprises 135 atoms, including the Zn<sup>2+</sup>-OH<sup>-</sup> cofactor, Tyr79, His130, His133, Arg135, Glu139 of CupA subunit, and Arg37 and Tyr41 of the NdhF3 subunit. The QM/MM boundary was described by a link-atoms approach. The model was structure optimised at the QM/MM level using the Adopted Basis Newton-Raphson (ABNR) algorithm, and allowing a 10 Å-sphere within the QM region to relax, followed by QM/MM dynamics at  $T = 310$  K using a 1 fs integration timestep. The QM region was described at the B3LYP-D3/def2-SVP(C,H,N,O)/def2-TZVP(Zn<sup>2+</sup>) level of theory, and the MM region was described by the CHARMM36 force field<sup>56</sup>. A minimum energy pathway scan between the Zn-H<sub>2</sub>O and Tyr41 was performed along a reaction coordinate modelled as linear combination of all proton donor and acceptor distances ( $R = r_1 - r_2 + r_3 - r_4$ , see Supplementary Fig. 5d). Our reported DFT models predict similar energetics for carbonic anhydrases as those reported in previous studies, cf.<sup>57</sup> and refs. therein and further QM/MM models of these systems were therefore not considered. All QM/MM calculations were performed using an in-house version of the CHARMM/TURBOMOLE interface<sup>58</sup>.

**Classical molecular dynamics simulations.** Classical molecular dynamics (MD) simulations were performed based on the cryo-EM structure of NDH-IMS. Missing loops in the antiporter-like and NdhS subunits were modelled with MODELLER9.19<sup>59</sup>. The cryo-EM atomic coordinates were relaxed to the electron density map using molecular dynamics flexible fitting (MDFF)<sup>60</sup>. The MDFF relaxed model was further embedded in a POPC membrane and solvated with TIP3P water molecules and 100 mM NaCl concentration. Plastoquinone was modelled based on previous work<sup>16</sup>. The total system comprises ca. 580,000 atoms. MD simulations were performed at constant  $T = 310$  K and  $p = 1$  bar in an NPT ensemble. The CHARMM36 force field<sup>56</sup> was used to model protein/lipid/water-ion atoms. Parameters for PQ, FeS centres, and the Zn-first coordination sphere were derived from DFT calculations. Long-range electrostatic interactions were described by the particle mesh Ewald (PME) approach. The system was gradually relaxed during the 2.5 ns with harmonic restraints, followed by 20 ns equilibration without restraints. An initial 1 fs integration timestep was used during the equilibration process, and 2 fs timestep during the remaining 250 ns production steps. MD simulations of the CO<sub>2</sub> diffusion process were performed based on 20 independent trajectories each 10 ns with the CO<sub>2</sub> molecule positioned along the putative cavity in the NdhF3 subunit, that was identified with CAVER<sup>61</sup>. All classical MD simulations were performed with NAMD<sup>62</sup>, and VMD<sup>63</sup> and UCSF Chimera<sup>39</sup> were used for analysis.

**Poisson-Boltzmann continuum electrostatics calculations.** pK<sub>a</sub> values of titratable residues in CupA/S/NdhF3 were estimated using the Adaptive Poisson-Boltzmann Solver (APBS)<sup>64</sup> and by Monte Carlo (MC) sampling of 2<sup>N</sup> protonation states with Karlsberg<sup>65,66</sup>. PBE/MC calculations can provide accurate estimation of pK<sub>a</sub> values in complex biochemical systems<sup>67,68</sup>, whereas constant pH-MD simulations that could further enhance the accuracy of the predictions<sup>69,70</sup> are outside the scope of the present work due to the large size and complex surroundings of NDH-IMS. The protein was described using partial charges, embedded in an inhomogeneous dielectric medium with a dielectric constant of 4 inside the protein and 80 for water. The interface between the protein and solvent was calculated by the molecular surface routine, as implemented in APBS, using a solvent probe radius of 1.4 Å, and modelling an implicit ionic strength of 100 mM potassium chloride. The pK<sub>a</sub> values were computed as a difference of electrostatic free energy shifts between a model compound in water and the model compound in the protein for 80 structures obtained from the MDFF relaxation.

## Data availability

Data supporting the findings of this manuscript are available from the corresponding authors upon reasonable request. A reporting summary for this Article is available as a Supplementary Information file. The source data underlying Figs. 2b, 3c, 4c–e and Supplementary Figs. 2a, b, 3c, 5c, d are provided as a Source Data file. EMDB accession codes: EMD-10513, PDB ID: 6TJV.

