





# Cryo-EM structures of engineered active $bc_1$ - $cbb_3$ type CIII<sub>2</sub>CIV super-complexes and electronic communication between the complexes

Stefan Steimle<sup>1</sup>, Trevor van Eeuwen <sup>2</sup>, Yavuz Ozturk<sup>1,3</sup>, Hee Jong Kim <sup>2</sup>, Merav Braitbard<sup>4</sup>, Nur Selamoglu<sup>1</sup>, Benjamin A. Garcia<sup>5</sup>, Dina Schneidman-Duhovny<sup>4</sup>, Kenji Murakami <sup>5</sup>✉ & Fevzi Daldal <sup>1</sup>✉

Respiratory electron transport complexes are organized as individual entities or combined as large supercomplexes (SC). Gram-negative bacteria deploy a mitochondrial-like cytochrome (cyt)  $bc_1$  (Complex III, CIII<sub>2</sub>), and may have specific  $cbb_3$ -type cyt  $c$  oxidases (Complex IV, CIV) instead of the canonical  $aa_3$ -type CIV. Electron transfer between these complexes is mediated by soluble ( $c_2$ ) and membrane-anchored ( $c_y$ ) cyts. Here, we report the structure of an engineered  $bc_1$ - $cbb_3$  type SC (CIII<sub>2</sub>CIV, 5.2 Å resolution) and three conformers of native CIII<sub>2</sub> (3.3 Å resolution). The SC is active in vivo and in vitro, contains all catalytic subunits and cofactors, and two extra transmembrane helices attributed to cyt  $c_y$  and the assembly factor CcoH. The cyt  $c_y$  is integral to SC, its cyt domain is mobile and it conveys electrons to CIV differently than cyt  $c_2$ . The successful production of a native-like functional SC and determination of its structure illustrate the characteristics of membrane-confined and membrane-external respiratory electron transport pathways in Gram-negative bacteria.

<sup>1</sup>Department of Biology, University of Pennsylvania, Philadelphia, PA, USA. <sup>2</sup>Biochemistry and Molecular Biophysics Graduate Group, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. <sup>3</sup>Institute of Biochemistry and Molecular Biology, Faculty of Medicine, Albert-Ludwigs University of Freiburg, Freiburg, Germany. <sup>4</sup>School of Computer Science and Engineering, Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel. <sup>5</sup>Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. ✉email: [kenjim@penmedicine.upenn.edu](mailto:kenjim@penmedicine.upenn.edu); [fdaldal@sas.upenn.edu](mailto:fdaldal@sas.upenn.edu)

Respiratory chains couple exergonic electron transport from nutrients to the terminal acceptor oxygen ( $O_2$ ) and generate a proton motive force used for ATP synthesis. Complex I (NADH dehydrogenase) and Complex II (succinate dehydrogenase) are the entry points into the chain of reducing equivalents, reducing quinone (Q) to hydroquinone ( $QH_2$ ). Complex III (cytochrome (cyt)  $bc_1$  or  $CIII_2$ ) oxidizes  $QH_2$  to reduce cyt  $c$ , which in turn is oxidized by Complex IV (cyt  $c$  oxidase or CIV) converting oxygen to water<sup>1</sup> (Fig. 1a).

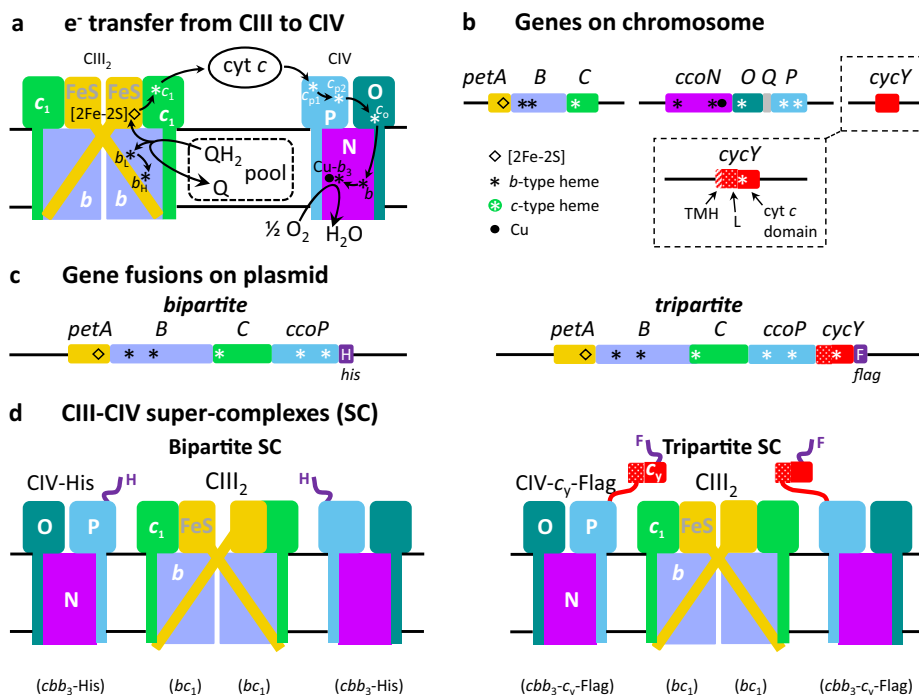
Respiratory complexes are evolutionarily conserved, and bacterial enzymes are structurally simple, consisting mainly of the catalytic subunits.  $CIII_2$  is a dimer with each monomer comprised of three subunits: the Rieske FeS (FeS) protein with a [2Fe-2S] cluster, cyt  $b$  with hemes  $b_H$  and  $b_L$ , and cyt  $c_1$  with heme  $c_1$  cofactors (Fig. 1a, b). The FeS protein external domain (FeS-ED) is mobile between the  $b$  (close to heme  $b_L$ ) and  $c$  (close to heme  $c_1$ ) positions<sup>2,3</sup>. Some bacterial species such as *Rhodobacter capsulatus*<sup>4</sup> and pathogens like *Helicobacter pylori*, *Campylobacter jejuni*<sup>5</sup>, and *Neisseria*<sup>6</sup> contain only a high oxygen affinity  $cbb_3$ -type CIV to support their micro-aerophilic growth. The  $cbb_3$ -type CIV is a monomer comprised of four subunits: CcoN with heme  $b$  and heme  $b_3$ -Cu binuclear center, CcoO with heme  $c_o$ , CcoQ, and CcoP with hemes  $c_{p1}$  and  $c_{p2}$  cofactors<sup>7</sup> (Fig. 1a, b).

Many Gram-negative bacteria contain soluble (e.g., cyt  $c_2$ ) and membrane-anchored (via transmembrane domains or fatty acids) cyts  $c$  (e.g., *R. capsulatus* cyt  $c_y$ <sup>8</sup>, *Paracoccus denitrificans* cyt  $c_{552}$ <sup>9</sup>,

and *Helicobacterium gestii* cyt  $c_{553}$ <sup>10</sup>) as electron carriers. Conversely, Gram-positive bacteria have no freely diffusing electron carrier but contain cyt  $c$  domains fused to their  $CIII_2$  like in *Mycobacterium smegmatis*<sup>11</sup>, or CIV like in *Bacillus subtilis*<sup>12</sup>. In *R. capsulatus*, both the diffusible cyt  $c_2$  and the membrane-anchored cyt  $c_y$  carry electrons from  $CIII_2$  to CIV in respiration<sup>13</sup>.

In recent years, the co-occurrence of individual complexes together with multi-enzyme supercomplexes (SCs) in energy-transducing membranes has become evident<sup>14–17</sup>. SCs may enhance catalytic efficiency through substrate/product channeling or minimize production of harmful reactive oxygen species to decrease cellular distress<sup>18</sup>. The structures of mitochondrial SCs, such as  $CIII_2$ CIV or its smaller variants  $CIII_2$  and  $CIII_2$ -CIV<sup>19</sup> are well known<sup>20,21</sup>. Some bacterial SCs (e.g., *P. denitrificans*<sup>22</sup> or *C. glutamicum*<sup>23</sup>) have also been characterized biochemically, but only the structure of the Gram-positive *M. smegmatis* SC ( $CIII_2$ CIV<sub>2</sub>) is available<sup>24,25</sup>.

As of yet, no respiratory SC structure has been determined for Gram-negative bacteria, the evolutionary precursors of mitochondria. Furthermore, SCs containing  $cbb_3$ -type ancient forms of CIV with primordial respiratory features remain unknown<sup>26</sup>. Such structural studies have been hampered due to unstable interactions between  $CIII_2$  and CIV, yielding trace amounts of “native” SCs. We have overcome this hurdle using a genetic approach and obtained large amounts of engineered SCs from the Gram-negative facultative phototroph *R. capsulatus*. Here we



**Fig. 1**  $QH_2$  oxidation and fused SCs. **a** Bifurcated electron transfer reaction conveys one electron from  $QH_2$  oxidation to the FeS protein (FeS, yellow) [2Fe-2S] cluster and another electron to cyt  $b$  (periwinkle) hemes  $b_L$  and  $b_H$ . Reduced [2Fe-2S] cluster transfers the electron to cyt  $c_1$  (green) heme  $c_1$  via the movement of the FeS protein from near heme  $b_L$  to near heme  $c_1$ . The subsequent electron transfer step from heme  $b_H$  to Q from the pool is not shown for clarity. A cyt  $c$  ( $c_2$  or  $c_y$ ) carries electrons from heme  $c_1$  to CIV, where electrons reach the heme-Cu (Cu- $b_3$ ) site, where  $O_2$  is reduced to  $H_2O$ , via the CcoP (P, light blue) hemes  $c_{p1}$  and  $c_{p2}$ , CcoO (O, dark green) heme  $c_o$ , and CcoN (N, purple) heme  $b$ . **b** *R. capsulatus*  $petABC$  encodes the structural genes of the  $bc_1$ -type  $CIII_2$  subunits, the FeS protein ( $petA$ , yellow), cyt  $b$  (B, periwinkle), and cyt  $c_1$  (C, green), and  $ccoNOP$  encodes those of the  $cbb_3$ -type CIV subunits, the CcoN ( $ccoN$ , purple), CcoO (O, dark green), CcoQ (Q, gray), and CcoP (P, light blue).  $cycY$  gene (red) encodes cyt  $c_y$ , and its 30-residue transmembrane helix (TMH), 69-residue linker (L), and 100-residue cyt  $c$  (cyt  $c$ ) domain are indicated. Heme cofactors of  $b$ - and  $c$ -type cyts are indicated by black and white asterisks, respectively, whereas diamond and dot designate the [2Fe-2S] cluster and Cu atom, respectively. **c** Plasmid-borne genetic fusions. The bipartite fusion (left) is formed by in-frame linking the 3'-end of  $petC$  to the 5'-end of  $ccoP$  and the tripartite fusion (right) by adding the linker and cyt  $c$  domain of  $cycY$  to the 3'-end of  $ccoP$ . Colors and cofactors are as in **a**, and the his (H) and flag (F) affinity tags (dark purple) of the bipartite and tripartite fusion subunits are shown. **d** Schematic depiction of bipartite (left) and tripartite (right) supercomplexes (SC). The bipartite SC encodes a  $CIII_2$  dimer fused on each side to a His-tagged CIV. The tripartite SC also contains the Flag-tagged cyt  $c$  domain of  $c_y$  (red) at the end of CcoP (blue).

report the cryo-electron microscopic (cryo-EM) structure of this functional *bc<sub>1</sub>-cbb<sub>3</sub>*-type SC (CIII<sub>2</sub>CIV, at 5.2 Å resolution), together with three structural conformers of native CIII<sub>2</sub> (at 3.3–4.2 Å resolution). The structures show that the membrane-bound cyt *c<sub>γ</sub>* and the diffusible cyt *c<sub>2</sub>* interact differently with CIV for electron transfer. This work illustrates the structural and functional features of a native-like CIII<sub>2</sub>CIV and its distinct membrane-confined and membrane-peripheral respiratory electron transport pathways in Gram-negative bacteria.

## Results

**Isolation and characterization of engineered functional SCs.** Earlier studies have indicated that in *R. capsulatus* CIII<sub>2</sub>, CIV, and cyt *c<sub>γ</sub>* are in close proximity to each other<sup>27</sup>. Analyses of native membranes of this species indicated barely detectable entities around ~450 kDa *M<sub>r</sub>*, larger than the masses of CIV monomers or CIII<sub>2</sub> dimers (Supplementary Fig. 1a, b). Similar entities were seen with dodecyl-maltoside or digitonin used to disperse wild-type membrane, but they were of low abundance and unstable, rendering their isolation and study difficult.

Earlier works of the assembly processes of CIII<sub>2</sub> and CIV showed that cyt *c<sub>1</sub>* interacts with cyt *b* to form a cyt *b-c<sub>1</sub>* subcomplex<sup>28</sup>, and similarly, CcoP associates with CcoNOQ subcomplex to yield an active CIV<sup>29</sup>. Thus we hypothesized that translationally fusing the C-terminus (C-ter) of cyt *c<sub>1</sub>* to the N-terminus (N-ter) of CcoP (on the inner (*n*) side of the membrane) as a cyt *c<sub>1</sub>*-CcoP fusion might produce a bipartite *bc<sub>1</sub>-cbb<sub>3</sub>*-type SC that may form a native-like SC and enrich possible SC assembly factors during protein isolation for biochemical and structural studies (left panels of Fig. 1c, d). Further, adding the 69-residue linker (L) and the 100-residue cyt domain of cyt *c<sub>γ</sub>* to the C-ter of cyt *c<sub>1</sub>*-CcoP (on the outer (*p*) side of the membrane) as a cyt *c<sub>1</sub>*-CcoP-*c<sub>γ</sub>* fusion might yield a tripartite SC with its attached electron carrier (right panels of Fig. 1c, d). This approach (see “Methods” and Supplementary Table 1) was successful and yielded high amounts of fused SCs of native-like electrophoretic mobilities (Supplementary Fig. 1a, b; a was overstained for IGA as compared to b). Indeed, these fused native-like SCs were functional in vivo and complemented a mutant lacking both CIII<sub>2</sub> and CIV for photosynthesis proficiency, which requires a functional CIII<sub>2</sub>, and colonies producing them exhibited CIV activity in situ (Supplementary Fig. 1c). Thus the functional properties of the native-like SCs were like those of native SCs.

The His-tagged bipartite and Flag-tagged tripartite SCs were purified by tag-affinity and size exclusion chromatography (SEC; “Methods”; Supplementary Fig. 1d, e). Analyses of isolated proteins showed the presence of large entities of *M<sub>r</sub>* ~450 kDa range (Supplementary Fig. 1d, e fractions A-1 and B-1 in insets). These fractions contained the cyt *c<sub>1</sub>*-CcoP (~65 kDa) or cyt *c<sub>1</sub>*-CcoP-*c<sub>γ</sub>* fusions (~80 kDa) (Supplementary Fig. 1f), and the protein bands were identified by mass spectrometry (MS) as subunits of CIII<sub>2</sub> and CIV (Supplementary Table 2). The fusion proteins contained covalently attached heme cofactor(s) as revealed by a cyt *c*-specific staining (Supplementary Fig. 1g). CcoQ (*M<sub>r</sub>* ~7 kDa) of CIV was absent in the SC preparations.

Purified SCs were characterized for their cyts *b* and *c* contents. The optical redox difference spectra were distinct from those of CIII<sub>2</sub><sup>30</sup> or CIV<sup>31</sup>, and the tripartite SC contained more *c*-type cyt than the bipartite SC, due to the additional cyt domain of cyt *c<sub>γ</sub>* (Supplementary Fig. 1h). The samples had both QH<sub>2</sub>:cyt *c* reductase (specific to CIII<sub>2</sub>) and cyt *c* oxidase (specific to CIV) activities (Supplementary Fig. 1i, j). Thus the native-like SCs were enzymatically active in vitro. Moreover, the tripartite SC also exhibited QH<sub>2</sub>:O<sub>2</sub> reductase (coupled CIII<sub>2</sub> + CIV) activity without addition of any electron carrier (e.g., horse heart cyt *c* or *R.*

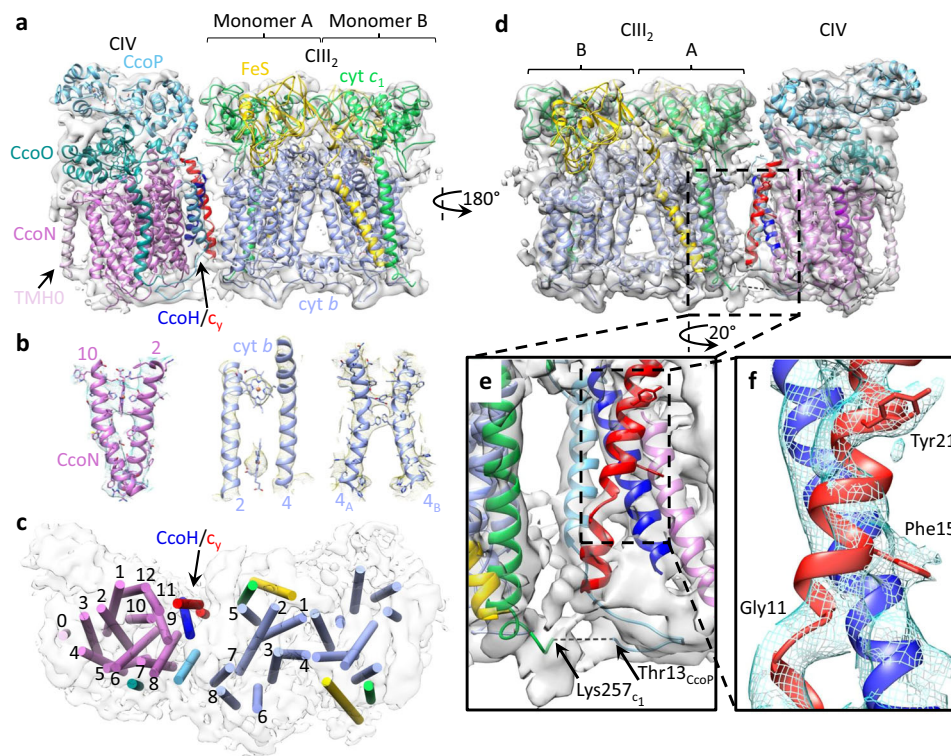
*capsulatus* cyt *c<sub>2</sub>*; Supplementary Fig. 1k), suggesting that the cyt domain of cyt *c<sub>γ</sub>* fused to cyt *c<sub>1</sub>*-CcoP transferred electrons from CIII<sub>2</sub> to CIV.

**Structures of the tripartite SCs.** Cryo-EM analysis of the tripartite SC preparations (Supplementary Fig. 1e, fraction B-1) showed that the initial three-dimensional (3D) classes were mainly of two different sizes (Supplementary Fig. 2, Box 1, left). The size (~180 Å length) and shape of the smaller particles suggested that these may correspond to a dimeric CIII<sub>2</sub> associated with a single CIV. Focused classification and processing of the subclass containing ~62,000 particles with the highest initial resolution and best discernable features yielded a tripartite CIII<sub>2</sub>CIV map (SC-1A, EMD-22228) at 6.1 Å resolution (Supplementary Fig. 2a, “Methods”), while another dataset produced a map (SC-1B, EMD-22230) at 7.2 Å resolution (Supplementary Fig. 2b and Supplementary Table 3). The larger particles (~250 Å length, Supplementary Fig. 2, Box 1, left) represented a dimeric CIII<sub>2</sub> flanked by two CIV, as expected based on two *c<sub>1</sub>*-CcoP-*c<sub>γ</sub>* subunits per CIII<sub>2</sub>. However, these particles were rare (~5000) and their map (SC-1C) could not be refined beyond ~10 Å resolution (Supplementary Fig. 2c).

A homology model of *R. capsulatus cbb<sub>3</sub>*-type CIV was built using as a template the highly homologous *Pseudomonas stutzeri* structure (PDB: 3MK7; 3.2 Å resolution) (“Methods,” Supplementary Table 3). The available CIII<sub>2</sub> model (PDB: 1ZRT; 3.5 Å resolution) was further refined (PDB: 6XI0; 3.3 Å resolution) based on the cryo-EM data obtained in this work (Supplementary Table 3). These models were fitted as rigid bodies into the maps SC-1A and SC-1B (Supplementary Fig. 3a), with correlation coefficients *CC<sub>box</sub>* of 0.75 and *CC<sub>box</sub>* of 0.71, respectively. The [2Fe-2S] clusters of the FeS proteins of CIII<sub>2</sub> could be recognized closer to *b* (heme *b<sub>L</sub>*) than to *c* (heme *c<sub>1</sub>*) positions but had lower occupancy and resolution likely due to conformational heterogeneity. The heterogeneity of the FeS-ED of CIII<sub>2</sub> in monomer B (away from CIV) was more pronounced than that in monomer A (adjacent to CIV) (Supplementary Fig. 3b). Lower resolutions of the CIII<sub>2</sub> FeS-ED portions were anticipated due to their mobility<sup>2,3</sup>.