Received: 13 October 2019; Accepted: 22 December 2019;

Published online: 24 January 2020

## References

- Cardona, T. Thinking twice about the evolution of photosynthesis. *Open Biol.* **9**, 180246 (2019).
- Nelson, N. & Ben-Shem, A. The complex architecture of oxygenic photosynthesis. *Nat. Rev. Mol. Cell Biol.* **5**, 971–982 (2004).
- Mitchell, P. Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature* **191**, 144–148 (1961).
- Yoshida, M., Muneyuki, E. & Hisabori, T. ATP synthase: A marvellous rotary engine of the cell. *Nat. Rev. Mol. Cell Biol.* **2**, 669–677 (2001).
- Crowley, T. J. & Berner, R. A. CO<sub>2</sub> and climate change. *Science* **292**, 870–872 (2001).
- Badger, M. R. & Price, G. D. CO<sub>2</sub> concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. *J. Exp. Bot.* **54**, 609–622 (2003).
- Price, G. D., Maeda, S., Omata, T. & Badger, M. R. Modes of active inorganic carbon uptake in the cyanobacterium, *Synechococcus sp.* PCC7942. *Funct. Plant Biol.* **29**, 131–149 (2002).
- Silverman, D. N. & Lindskog, S. The catalytic mechanism of carbonic anhydrase: implications of a rate-limiting protolysis of water. *Acc. Chem. Res.* **21**, 30–36 (1988).
- Peltier, G., Aro, E. -M. & Shikanai, T. NDH-1 and NDH-2 plastoquinone reductases in oxygenic photosynthesis. *Annu. Rev. Plant Biol.* **67**, 55–80 (2016).
- Wang, H. et al. Rubisco condensate formation by CcmM in β-carboxysome biogenesis. *Nature* **566**, 131–135 (2019).
- Baradaran, R., Berrisford, J. M., Minhas, G. S. & Sazanov, L. A. Crystal structure of the entire respiratory complex I. *Nature* **494**, 443–448 (2013).
- Agip, A. A. et al. Cryo-EM structures of complex I from mouse heart mitochondria in two biochemically defined states. *Nat. Struct. Mol. Biol.* **25**, 548–556 (2018).
- Fiedorczuk, K. et al. Atomic structure of the entire mammalian mitochondrial complex I. *Nature* **538**, 406–410 (2016).
- Zickermann, V. et al. Mechanistic insight from the crystal structure of mitochondrial complex I. *Science* **347**, 44–49 (2015).
- Schuller, J. M. et al. Structural adaptations of photosynthetic complex I enable ferredoxin-dependent electron transfer. *Science* **363**, 257–260 (2019).
- Saura, P. & Kaila, V. R. I. Molecular dynamics and structural models of the cyanobacterial NDH-1 complex. *Biochim. Biophys. Acta—Bioenerg.* **1860**, 201–208 (2019).
- Laughlin, T. G., Bayne, A. N., Trempe, J. F., Savage, D. F. & Davies, K. M. Structure of the complex I-like molecule NDH of oxygenic photosynthesis. *Nature* **566**, 411–414 (2019).
- Yu, H. et al. Structure of an ancient respiratory system. *Cell* **173**, 1636–1649, e16 (2018).
- Kaila, V. R. I. Long-range proton-coupled electron transfer in biological energy conversion: Towards mechanistic understanding of respiratory complex I. *J. R. Soc. Interface* **15**, 20170916 (2018).
- Di Luca, A., Gamiz-Hernandez, A. P. & Kaila, V. R. I. Symmetry-related proton transfer pathways in respiratory complex I. *Proc. Natl. Acad. Sci. USA* **114**, E6314–E6321 (2017).
- Korste, A., Wulforst, H., Ikegami, T., Nowaczyk, M. M. & Stoll, R. Solution structure of the NDH-1 complex subunit CupS from *Thermosynechococcus elongatus*. *Biochim. Biophys. Acta—Bioenerg.* **1847**, 1212–1219 (2015).
- Ferraroni, M. et al. Crystal structure of a zinc-activated variant of human carbonic anhydrase I, CA I Michigan 1: Evidence for a second zinc binding site involving arginine coordination. *Biochemistry* **41**, 6237–6244 (2002).
- Kurisu, G., Zhang, H., Smith, J. L. & Cramer, W. A. Structure of the cytochrome b<sub>6</sub>f complex of oxygenic photosynthesis: tuning the cavity. *Science* **302**, 1009–1014 (2003).
- Rowlett, R. S. et al. Kinetic characterization of wild-type and proton transfer-impaired variants of β-carbonic anhydrase from *Arabidopsis thaliana*. *Arch. Biochem. Biophys.* **404**, 197–209 (2002).
- Simon, S. A. & Gutknecht, J. Solubility of carbon dioxide in lipid bilayer membranes and organic solvents. *Biochim. Biophys. Acta—Bioenerg.* **596**, 352–358 (1980).
- Mangan, N. M., Flamholz, A., Hood, R. D., Milo, R. & Savage, D. F. pH determines the energetic efficiency of the cyanobacterial CO<sub>2</sub> concentrating mechanism. *Proc. Natl. Acad. Sci. USA* **113**, E5354–E5362 (2016).
- Tagawa, K. & Arnon, D. I. Oxidation-reduction potentials and stoichiometry of electron transfer in ferredoxins. *Biochim. Biophys. Acta—Bioenerg.* **153**, 602–613 (1968).
- Golbeck, J. H. & Kok, B. Redox titration of electron acceptor Q and the plastoquinone pool in Photosystem II. *Biochim. Biophys. Acta—Bioenerg.* **547**, 347–360 (1979).
- Ritchie, J. Membrane potential and pH Control in The Cyanobacterium *Synechococcus R-2 (Anacystis nidulans)* PCC 7942. *J. Plant Phys.* **137**, 409–418 (1991).
- Lambert, A. J. & Brand, M. D. Superoxide production by NADH:ubiquinone oxidoreductase (complex I) depends on the pH gradient across the mitochondrial inner membrane. *Biochem. J.* **382**, 511–517 (2004).
- Di Luca, A., Mühlbauer, M. E., Saura, P. & Kaila, V. R. I. How inter-subunit contacts in the membrane domain of complex I affect proton transfer energetics. *BBA—Bioenerg.* **1859**, 734–741 (2018).
- Brandt, U. Adaptations of an ancient modular machine. *Science* **363**, 230–231 (2019).
- Wulforst, H., Franken, L. E., Wessinghage, T., Boekema, E. J. & Nowaczyk, M. M. The 5kDa protein NdhP is essential for stable NDH-1L assembly in *Thermosynechococcus elongatus*. *PLoS ONE* **9**, e103584 (2014).
- Zivanov, J., Nakane, T. & Scheres, S. H. W. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCr* **6**, 5–17 (2019).



35. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
36. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, pii: e42166 (2018).
37. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
38. Tan, Y. Z. et al. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods* **14**, 793–796 (2017).
39. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
40. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **60**, 2126–2132 (2004).
41. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. Sect. D. Struct. Biol.* **74**, 531–544 (2018).
42. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **66**, 12–21 (2010).
43. Barad, B. A. et al. EMRinger: side chain-directed model and map validation for 3D cryo-electron microscopy. *Nat. Methods* **12**, 943–946 (2015).
44. Delano, W. L. The PyMOL molecular graphics system. *Schrodinger* <http://www.pymol.sourceforge.net> (2002).
45. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).
46. Kim, C. U. et al. Tracking solvent and protein movement during CO<sub>2</sub> release in carbonic anhydrase II crystals. *Proc. Natl Acad. Sci.* **113**, 5257–5262 (2016).
47. Kimber, M. S. The active site architecture of *Pisum sativum* beta-carbonic anhydrase is a mirror image of that of alpha-carbonic anhydrases. *EMBO J.* **19**, 1407–1418 (2000).
48. Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98**, 5648–5652 (1993).
49. Lee, C., Yang, W. & Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **37**, 785–789 (1988).
50. Schäfer, A., Horn, H. & Ahlrichs, R. Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *J. Chem. Phys.* **97**, 2571–2577 (1992).
51. Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *J. Chem. Phys.* **132**, 154104–1–19 (2010).
52. Klamt, A. & Schüürmann, G. COSMO: A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc. Perkin Trans. 2*, 799–805 (1993).
53. Plessow, P. Reaction path optimization without NEB springs or interpolation algorithms. *J. Chem. Theory Comput.* **9**, 1305–1310 (2013).
54. E, W., Ren, W. & Vanden-Eijnden, E. String method for the study of rare events. *Phys. Rev. B – Condens. Matter Mater. Phys.* **66**, 523011–523014 (2002).
55. Ahlrichs, R., Bär, M., Häser, M., Horn, H. & Kölmel, C. Electronic structure calculations on workstation computers: The program system turbomole. *Chem. Phys. Lett.* **162**, 165–169 (1989).
56. Huang, J. & Mackerell, A. D. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J. Comput. Chem.* **34**, 2135–2145 (2013).
57. Riccardi, D., Yang, S. & Cui, Q. Proton transfer function of carbonic anhydrase: Insights from QM/MM simulations. *Biochimica et Biophysica Acta – Proteins Proteom.* **1804**, 342–351 (2010).
58. Riahi, S. & Rowley, C. N. The CHARMM-TURBOMOLE interface for efficient and accurate QM/MM molecular dynamics, free energies, and excited state properties. *J. Comput. Chem.* **35**, 2076–2086 (2014).
59. Webb, B. & Sali, A. Comparative protein structure modeling using modeller. *Curr. Protoc. Bioinforma.* **54**, 5.6.1–5.6.37 (2016).
60. Trabuco, L. G., Villa, E., Mitra, K., Frank, J. & Schulten, K. Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure* **16**, 673–683 (2008).
61. Chovancova, E. et al. CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. *PLoS Comput. Biol.* **8**, e1002708 (2012).
62. Phillips, J. C. et al. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
63. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
64. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl Acad. Sci.* **98**, 10037–10041 (2002).
65. Kieseritzky, G. & Knapp, E. W. Optimizing pKa computation in proteins with pH adapted conformations. *Proteins Struct. Funct. Genet.* **71**, 1335–1348 (2008).
66. Rabenstein, B. & Knapp, E. W. Calculated pH-dependent population and protonation of carbon-monooxygen myoglobin conformers. *Biophys. J.* **80**, 1141–1150 (2001).
67. Meyer, T. & Knapp, E. W. pKa values in proteins determined by electrostatics applied to molecular dynamics trajectories. *J. Chem. Theory Comput.* **11**, 2827–2840 (2015).
68. Kieseritzky, G. & Knapp, E. W. Improved pKa prediction: combining empirical and semimicroscopic methods. *J. Comput. Chem.* **29**, 2575–2581 (2008).
69. Baptista, A. M., Teixeira, V. H. & Soares, C. M. J. Constant-pH molecular dynamics using stochastic titration. *Chem. Phys.* **117**, 4184–4200 (2002).
70. Swails, J. M., York, D. M. & Roitberg, A. E. Constant pH replica exchange molecular dynamics in explicit solvent using discrete protonation states: implementation, testing, and validation. *J. Chem. Theory Comput.* **10**, 1341–1352 (2014).
71. Warnau, J. et al. Redox-coupled quinone dynamics in the respiratory complex I. *Proc. Natl Acad. Sci. USA* **115**, E8413–E8420 (2018).