Superimposition of the CIII<sub>2</sub> portions of SC-1A and SC-1B maps showed that CIV was in different orientations in different maps (Supplementary Fig. 3c). The two extreme locations of CIV with respect to CIII<sub>2</sub> were displaced from each other by a translation of ~3 Å and a rotation of ~37° (Supplementary Fig. 3d, e; SC-1A in red and SC-1B in blue). Other subclasses identified in 3D classifications showed CIV in various orientations between those seen in SC-1A and SC-1B maps. This CIII<sub>2</sub>CIV interface flexibility is attributed to the limited interaction between the CcoP (N-ter transmembrane helix (TMH)) of CIV and the cyt *b* (TMH7) of CIII<sub>2</sub> (Fig. 2, top view). In the interface regions of SC-1A and SC-1B maps, additional weaker features that are not readily attributable to CIII<sub>2</sub> and CIV structures were observed. However, no membrane-external feature that might correspond to cyt domain of cyt *c<sub>γ</sub>*, which is an integral part of the cyt *c<sub>1</sub>*-CcoP-*c<sub>γ</sub>* subunit, could be discerned in these maps.

**Structure of bipartite SC supplemented with cyt *c<sub>γ</sub>*.** The bipartite SC preparations (Supplementary Fig. 1d, fraction A-1) were supplemented with either purified full-length cyt *c<sub>γ</sub>* or its soluble variant lacking the TMH (i.e., cyt *S-c<sub>γ</sub>*)<sup>32</sup> to yield the bipartite SC + *c<sub>γ</sub>* and SC + *S-c<sub>γ</sub>* samples. Following SEC, the elution fractions showed that only the intact cyt *c<sub>γ</sub>*, but not the cyt *S-c<sub>γ</sub>*, remained associated with the SC (Supplementary Fig. 4a). Thus the cyt domain of cyt *c<sub>γ</sub>* does not bind tightly to, and its TMH is required for its association with, the SC.



**Fig. 2 Cryo-EM structure of CIII<sub>2</sub>CIV.** **a** The structures of CIII<sub>2</sub> (PDB: 6XIO) and the homology model of *R. capsulatus* *cbb*<sub>3</sub>-type CIV were fitted into the cryo-EM map SC-2A (gray), depicted as a side view. All subunits are colored and labeled as in Fig. 1, and the arrow at the edge of CcoN of CIV is for the extra N-ter TMH (TMH0, light purple) specific to *R. capsulatus*. **b** Representative regions of the cryo-EM map showing map quality and model fitting. The TMH2 and TMH10 of CcoN show heme *b* and some bulky side chains. The protein backbone and hemes *b<sub>L</sub>* and *b<sub>H</sub>* are resolved between *cyt b* TMH2–TMH4 (compare to Fig. 3b, CIII<sub>2</sub> map at 3.3 Å). Large side chains are visible between the TMH4s of CIII<sub>2</sub> monomers A and B (4<sub>A</sub> and 4<sub>B</sub>). **c** Top view of CIII<sub>2</sub>CIV with TMHs depicted as cylinders and colored as in **a**. The TMHs of *cyt b* (only CIII<sub>2</sub> monomer A) and CcoN of CIV are numbered, and the TMHs of the FeS protein (yellow), *cyt c<sub>1</sub>* (green), CcoO (dark green), CcoP (light blue), CcoH/*c<sub>γ</sub>* (blue/red with an arrow), and CcoN TMH0 (light purple) are shown. **d** 180° rotated view for the back view of CIII<sub>2</sub>CIV interface. The two extra TMHs at the interface are those of CcoH (blue) and *cyt c<sub>γ</sub>* (red). **e** Enlarged view of CIII<sub>2</sub>CIV interface. The view is slightly rotated relative to **d** for better visibility of CcoP TMH in the background (light blue). CcoN TMH9 is next to CcoH and *cyt c<sub>γ</sub>* TMHs. The fusion region between *cyt c<sub>1</sub>* and CcoP is indicated at the bottom, with *cyt c<sub>1</sub>* C-ter (green) and CcoP N-ter (light blue) with their respective end residues (Lys257<sub>*c<sub>1</sub>*</sub> and Thr13<sub>CcoP</sub>) (only resolved portions). The 12 N-ter CcoP residues connecting these two chains (dashed line) are not resolved. **f** Enlarged view showing close interaction between CcoH and *cyt c<sub>γ</sub>* TMHs. Characteristic features of *cyt c<sub>γ</sub>* TMH (NH<sub>2</sub>-Gly11xxxPhe15xxxxTyr21-COOH) are used to determine the registration, with the helix break induced by Gly11, and bulky side chains for Phe15 and Tyr21 clearly visible.

The cryo-EM analyses of the bipartite SC + *c<sub>γ</sub>* yielded a map (SC-2A, EMD-22227) at 5.2 Å resolution (Supplementary Fig. 5a, b), with local resolutions ranging from 4.3 to 8.0 Å (Supplementary Fig. 6a, c). The homology model of CIV and the refined model of CIII<sub>2</sub> (PDB: 6XIO) were fitted as rigid bodies into SC-2A (Fig. 2a), with a correlation coefficient CC<sub>box</sub> of 0.74. Comparison of the maps SC-2A (bipartite CIII<sub>2</sub>CIV) with SC-1A (tripartite CIII<sub>2</sub>CIV) showed that they were similar, with root-mean-square deviation of 1.6 Å. Hence, they are collectively referred to as CIII<sub>2</sub>CIV, irrespective of their bipartite or tripartite origins.

The dimensions (~155 × 60 × 90 Å, *L* × *W* × *H*) of the slightly curved CIII<sub>2</sub>CIV structure were consistent with a CIII<sub>2</sub> dimer associated with one CIV. On SC-2A map (5.2 Å resolution), some large aromatic side chains could be readily discerned (Fig. 2b). Of the TMHs seen, 34 were accounted for by two FeS proteins, two *cyt b*, and two *cyt c<sub>1</sub>* (2, 16, and 2 TMHs per dimer, respectively) and single CcoN, CcoO, and CcoP (12, 1, and 1 TMHs, respectively) (Fig. 2c). The features corresponding to the heme cofactors of CIII<sub>2</sub>CIV were attributed to hemes *b<sub>H</sub>* and *b<sub>L</sub>* of *cyt b*, heme *c<sub>1</sub>* of *cyt c<sub>1</sub>*, and to hemes *b* and *b<sub>3</sub>* of CcoN, heme *c<sub>0</sub>* of CcoO, and hemes *c<sub>p1</sub>* and *c<sub>p2</sub>* of CcoP proteins. Like the tripartite maps, in this case also the FeS-EDs of CIII<sub>2</sub> had lower resolution and the [2Fe-2S] clusters were located closer to heme

*b<sub>L</sub>* (**b** position). An additional TMH was observed at the distal end of CIV (Fig. 2a, rotated 180° in Fig. 2d) close to CcoN TMH3 and TMH4 (Fig. 2c). Due to its location, this TMH (depicted in Fig. 2 as an ab initio model of CcoN residues Arg25 to Leu48) was tentatively attributed to the extra N-ter TMH (TMH0) of CcoN, specific to *R. capsulatus*.

The interface of CIII<sub>2</sub>CIV is delimited by CcoN TMH8 and TMH9, CcoP TMH, *cyt b* TMH5 and TMH7, and *cyt c<sub>1</sub>* TMH of monomer A, with the closest interaction being between CcoP TMH and *cyt b* TMH7 (Fig. 2a, c). Two interacting TMHs of unknown identities (Fig. 2d, in red and blue) and an inter-complex connection are seen at the interface. This connection linking CIII<sub>2</sub> to CIV is at the *n* face of the membrane, near the TMHs of *cyt c<sub>1</sub>* and CcoP (Fig. 2e, Lys257<sub>*c<sub>1</sub>*</sub> and Thr13<sub>CcoP</sub>). It is tentatively attributed to the 13 amino acid residues spanning the fusion joint between the C-end of PetC and N-end of CcoP (with no extra amino acid inserted in between) that are not clearly observed.

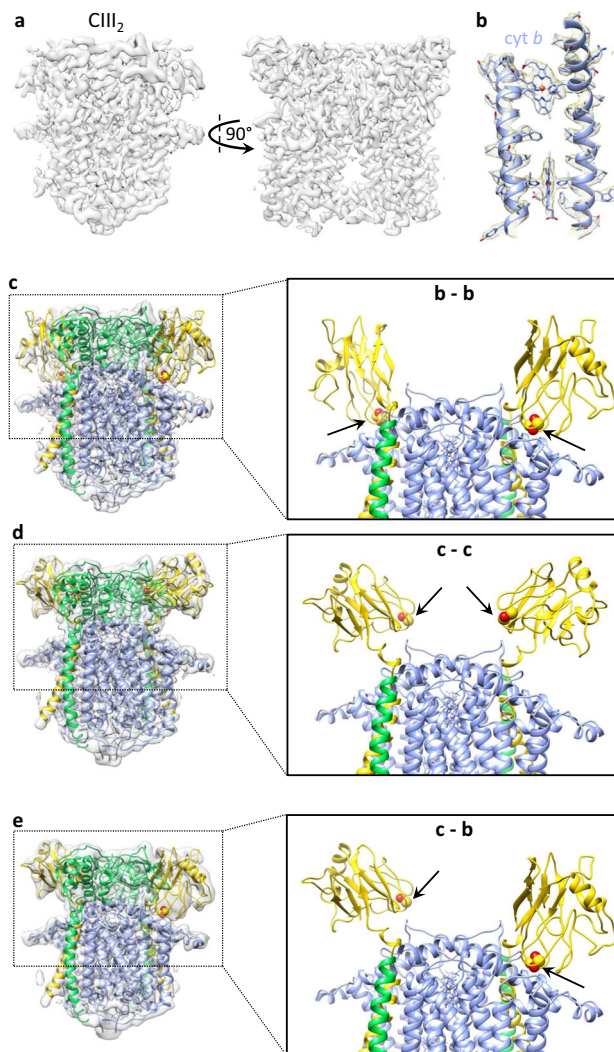
**Assembly factor CcoH and *cyt c<sub>γ</sub>* TMHs are located at CIII<sub>2</sub>-CIV interface.** A co-evolution-based approach predicting the residue–residue contacts in protein–protein interactions,

RaptorX-ComplexContact<sup>33</sup>, was used to identify the TMHs at CIII<sub>2</sub>CIV interface. The single TMH containing CIV-related proteins (i.e., CcoQ subunit, CcoS and CcoH assembly factors<sup>29</sup>, and cyt *c*<sub>γ</sub><sup>8</sup>) were analyzed against the subunits of CIII<sub>2</sub> and CIV. Significant predictions of interacting residue pairs (confidence value >0.5) were found only between CcoN (primarily TMH9) and the putative N-term TMH of CcoH (Supplementary Fig. 7a, b). An ab initio model of CcoH TMH was docked via Patch-Dock<sup>34</sup> onto CIV using the predicted residue-residue contacts as distance restraints (15 Å threshold; “Methods”). The top scoring 50 models converged to a single cluster at the location of the unknown TMH, next to CcoN TMH9 (Supplementary Fig. 7c). Examination of the interactions between CcoH TMH and CcoN TMH9 showed multiple co-evolutionarily conserved residues in close contacts (Supplementary Fig. 7b, d). Earlier studies had indicated that CcoH can be cross-linked by disuccinimidyl suberate (spacer length ~11 Å) to CcoP and CcoN<sup>35</sup>. Thus the TMH close to CcoN TMH9 (Fig. 2e, blue TMH) is tentatively assigned to the assembly factor CcoH.

A clear difference between the maps SC-2A (bipartite CIII<sub>2</sub>CIV + *c*<sub>γ</sub>) and SC-1A (tripartite CIII<sub>2</sub>CIV) was seen at the interface. The unidentified TMHs were barely visible in SC-1A but highly enhanced in SC-2A (Fig. 2e), indicating higher occupancy. The observation that only the native cyt *c*<sub>γ</sub> binds to bipartite SC via its TMH suggested that the TMH (red in Fig. 2e) next to CcoH TMH (blue in Fig. 2e) corresponds to the membrane anchor of cyt *c*<sub>γ</sub>. This explanation is supported by fact that the bipartite CIII<sub>2</sub>CIV + *c*<sub>γ</sub> is supplemented with full-length cyt *c*<sub>γ</sub> while the tripartite SC contains only the fused cyt domain of *c*<sub>γ</sub>. Indeed, landmark densities corresponding to the helix-breaking Gly11 and the correctly spaced bulky side chains of Phe15 and Tyr21 of cyt *c*<sub>γ</sub> TMH (NH<sub>2</sub>-xxxGly11xxxPhe15xxxx-Tyr21-COOH) were discerned (Fig. 2f).

We noted that some CIII<sub>2</sub>CIV + *c*<sub>γ</sub> subclasses exhibited a weak feature on the *p* side of the membrane that may reflect the cyt domain of *c*<sub>γ</sub> (Supplementary Fig. 5g, SC-2B). However, this feature could not be refined to high resolution, consistent with the weak binding of cyt domain of *c*<sub>γ</sub> to CIII<sub>2</sub>CIV. Moreover, upon addition of cyt *c*<sub>γ</sub>, the predominant conformation of CIV in the bipartite CIII<sub>2</sub>CIV + *c*<sub>γ</sub> (Supplementary Fig. 5a, b, SC-2A) shifted toward that seen in map SC-1A of tripartite SC (Supplementary Fig. 2a), with no major class corresponding to SC-1B (Supplementary Fig. 2b). This suggested that local interactions between the CcoH and cyt *c*<sub>γ</sub> TMHs and CIV decreased the interface flexibility of CIII<sub>2</sub>CIV (Fig. 2e).

**Cryo-EM structures of *R. capsulatus* native CIII<sub>2</sub>.** The bipartite SC + *c*<sub>γ</sub> samples also contained smaller particles (~110 Å length) (Supplementary Fig. 2, Box 2) that were the size of CIII<sub>2</sub> (Supplementary Fig. 5c, d). Analyses of these particles using C2 symmetry led to the map CIII<sub>2</sub> at 3.3 Å resolution for native CIII<sub>2</sub> (Supplementary Fig. 5e), with local resolutions ranging from 3.0 to 4.0 Å (Supplementary Fig. 6b, d and Supplementary Table 3). The FeS-ED parts showed lower occupancy and resolution compared to the rest of the map, indicating conformational heterogeneity. When similar analyses were carried out without imposing C2 symmetry, three distinct maps were obtained for CIII<sub>2</sub> (CIII<sub>2</sub> c-c, CIII<sub>2</sub> b-c, and CIII<sub>2</sub> b-b at 3.8, 4.2, and 3.5 Å resolutions, respectively; Supplementary Fig. 5f). These maps were superimposable with respect to cyt *b* and cyt *c*<sub>1</sub> subunits, except for the FeS-ED portions. The CIII<sub>2</sub> structures depicted by the CIII<sub>2</sub> b-b (Fig. 3a–c) and CIII<sub>2</sub> c-c (Fig. 3d) maps exhibited overall C2 symmetry, but in the former the FeS-EDs were located in **b**, whereas in the latter in **c** position<sup>3</sup>. Notably, the third structure (CIII<sub>2</sub> b-c) was asymmetric, with the FeS-ED of one



**Fig. 3 Structures of native CIII<sub>2</sub> conformers.** **a** Cryo-EM map CIII<sub>2</sub> b-b (EMD-22226) with both FeS proteins in **b** position. **b** Representative region of CIII<sub>2</sub> b-b demonstrating map quality and model fitting. TMH2 and TMH4 of cyt *b* with hemes *b*<sub>L</sub> and *b*<sub>H</sub> are shown. **c–e** Maps and models showing different conformations of the FeS proteins. The left panels show the CIII<sub>2</sub> structure fitted into the different maps with the subunit colorings (cyt *b* in periwinkle, cyt *c*<sub>1</sub> in green, and the FeS protein in yellow) as in Fig. 2. The right panels show the top half of the models with the membrane-external domain of cyt *c*<sub>1</sub> omitted to better visualize the b-b, c-c, and c-b positions of FeS-EDs. The [2Fe-2S] clusters are shown as yellow-red spheres and indicated by arrows. **c** Structure of native CIII<sub>2</sub> with both FeS-EDs in **b** position (map CIII<sub>2</sub> b-b, EMD-22226; PDB: 6XKV). **d** Structure of native CIII<sub>2</sub> with both FeS-EDs in **c** position (map CIII<sub>2</sub> c-c, EMD-22224; PDB: 6XKT). **e** Structure of native CIII<sub>2</sub> with one FeS-ED in **c** and one in **b** position (map CIII<sub>2</sub> b-c, EMD-22225; PDB: 6XKU).

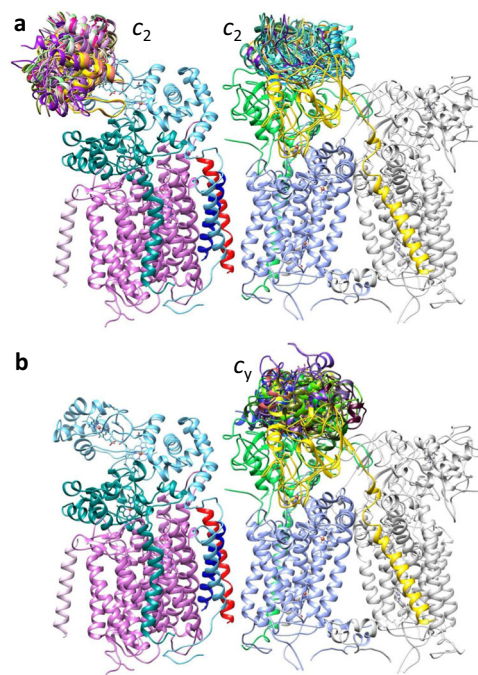
monomer being in **c** and the other one in **b** positions (Fig. 3e). Such asymmetric structures of native CIII<sub>2</sub> have not been observed frequently, although they are proposed to occur during QH<sub>2</sub> oxidation<sup>36–38</sup>.

**Interactions of cyt *c*<sub>2</sub> and cyt *c*<sub>γ</sub> with CIII<sub>2</sub>CIV.** As the tripartite or bipartite SC maps contained no feature that could clearly represent cyt domain of *c*<sub>γ</sub>, the interaction interfaces between SC and its electron carrier remained unknown. Thus a docking approach combined with crosslinking mass spectrometry

(XL-MS) was used to characterize these protein-protein interactions. XL-MS has been used successfully *in vivo* with whole cells<sup>39–42</sup> or *in vitro* with purified materials<sup>43,44</sup>, and PatchDock<sup>34</sup> is an efficient rigid docking method that maximizes geometric shape complementarity. First, the co-crystal structure (PDB: 3CX5) of yeast CIII<sub>2</sub> with its electron carrier iso-1 cyt *c*<sup>45</sup> was used as a template (homology between yeast cyt *c*<sub>1</sub> and *R. capsulatus* cyt *c*<sub>1</sub>: 31% identity and 58% similarity; iso-cyt *c* and cyt *c*<sub>2</sub>: 25% identity and 56% similarity) to model the binding of cyt *c*<sub>2</sub> on bacterial CIII<sub>2</sub>. The *R. capsulatus* cyt *c*<sub>1</sub> (PDB: 6X10) and cyt *c*<sub>2</sub> (PDB: 1C2N) structures were superimposed with their yeast counterparts on the co-crystal structure, and a model with a single cyt *c*<sub>2</sub> docked to a monomer of CIII<sub>2</sub> was generated (Supplementary Fig. 4e). This docking model was then experimentally validated using the protein crosslinker 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methyl-morpholinium chloride (DMTMM). The crosslinks (XLS) were identified using MeroX<sup>44,46</sup> search engine, and among them those of high confidence were cross-validated using FindXL<sup>43,47</sup>. The intra-subunit XLS within CIII<sub>2</sub>CIV provided controls (Supplementary Fig. 4c, d, “Methods”). The three XLS between cyt *c*<sub>1</sub> and cyt *c*<sub>2</sub> supplied distance restraints for docking cyt *c*<sub>2</sub> to CIII<sub>2</sub> using PatchDock (Supplementary Table 4 and Supplementary Fig. 4f), and the docking models (top 50) clustered at a single region per monomer of CIII<sub>2</sub> (Fig. 4a, right). This region overlapped with the binding site of cyt *c*<sub>2</sub> defined by the model obtained by alignment to the yeast co-crystal structure (Supplementary Fig. 4g). The distance from cyt *c*<sub>2</sub> heme-Fe to cyt *c*<sub>1</sub> heme-Fe was ~16.8 Å for the co-crystal-derived model, while comparable distances were between ~13.8 and 20.4 Å with the docking models. Thus docking via PatchDock integrating XL-MS-mediated distance restraints defined reliably, but with limited accuracy, the interaction region of cyt *c*<sub>2</sub> on CIII<sub>2</sub>.

No information about the binding of cyt *c*<sub>2</sub> and *ccb*<sub>3</sub>-type CIV was previously available. First, this binding was shown biochemically (Supplementary Fig. 4b) and then subjected to XL-MS using DMTMM. The XLS between the proteins (1 between cyt *c*<sub>2</sub> and CcoP and 8 between cyt *c*<sub>2</sub> and CcoO) provided distance restraints for docking cyt *c*<sub>2</sub> to CIV (Supplementary Table 4). The docking models (top 50) clustered in a single region of CIV (Fig. 4a, left), closer to heme *c*<sub>p2</sub> (*c*<sub>2</sub> heme-Fe to *c*<sub>p2</sub> heme-Fe: ~15.2–35.6 Å) than heme *c*<sub>p1</sub> (*c*<sub>2</sub> heme-Fe to *c*<sub>p1</sub> heme-Fe: ~23.0–42.0 Å) of CcoP (Fig. 5). The two cyt *c*<sub>2</sub>-binding regions on CIII<sub>2</sub>CIV are highly distant from each other (closest *c*<sub>2</sub> heme-Fe on CIII<sub>2</sub> to that on CIV is ~69 Å; Fig. 5a).