## Acknowledgements

We thank H. Wulforst for construction of the CupS-TS mutant, M. Völkel for excellent technical assistance and Petra Düchting (Department MGPP, Ruhr University Bochum, Germany) for ICP-OES analysis. J.M.S. and S.K.S. would like to acknowledge E. Conti for unconditional support. J.M.S. is grateful to B.D. Engel, J.M. Plitzko, and W. Baumeister for access to the cryo-EM infrastructure and early career support. This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme/grant agreement no. 715311 (to V.R.I.K.), the DFG research unit FOR2092 (836/3-2 to M.M.N.), the DFG priority programme 2002 (NO 836/4-1 to M.M.N.) and the DFG grant NO 836/1-1 (to M.M.N.). This work was also supported by Emergence of Life initiative (TRR235 to V.R.I.K.) and by the Knut and Alice Wallenberg Foundation (to V.R.I.K.). Computational resources were provided by SuperMuc/Leibniz-Rechenzentrum (LRZ, grant: pn34he) and PRACE (project nr: 2018194738), awarding us access to MareNostrum at the Barcelona Supercomputing Centre (BSC), Spain (project: pr1ejk).

## Author contributions

J.T. and M.M.N. isolated and characterised the photosynthetic complex I; P.S., A.P.G.H. and V.R.I.K. performed molecular simulations; J.M.S. and S.K.S. prepared cryo-EM grids and collected data; J.M.S. and S.K.S. processed cryo-EM data; J.M.S., S.K.S. and V.R.I.K. built cryo-EM models; J.M.S., S.K.S. and V.R.I.K. analysed and interpreted the cryo-EM models; J.M.S., P.S., J.T., S.K.S., A.P.G.H., G.K., M.M.N. and V.R.I.K. analysed the data; J.M.S., M.M.N. and V.R.I.K. directed the project; V.R.I.K. wrote the manuscript with input from all authors.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41467-020-14347-4>.

**Correspondence** and requests for materials should be addressed to J.M.S., M.M.N. or V.R.I.K.

**Peer review information** *Nature Communications* thanks Marcelo Martí, Thorsten Friedrich and Janet Vonck 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.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020