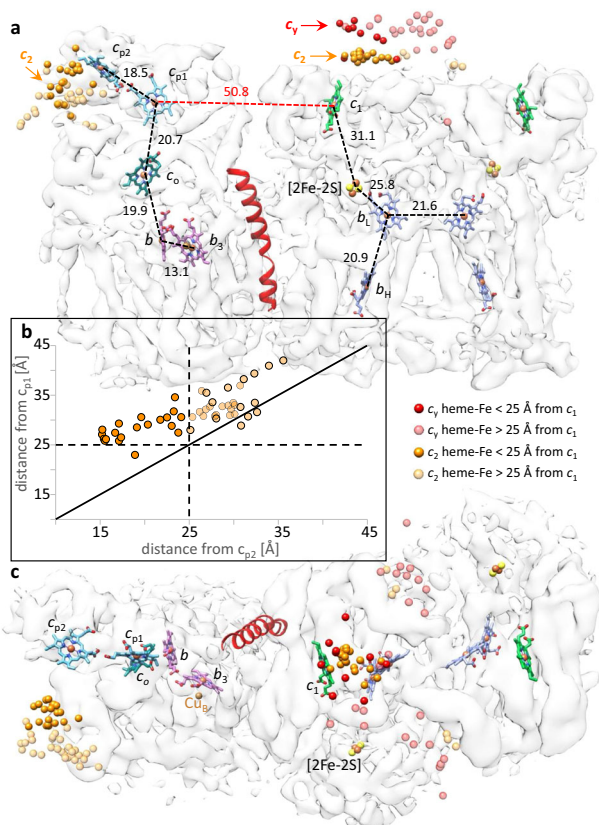
Next, the binding interactions between cyt domain of *c*<sub>γ</sub> and CIII<sub>2</sub>CIV were sought using the longer, primary amine reactive disuccinimidyl dibutyric urea (DSBU) together with the shorter, carboxylic acid reactive DMTMM. The use of two crosslinkers with different chemical properties and spacer lengths provided better coverage and tighter distance restraints, as the SC interface and the FeS-ED of CIII<sub>2</sub> were, and the cyt domain of *c*<sub>γ</sub> was expected to be, flexible (see below). Similar to DMTMM, DSBU yielded multiple intra-subunit XLS within the subunits of CIII<sub>2</sub>CIV as experimental controls (Supplementary Table 4 and Supplementary Fig. 4c, d). Six XLS (five cyt *c*<sub>γ</sub> to cyt *c*<sub>1</sub> and one cyt *c*<sub>γ</sub> to FeS protein) with DMTMM and four XLS (only cyt *c*<sub>γ</sub> to FeS protein) with DSBU were found (Supplementary Table 4). No XL was observed between cyt *c*<sub>γ</sub> and CIV, suggesting that cyt *c* domain of *c*<sub>γ</sub> is closer to CIII<sub>2</sub> in CIII<sub>2</sub>CIV. In this case, the models (top 50) generated by PatchDock clustered in two binding regions for cyt domain of *c*<sub>γ</sub> on each CIII<sub>2</sub> monomer of SC. One of the clusters was on top of cyt *c*<sub>1</sub> and overlapped with cyt *c*<sub>2</sub> cluster (Fig. 4b), whereas the other one was between cyt *c*<sub>1</sub> and the FeS-ED near CIII<sub>2</sub> inter-monomer space (Supplementary Fig. 4h–j). These two clusters are best visible in a top view



**Fig. 4** Binding regions of cyt *c*<sub>2</sub> and cyt *c*<sub>γ</sub> on CIII<sub>2</sub>CIV. The binding regions were defined by XL-MS-guided docking, and the subunits of CIII<sub>2</sub>CIV are colored as in Fig. 2 except the monomer B of CIII<sub>2</sub> shown in light gray. Only the binding regions on monomer A are shown. **a** Cyt *c*<sub>2</sub> (PDB: 1C2N) was docked onto CIII<sub>2</sub> and CIV using PatchDock guided by DMTMM-generated XLS as distance restraints. This yielded one cluster of models on CIV and one per monomer of CIII<sub>2</sub>. **b** A model of cyt domain of *c*<sub>γ</sub>, generated using *P. denitrificans* cyt *c*<sub>552</sub> structure (PDB: 3M97) as a template (RMSD between template and model: 0.2 Å) was docked on CIII<sub>2</sub> as in **a**, except that both DMTMM- and DSBU-generated XLS provided distance restraints. Two binding clusters for cyt domain of *c*<sub>γ</sub> per monomer of CIII<sub>2</sub> were found. The two clusters are located behind each other on a side view but are clearly visible on top views (Supplementary Fig. 4h–j, labeled 1 and 2). Only cluster 1 that is closer to cyt *c*<sub>1</sub> and overlapping with the binding region of cyt *c*<sub>2</sub> is shown. In all cases, ten representative models are shown to visualize the clusters of binding models. No binding region for cyt *c*<sub>γ</sub> on CIV could be defined since no XL was found between the proteins.

(Supplementary Fig. 4j), and the distances between cyt *c*<sub>γ</sub> heme-Fe and cyt *c*<sub>1</sub> heme-Fe of CIII<sub>2</sub> monomer A are between 13.8 and 47.1 Å (Fig. 5). To further support these binding locations, we searched in our cryo-EM datasets 3D classes that contained extra features attributable to cyt domain of *c*<sub>γ</sub>, and a minor class (~18,000 particles; Supplementary Fig. 5g) showing such an extra feature between CIV and CIII was found (Supplementary Fig. 4h, i).

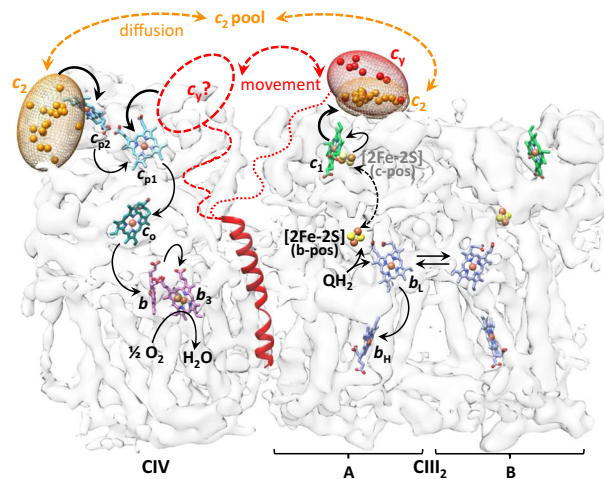
To better define the clusters of cyt domain of *c*<sub>γ</sub>, the docking analyses were repeated using the three different conformers of native CIII<sub>2</sub> (Fig. 3c–e, CIII<sub>2</sub> b-b, c-c, and b-c). Interestingly, when the FeS-EDs are in **c** position (CIII<sub>2</sub> c-c), the docking models gathered as a single cluster on cyt *c*<sub>1</sub>, slightly displaced toward the FeS-ED of the same monomer (Supplementary Fig. 8a–c). However, when the FeS-EDs are in **b** position (CIII<sub>2</sub> b-b), such models were more spread out (Supplementary Fig. 8d–f). The third model with one FeS-ED in **c** and the other in **b** positions yielded the expected clustering pattern depending on the local FeS-ED conformation. In CIII<sub>2</sub>CIV both FeS-EDs being in the **b** position, we assumed that cyt domain of *c*<sub>γ</sub> docks similarly to that seen with CIII<sub>2</sub> b-b and that spreading of docking originates from the variable conformations of the FeS-EDs in SC (Supplementary Fig. 8). Heme *c*<sub>1</sub> and not the FeS protein being



**Fig. 5 Organization of CIII<sub>2</sub>CIV cofactors and redox partners.** **a** The cofactors are shown inside the transparent map SC-2A of CIII<sub>2</sub>CIV with the same subunit colors as in Fig. 2: hemes *b*<sub>L</sub> and *b*<sub>H</sub> (periwinkle), heme *c*<sub>1</sub> (green), hemes *c*<sub>p1</sub> and *c*<sub>p2</sub> (light blue), heme *c*<sub>o</sub> (dark green), hemes *b* and *b*<sub>3</sub> (purple). The [2Fe-2S] clusters are shown as yellow-red spheres. In all cases, the distances (heme-Fe to heme-Fe) between the heme cofactors are indicated. The positions of docked cyt *c*<sub>2</sub> and cyt domain of *c*<sub>y</sub> are indicated as orange (heme *c*<sub>2</sub>) and red (heme *c*<sub>y</sub>) spheres, respectively, representing their heme-Fe atoms. All heme-Fe atoms corresponding to the top 50 docking positions for cyt *c*<sub>2</sub> on CIV are shown as solid (<25 Å) or transparent (>25 Å) spheres, depending on their distances to heme *c*<sub>p2</sub>. In the case of CIII<sub>2</sub>, only the docking positions of cyt *c*<sub>2</sub> and cyt *c*<sub>y</sub> on monomer A and between the monomers A and B are shown, omitting those located on monomer B. The TMH of cyt *c*<sub>y</sub> is shown in red at CIII<sub>2</sub>CIV interface. **b** The heme-Fe atoms of all 50 cyt *c*<sub>2</sub> models docked onto CIV are plotted as function of their distances from heme *c*<sub>p1</sub> and heme *c*<sub>p2</sub>, with the Fe atoms within 25 Å shown as solid spheres and those beyond 25 Å as transparent spheres. The vast majority of heme-Fe atoms of docked cyt *c*<sub>2</sub> models are closer to heme *c*<sub>p2</sub> than heme *c*<sub>p1</sub> of CIV. **c** Top view of the map shown in **a** is presented to better visualize the distribution of the docked cyt domain of *c*<sub>y</sub> on monomer A and between the monomers A and B. In all cases, the heme-Fe atoms are depicted by spheres and colored as indicated above and on the figure.

the electron exit site of CIII<sub>2</sub><sup>48,49</sup>, the cluster on cyt *c*<sub>1</sub> was taken as the productive binding region of cyt domain of *c*<sub>y</sub>.

Examination of all pertinent distances between the cofactors of CIII<sub>2</sub>CIV (Fig. 5a) indicates that the binding region of cyt domain of *c*<sub>y</sub> near heme *c*<sub>1</sub> is far from CIV. The large distance (~50.8 Å) separating cyt *c*<sub>1</sub> heme-Fe of CIII<sub>2</sub> monomer A from CcoP *c*<sub>p1</sub> heme-Fe (the closest compared with heme *c*<sub>p2</sub>) makes it impossible to define a position for cyt *c*<sub>y</sub> close enough to heme *c*<sub>1</sub> reducing it, and heme *c*<sub>p1</sub> oxidizing it, while transferring electrons from CIII<sub>2</sub> to CIV. The distance constraint, the inability to resolve the cyt domain of cyt *c*<sub>y</sub> by cryo-EM, and the more



**Fig. 6 Proposed cyt *c*<sub>2</sub> and cyt *c*<sub>y</sub> binding regions of CIII<sub>2</sub>CIV and electron transfer pathways.** The likely binding regions of cyt *c*<sub>2</sub> and cyt *c* domains of *c*<sub>y</sub> (orange and red ellipsoids, respectively), defined by XL-MS-guided docking, are depicted by the distributions of their heme-Fe atoms on the transparent map SC-2A of CIII<sub>2</sub>CIV. Only the positions that are within 25 Å of heme *c*<sub>1</sub> of CIII<sub>2</sub> or heme *c*<sub>p2</sub> of CIV are indicated. The cofactors together with the TMH of cyt *c*<sub>y</sub> are shown as in Fig. 5. The linker region (indicated by dotted or dashed lines) between the TMH and the cyt domain of *c*<sub>y</sub> is not resolved in the cryo-EM map. The proposed electron transport pathways are shown by thicker black arrows: upon QH<sub>2</sub> oxidation by CIII<sub>2</sub>, cyt *c*<sub>y</sub> that is integral to CIII<sub>2</sub>CIV receives an electron from heme *c*<sub>1</sub>. It then moves (double-headed dashed red arrow) to an undefined binding region (dashed oval with *c*<sub>y</sub>?) on CIV, where it delivers the electron to the nearest heme *c*<sub>p1</sub> of CIV. Similarly, cyt *c*<sub>2</sub> that is peripheral to CIII<sub>2</sub>CIV also receives an electron from heme *c*<sub>1</sub>, diffuses away to reach CIV, and conveys it to heme *c*<sub>p2</sub>. The canonical electron transfers occurring from QH<sub>2</sub> to heme *c*<sub>1</sub> in CIII<sub>2</sub>, and from heme *c*<sub>p1</sub> to O<sub>2</sub> in CIV, are indicated by thinner arrows. The double-headed dashed black arrow depicts the movement of the [2Fe-2S] of FeS protein from the **b** position (b-pos, black, solid spheres) to the **c** position (c-pos, gray, transparent spheres) in CIII<sub>2</sub> during QH<sub>2</sub> oxidation. A possible electron equilibration between the two heme *b*<sub>L</sub> of CIII<sub>2</sub> is indicated by double arrows, and the electron transfer steps subsequent to heme *b*<sub>H</sub> reduction are not shown.

frequent XLs to CIII<sub>2</sub> led us to conclude that cyt *c*<sub>y</sub> oscillates within CIII<sub>2</sub>CIV during QH<sub>2</sub>:O<sub>2</sub> oxidoreduction (Fig. 6).

## Discussion

Prior to this work, no information was available neither on the occurrence or structure of any *cbb*<sub>3</sub>-type CIV containing SC nor on its interactions with physiological electron carriers. Here we describe the cryo-EM structures of engineered native-like active *bc*<sub>1</sub>-*cbb*<sub>3</sub>-type respiratory SCs, together with three structural conformers of native CIII<sub>2</sub> at high resolution, from the Gram-negative, facultative phototroph *R. capsulatus*. Although the 3D structures of bacterial *bc*<sub>1</sub>-type CIII<sub>2</sub> are available, native CIII<sub>2</sub> heterodimers have not been reported. Similarly, only one structure was available for *cbb*<sub>3</sub>-type CIV<sup>50</sup>. Members of this subfamily of heme-Cu:O<sub>2</sub> reductases are widespread among bacteria and essential for microaerobic processes, including anaerobic photosynthesis and bacterial infection<sup>4</sup>.

The facultative CIII<sub>2</sub>CIV of *R. capsulatus* is more flexible, possibly due to the involvement of CIII<sub>2</sub> into multiple metabolic pathways, unlike the highly rigid obligate CIII<sub>2</sub>CIV<sub>2</sub> SC of *Actinobacteria* dedicated to respiration<sup>24,25</sup>. In species like *R. capsulatus*, the dual function of CIII<sub>2</sub>, interacting with the photochemical reaction center during photosynthesis and with

CIV during respiration, may need this natural flexibility to allow metabolic adaptations for growth fitness. The native-like SCs, obtained by genetically fusing a subunit of CIII<sub>2</sub> to another one of CIV, are fully active both in vivo and in vitro and exhibit increased stability. These SCs are produced in cells without any harmful effect, mature properly, and support growth of mutants lacking either CIII<sub>2</sub> or CIV or both. Moreover, detergent-dispersed membranes and affinity-purified SCs exhibit both CIII<sub>2</sub> and CIV, as well as CIII<sub>2</sub>-CIV coupled, enzymatic activities. Clearly this innovative approach, although yielded heterogeneous preparations, containing mixtures of CIII<sub>2</sub>CIV<sub>2</sub>, CIII<sub>2</sub>CIV, and CIII<sub>2</sub> particles, allowed resolution of their structures and defined the likely binding regions of cyt *c*<sub>2</sub> and cyt *c*<sub>γ</sub>. Insertion of different spacers at the cyt *c*<sub>1</sub>-CcoP fusion junction and over-expression of the subunits and related assembly components could not further improve this heterogeneity (“Methods”). Certainly, it remains unknown to what extent the structural and functional properties of the native-like SCs described here are identical to their “native” counterparts, which are currently unavailable.

The structures of the tripartite CIII<sub>2</sub>CIV, which contains a fused cyt *c* domain<sup>39</sup>, and the bipartite CIII<sub>2</sub>CIV + *c*<sub>γ</sub>, which has a bound native cyt *c*<sub>γ</sub>, were highly similar at sub-nanometer resolution (~7.2–5.2 Å, respectively). Their flexibility was due to limited protein–protein interactions at their interface, where the TMHs of cyt *c*<sub>γ</sub> and CcoH are located (Fig. 2). Previously, neither the exact location nor the mobility of cyt *c*<sub>γ</sub> between CIII<sub>2</sub> and CIV were known. The SC structure shows that the N-terminal TMH of cyt *c*<sub>γ</sub> is locked at the interface, allowing mobility of its cyt domain within CIII<sub>2</sub>CIV (Fig. 6). The linker region of cyt *c*<sub>γ</sub> remains unresolved but is long enough to reach both CIII<sub>2</sub> and CIV. Early studies with *R. capsulatus* cyt *c*<sub>γ</sub> had shown that a full-length linker (69 residues) is needed for rapid (~50 μs) electron transfer from CIII<sub>2</sub> to the photosynthetic reaction center<sup>51</sup>, whereas a shorter linker (~45 residues) is proficient for respiratory electron transfer to CIV<sup>52</sup>.

In native CIII<sub>2</sub> conformers, different positions of the [2Fe-2S] cluster bearing FeS-EDs were observed. Crystallographic structures have often depicted bacterial CIII<sub>2</sub> as symmetrical homodimers<sup>3,53,54</sup>. These structures were obtained in the presence of inhibitors constraining FeS-EDs near heme *b*<sub>L</sub> or used mutants stabilizing it on cyt *b* surface. Alternatively, they contained crystal contacts restricting the FeS-ED movement<sup>55</sup>. To our knowledge, no native heterodimeric CIII<sub>2</sub> structure of bacterial origin with different conformations of its FeS-EDs has been reported. Only recent cryo-EM structures of mitochondrial SCs have shown different maps for CIII<sub>2</sub> FeS-EDs<sup>17,56</sup>. Thus native CIII<sub>2</sub> is not always a symmetric homodimer, and the FeS-ED of each monomer is free to move independently from the other. The Q-cycle models describe the mechanism of CIII<sub>2</sub> catalysis by two turnovers of a given monomer, with the different positions of the FeS-ED protein attributed to different catalytic steps<sup>3,48,49</sup>. Emerging asymmetric structures of bacterial and mitochondrial native CIII<sub>2</sub> obtained by cryo-EM in the absence of inhibitors or mutations, combined with the inter-monomer electron transfer between the hemes *b*<sub>L</sub> of the monomers<sup>57,58</sup>, now provide a glimpse into plausible “heterodimeric Q-cycle” mechanism(s)<sup>37,38</sup>. Accordingly, the FeS-EDs of CIII<sub>2</sub> monomers might cycle between homo- and hetero-dimeric conformations during the Q-cycle.

This work defined hitherto unknown structural interactions between the cyt *c*<sub>2</sub> or cyt *c*<sub>γ</sub> and native-like CIII<sub>2</sub>CIV (Fig. 6). The structures indicate that the distances separating heme *c*<sub>1</sub> of CIII<sub>2</sub> monomer A and hemes *c*<sub>p1</sub> or *c*<sub>p2</sub> of CIV are too large (Fig. 5) for direct microsecond time scale electronic communication<sup>59</sup> to sustain the turnover rate of CIII<sub>2</sub>CIV. Thus, even when CIII<sub>2</sub> and

CIV form a SC, an electron carrier is required for QH<sub>2</sub>:O<sub>2</sub> oxidoreduction. The binding location of cyt *c*<sub>2</sub> on CIV defined here, the redox midpoint potentials (*E*<sub>m</sub>) of the cofactors, and the distances separating them (Fig. 5a) suggest that cyt *c*<sub>2</sub> would confer electrons to the closer heme *c*<sub>p2</sub>, rather than the more distant heme *c*<sub>p1</sub>, of CcoP. This will then initiate canonical electron transfer via heme *c*<sub>p1</sub>, heme *c*<sub>o</sub>, and heme *b* to heme *b*<sub>3</sub>-Cu<sub>B</sub> site for O<sub>2</sub> reduction<sup>60,61</sup> (Fig. 6). For purified *R. capsulatus* proteins, the *E*<sub>m</sub> values of cyt *c*<sub>2</sub>, cyt *c*<sub>1</sub><sup>30</sup>, and heme *c*<sub>o</sub> of CcoO are known<sup>31</sup>. The *E*<sub>m</sub> values of *R. capsulatus* CcoP hemes *c*<sub>p1</sub> and *c*<sub>p2</sub> are unknown, but based on similar *E*<sub>m</sub> values of heme *c*<sub>o</sub> for *Bradyrhizobium japonicum* and *R. capsulatus*, they are expected to be similar<sup>62</sup>.

The likely interaction region of cyt *c*<sub>γ</sub> on cyt *c*<sub>1</sub> of CIII<sub>2</sub>, which is close to that of cyt *c*<sub>2</sub>, is now known but that on CIV remains less well defined. The distance separating the redox centers is a major factor that controls the rate of electron transfer<sup>59</sup>. The binding region of cyt domain of *c*<sub>γ</sub> on CIII<sub>2</sub> suggests that reduced cyt *c*<sub>γ</sub>, upon its movement to CIV, might preferentially convey electrons to the closer heme *c*<sub>p1</sub> than heme *c*<sub>p2</sub> of CcoP (Fig. 6). If so, under physiological conditions, heme *c*<sub>p1</sub> would be the primary receiver of electrons derived from QH<sub>2</sub> oxidation by CIII<sub>2</sub>, forming a fully membrane-confined electronic wiring within CIII<sub>2</sub>CIV. In contrast, cyt *c*<sub>2</sub> carries electrons from heme *c*<sub>1</sub> to heme *c*<sub>p2</sub> via free diffusion. This membrane-external pathway might accommodate electrons not only from QH<sub>2</sub> but also from other donors distinct from CIII<sub>2</sub>. As such, reduction of cyt *c*<sub>2</sub> during methylamine oxidation<sup>63</sup> or degradation of sulfur containing amino acids, converting toxic sulfite (SO<sub>3</sub><sup>2-</sup>) to sulfate (SO<sub>4</sub><sup>2-</sup>) by sulfate oxidase<sup>64</sup>, might provide electrons to CIV, contributing to energy production.

In summary, the architecture of an engineered native-like CIII<sub>2</sub>CIV along with its interactions with its physiological redox partners establishes salient structural and functional features of two distinct respiratory electron transport pathways (membrane-confined and membrane-external) that operate between CIII<sub>2</sub> and CIV in Gram-negative bacteria.

## Methods

**Bacterial strains, growth conditions, and primers used.** Bacterial strains, plasmids, and primers used are listed in Supplementary Table 1. LB medium supplemented as appropriate with ampicillin, gentamicin, kanamycin, or tetracycline at 100, 12, 50, or 12.5 μg/mL, respectively, was used for growing *Escherichia coli* strains at 37 °C. *R. capsulatus* strains were grown chemoheterotrophically under semi-aerobic/dark conditions at 35 °C on enriched (MPYE) medium, supplemented as needed with gentamicin, kanamycin, spectinomycin, or tetracycline at 3, 10, 10, or 2.5 μg/mL, respectively. Colonies were stained for *cbh*<sub>3</sub>-type CIV activity by incubating plates with a 1:1 (v/v) mixture of 35 mM 1-naphthol and 30 mM *N,N*-dimethyl-1,4-phenylenediamine (NADI staining)<sup>65</sup>.

**Construction of *petABC::ccoP*-His<sub>8</sub> (cyt *bc*<sub>γ</sub>-CcoP) fusion.** Using the primers Fw-ccoP (StuI) and Rv-ccoP (HindIII) (Supplementary Table 1), *petABC::ccoP*-His<sub>8</sub> fusion was constructed by ligating in-frame the PCR-amplified 1.13 kb StuI-HindIII fragment containing *ccoP* (with a C-terminal His<sub>8</sub> tag) to the 3' end of *petABC*, after elimination of *petC* stop codon and *ccoP* start codon on StuI-HindIII digested plasmid pMTS1, to yield pYO60 (Supplementary Table 1). The 3.85 kb BamHI fragment of pYO60 carrying *petABC::ccoP* fusion was transferred to the plasmids pBSII and pRK415 using the same sites, yielding pYO63 and pYO76, respectively. The linker-spaced versions of *petABC::ccoP* fusion (*petABC::L2-ccoP*, *petABC::L3-ccoP*, and *petABC::L4-ccoP*) were also constructed by exchanging the StuI-HindIII fragment of pYO63 carrying *petABC-ccoP* with the PCR-amplified linker added versions, yielding pYO77, pYO78, and pYO80, respectively. L2, L3, and L4 linkers were introduced via the primers F-L2, F-L3, and F-L4 (Supplementary Table 1) and contained the amino acid sequences of NH<sub>2</sub>-GGSGGGSGGSG-COOH, NH<sub>2</sub>-GGSGGGSGGGSG-COOH, and NH<sub>2</sub>-ASIAGGRTASGP-COOH, respectively. The 3.9 kb KpnI-XbaI fragments carrying these linker added versions were cloned to pRK415, yielding pYO81, pYO82, and pYO83, respectively. All constructs were subsequently verified by DNA sequencing. As the protein yields and enzymatic activities of all constructs were similar, only pYO76, which has no linker (i.e., native-like fusion), was used for subsequent work.



**Construction of *petABC::ccoP::cycY-Flag* (cyt *bc*<sub>1</sub>-CcoP-cyt *c*<sub>Y</sub>-Flag).** Using the primers Fw-cp-cy (Ball) and Rv-cp-cy (BglII + HinIII), the C-terminally Flag-tagged *cycY* lacking its N-terminal TMH anchor (i.e., amino acid residues 1–30) was PCR amplified. The 619-bp-long PCR fragment containing *cycY-Flag* without its anchor was digested with Ball-HinIII and exchanged with its counterpart that encompasses the 3' end of *ccoP* on pYO63 to yield the *petABC::ccoP::cycY-Flag* fusion carried by pYO91. The 4.5 kb KpnI-XbaI fragment of pYO91 containing this fusion was cloned into pRK415 using the same sites to yield pYO92 (Supplementary Table 1).

**Inactivation of CIII<sub>2</sub> (cyt *bc*<sub>1</sub>) and CIV (*ccb3*-type Cox) structural genes.** The  $\Delta$  (*petABC::gm*) deletion-insertion allele carried by pYO34<sup>32</sup> was transferred using the gene transfer agent into the chromosome of *R. capsulatus* strain MG1 (*ccoP::kan*) to yield YO12, a mutant background lacking both CIII<sub>2</sub> and the CcoP subunit of CIV. A similar strain, M7G-CBC1 that lacks both CIII<sub>2</sub> and CcoP but over-producing CcoN and CcoO of CIV<sup>31</sup>, was also used.

**Construction of strains producing bipartite and tripartite fusion SCs.** The plasmids pYO76 and pYO92 encoding the bipartite and tripartite SCs, respectively, were conjugated into YO12 and M7G-CBC1 using triparental mating to yield pYO76/YO12 and pYO92/M7G-CBC1 strains as in ref. <sup>65</sup> for protein purification.

**Purification of fusion SCs.** *R. capsulatus* cells (~25–30 g from 8 L media) were resuspended in a final volume of 35–40 mL of buffer A (50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 20 mM EDTA, 1 mM aminocaproic acid, and 1 mM 4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride (AEBSF) supplemented with Pierce Protease Inhibitor (Thermo Scientific, 1 mini tablet per 30 mL). Cells were disrupted by three passages through a French Pressure Cell (SLM Aminco) at 13,000 psi in the presence of 5–10 mg of DNase I (GoldBio), and cell debris was removed by centrifugation at 27,000 × g for 30 min<sup>31</sup>. Chromatophore membranes were sedimented by centrifugation at 190,000 × g for 2 h and resuspended in buffer A. The protein concentration was determined using the Pierce Protein BCA Assay Kit (Thermo Scientific), the suspension was supplemented with 10% glycerol and adjusted to a protein concentration of 10 mg/mL with buffer A. Membrane proteins were solubilized with 2% n-dodecyl- $\beta$ -D-maltoside (DDM; Anatrace) for 15 min on ice, and non-solubilized materials were sedimented by centrifugation at 100,000 × g for 15 min. The supernatant was subjected to chromatography, as appropriate.

The His-tagged bipartite SC was purified by anion exchange chromatography on Bio-Gel A-50 (Bio-Rad), followed by affinity chromatography on Ni-Sepharose High Performance (GE Healthcare). Solubilized proteins were loaded onto a 40 mL Bio-Gel column equilibrated with Tris buffer (50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1% glycerol, and 0.01% DDM) and washed with three column volumes (CVs) of the buffer. After another washing step with 3 CVs of Tris buffer containing 150 mM NaCl to eliminate weakly bound proteins, elution was carried out with a linear, five CV gradient from 150 to 400 mM NaCl in Tris buffer. Fractions containing both cyt *c* reductase and cyt *c* oxidase activities were combined and loaded onto a 7-mL Ni-Sepharose column equilibrated with the binding buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 20 mM imidazole, 1% glycerol, and 0.01% DDM). After washing with 30 mL of binding buffer, the column was eluted with a linear 100 mL gradient from 20 to 260 mM imidazole.

The Flag-tagged tripartite SC was purified by affinity chromatography using the Anti-Flag affinity gel (Bimake). The supernatant containing solubilized proteins was loaded onto a 1 mL column equilibrated with TBS buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.01% DDM, and 1 mM AEBSF), washed with 10 mL TBS, and eluted with 5 mL of Flag (DYKDDDDK) peptide (100  $\mu$ g/mL) (Sigma) in TBS buffer.

Eluted proteins were concentrated (Amicon Ultra-15, 30 kDa molecular weight cut-off (MWCO)) (Millipore) to a final volume of 150–200  $\mu$ L and loaded onto a Superose 6 Increase 10/300 GL (GE Healthcare) size exclusion column equilibrated with TBS buffer and eluted with the same buffer at a flow rate of 0.4 mL/min. Peak fractions were combined and concentrated, and PD MiniTrap G-25 columns (GE Healthcare) were used for buffer exchanges, as needed.

**Biochemical characterization of fusion SCs.** Native polyacrylamide gel electrophoresis (PAGE) was performed using 4–13% gels<sup>66</sup>, and sodium dodecyl sulfate (SDS)-PAGE used 12.5 or 18% gels<sup>67</sup>. Gels were stained with Coomassie Brilliant Blue (Bio-Rad) or colloidal silver<sup>68</sup>. Immunoblot analyses were performed using polyclonal antibodies specific of *R. capsulatus* cyt *b*<sup>28</sup> and monoclonal anti-rabbit-IgG-alkaline phosphatase (Sigma) used at 1:10,000 dilution. CIV in-gel activity was revealed by incubating native gels with 0.5 mg/mL 3,3'-diaminobenzidine in 50 mM Na-phosphate, pH 7.2<sup>69</sup>. The *c*-type cyts were revealed by their heme peroxidase activity, using 3,3',5,5'-tetramethylbenzidine (TMBZ) and hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>)<sup>70</sup>. Briefly, gels were washed with 0.25 M Na-acetate buffer (pH 5) and incubated with 6.3 mM TMBZ in 30% methanol and 0.25 M Na-acetate, pH 5.0. After 1-h incubation, 0.4% H<sub>2</sub>O<sub>2</sub> were added to reveal the peroxidase activity of *c*-type cyts.

Reduced *minus* oxidized spectra were recorded in 50 mM MOPS, pH 7.0 using a Cary 60 UV-Vis spectrophotometer (Agilent). Usually, 50  $\mu$ g of purified protein

was fully oxidized by adding a grain of K<sub>3</sub>[Fe(CN)<sub>6</sub>] and then gradually reduced with a few grains of sodium ascorbate or sodium dithionite, as appropriate. Protein concentrations were determined with a NanoDrop 2000c (Thermo Scientific) using the A<sub>280</sub> method<sup>71</sup>.

Cyt *c* reductase (cyt *bc*<sub>1</sub>) activity was determined as described<sup>72</sup>. Accordingly, 10 mM stock solution of 2,3-dimethoxy-5-methyl-6-decyl-1,4-benzoquinone (DB) (Sigma) in dimethyl sulfoxide was reduced to DBH<sub>2</sub> with a few grains of sodium borohydride and excess borohydride was quenched by adding HCl to a final pH of 6.0. For assays, 40  $\mu$ M DBH<sub>2</sub> were added to 25  $\mu$ M horse heart cyt *c* (Sigma) in 500  $\mu$ L assay buffer (40 mM sodium phosphate, pH 7.4, 20 mM sodium malonate, 0.5 mM EDTA, 0.5 mM KCN, and 0.01% DDM) in a stirred cuvette, and the non-enzymatic rate of cyt *c* reduction was recorded at 550 nm for 1 min (Cary 60 UV-Vis spectrophotometer, Agilent). The reaction was initiated by adding 1–2  $\mu$ g of purified protein and cyt *c* reduction was monitored. The enzymatic activity was inhibited by adding 1  $\mu$ M stigmatellin (Fluka), inhibitor of cyt *bc*<sub>1</sub>-type CIII<sub>2</sub>.

The CIV (cyt *ccb3*-Cox) activity was determined as described<sup>73</sup>. Reduced horse heart cyt *c* (Sigma) was prepared by incubating a 1–2 mM stock solution with 10 mM sodium dithionite for 15 min at room temperature, excess dithionite was removed by gel filtration using a PD-10 desalting column (GE Healthcare), and the final concentration of reduced cyt *c* was calculated based on its absorption at 550 nm and an  $\epsilon = 20 \text{ mM}^{-1} \text{ cm}^{-1}$ . In all, 20  $\mu$ M reduced cyt *c* in 500  $\mu$ L assay buffer (10 mM Tris-HCl, pH 7.0, 100 mM KCl) were prepared in a stirred cuvette (Cary 60 UV-Vis spectrophotometer), and the baseline was recorded for 1 min. The reaction was started by adding 1–2  $\mu$ g of purified protein, and cyt *c* oxidation monitored at 550 nm. The enzymatic activity was inhibited by addition of 250  $\mu$ M KCN as an inhibitor of *ccb3*-Cox.

DBH<sub>2</sub>-dependent oxygen consumption was monitored using a mini Clark-type oxygen electrode (Instech Laboratories, PA) at 25 °C. The assay was performed in the same buffer as that used for cyt *c* reductase activity (40 mM sodium phosphate, pH 7.4, 20 mM sodium malonate, 0.5 mM EDTA, and 0.01% DDM) but without KCN. A baseline was recorded after adding 100  $\mu$ M DBH<sub>2</sub> to 1 mL of assay buffer, and the reaction was started by adding 30  $\mu$ g of purified protein. As needed, purified protein was pre-incubated with a 1:1 (w/w) mixture of *E. coli* polar lipids (Avanti), which improved the activity. In all, 20  $\mu$ M horse heart cyt *c* (Sigma), 250  $\mu$ M KCN, or 1  $\mu$ M of stigmatellin were used as specificity controls. For all measurements, three biological and three technical repeats are used.

**Identification of protein subunits by MS.** Protein subunits were identified by MS using appropriate gel bands reduced (dithiothreitol (DTT); Sigma), alkylated (iodoacetamide; Bio-Rad) and subjected to in-gel trypsin digestion (Promega, Sequencing Grade Modified Trypsin) overnight at 37 °C. Peptides eluted from the gel samples were dried, desalted using ZipTips (Millipore U-C18 P10, Millipore), lyophilized, and stored at –80 °C. They were resuspended in 10  $\mu$ L 5% acetonitrile (ACN)/0.1% formic acid prior to MS, using either a LCQ Deca XP + ion trap, or a Q-Exactive Quadrupole-Orbitrap mass spectrometer (both from Thermo Fisher Scientific)<sup>74</sup>. The LCQ Deca XP+ mass spectrometer was coupled to a Thermo-Dionex LC Packings Ultimate Nano HPLC system controlled by the Thermo Xcalibur version 2.0 software. Peptides were separated on a 15 cm C18 nano-column (Thermo-Dionex, NAN-75-15-03-PM) using a 45-min linear gradient from 4 to 40% buffer B (100% ACN, 0.1% formic acid), followed by a 7-min gradient from 40 to 80% buffer B and 8-min wash with 80% B (constant flow rate 150 nL/min). MS/MS data were acquired in data-dependent analysis mode with dynamic exclusion enabled (repeat count: 3, exclusion duration: 3 min). Full MS survey scans (mass range 300–2000 *m/z*) were followed by MS/MS fragmentation (normalized collision energy 35) of the top 3 most intense ions. The Q-Exactive Quadrupole-Orbitrap mass spectrometer was coupled to an Easy-nLC<sup>1000</sup> nano HPLC (Thermo Fisher Scientific), and samples were loaded in buffer A (0.1% formic acid) onto a 20-cm-long fused silica capillary column (75  $\mu$ m ID), packed with reversed-phase Repro-Sil Pur C18-AQ 3  $\mu$ m resin (Dr. Maisch GmbH, Ammerbuch, Germany). Peptides were eluted using a 45-min linear gradient from 4 to 40% buffer B (100% ACN, 0.1% formic acid), followed by a 7-min gradient from 40 to 80% buffer B and 8-min wash with 80% B (constant flow rate 300 nL/min). The Q-Exactive was operated in data-dependent acquisition mode with dynamic exclusion enabled (repeat count: 1, exclusion duration: 20 s). Full MS survey scans (mass range 300–1600 *m/z*) at high resolution (70 000 at 200 *m/z*) were followed by MS/MS fragmentation of the top 15 most intense ions with higher energy collisional dissociation at a normalized collision energy of 22 (resolution 17 500 at 200 *m/z*). Dual lock mass calibration was enabled with 371.101233 and 445.120024 *m/z* background ions.

MS spectra were searched against the *R. capsulatus* protein database (<https://www.uniprot.org>, last modified 01/15/2020; *R. capsulatus* (strain ATCC BAA-309/NBRC 16581/SB1003)) using Proteome Discoverer 1.4 (Thermo Fisher Scientific) with Sequest-HT search engine. Search parameters were set to full trypsin digestion, with maxima of three missed cleavages and three modifications per peptide. Oxidation of methionine (+16 Da) and carbamidomethylation of cysteine (+57 Da) were selected as dynamic modifications. Precursor and fragment ion tolerances were set to 2 and 1 Da, respectively, for the LCQ DecaXP+, and 10 ppm and 0.6 Da, respectively, for the Q-Exactive data. False discovery rates (FDRs) by target-decoy search were set to 0.01 and X<sub>corr</sub> filter based on charge (*z*) were: >2 for

$z = 2$ ,  $>2.5$  for  $z = 3$ , and  $>2.6$  for  $z = 4$ <sup>74</sup>. All identifications are listed in Supplementary Table 2.

**Binding of cyt  $c_7$  and soluble cyt S- $c_7$  to purified proteins.** Binding of purified cyt  $c_7$ <sup>75</sup> or a variant of it containing only the soluble cyt domain (cyt S- $c_7$ ) (residues 99–199) to the bipartite SC was assayed by mixing a 5–10-fold molar excess of purified cyt  $c_7$  with 125  $\mu\text{g}$  of purified SC in a final volume of 150  $\mu\text{L}$  TBS buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, and 0.01% DDM). The mixture was separated by chromatography using Superose 6 Increase size exclusion column, and elution fractions were analyzed by SDS-PAGE followed by silver staining. Cyt S- $c_7$  was purified from the *R. capsulatus* strain pYO135/FJ2-R4 (Supplementary Table 1) grown under photosynthetic conditions to maximize its yield<sup>32</sup>. Cells were washed with 20 mM Tris-HCl, pH 8.0, and resuspended 1:5 (w/v) in the same buffer supplemented with 50 mM NaCl. Polymyxin B sulfate (1 mg/mL) was added, and the cell suspension was incubated for 75 min on ice with gentle stirring. After centrifugation for 20 min at  $10,000 \times g$  followed by 3 h at  $150,000 \times g$ , the supernatant was collected and concentrated to a final volume of 2 mL (Amicon Ultra-15 3 kDa MWCO) (Millipore). Aliquots of 300  $\mu\text{L}$  were loaded onto Superose 6 Increase (GE Healthcare) equilibrated in 30 mM Tris-HCl, pH 8.0, and eluted in the same buffer. Fractions eluting after 20 mL were combined and concentrated using Amicon Ultra-15 3 kDa MWCO filters (Millipore).

**Binding of cyt  $c_2$  to purified proteins.** Binding of cyt  $c_2$  to *ccb*<sub>3</sub>-type CIV was determined by mixing 2-fold molar excess of purified *R. capsulatus* cyt  $c_2$ <sup>76</sup> with 300  $\mu\text{g}$  of purified CIV<sup>31</sup> under low-salt conditions (20 mM Tris-HCl, pH 7.4, 1 mM NaCl, and 0.01% DDM). The mixture (250  $\mu\text{L}$  total volume) was loaded onto a Superose 6 Increase (GE Healthcare) sizing column equilibrated with the same buffer to separate the proteins. Elution fractions were concentrated (Amicon Ultra-15, 3 kDa MWCO) (Millipore), and a volume containing 1–5  $\mu\text{g}$  of protein was analyzed by SDS-PAGE followed by silver staining.

**Protein crosslinking using chemical crosslinkers.** For crosslinking, 180  $\mu\text{g}$  of purified bipartite SC at a concentration of 1 mg/mL in phosphate-buffered saline buffer (50 mM Na-Phosphate, pH 7.4, 150 mM NaCl, and 0.01% DDM) mixed with 5-fold molar excess (90  $\mu\text{g}$ ) of purified cyt  $c_7$ <sup>75</sup> were supplemented with 6 mM DSBU (Thermo Fisher Scientific) and incubated on ice for 2 h. The reaction was quenched by adding 50 mM of ammonium bicarbonate, and the mixture was analyzed by MS. In all, 300  $\mu\text{g}$  of purified *ccb*<sub>3</sub>-type CIV<sup>31</sup> were mixed with a 2-fold molar excess of purified cyt  $c_2$ <sup>76</sup> at a final protein concentration of 200  $\mu\text{g}$ /mL under low-salt conditions (20 mM Na-phosphate, pH 7.4, 1 mM NaCl, and 0.01% DDM), and incubated with 20 mM 4-(4,6-Dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride (DMTMM) (Sigma) for 1 h at room temperature. Similarly, 300  $\mu\text{g}$  of purified *bc*<sub>1</sub>-type CIII<sub>2</sub> were mixed with 2-fold molar excess of purified cyt  $c_2$  and treated with DMTMM as above. In both cases, the reaction was stopped by removing excess DMTMM with a PD MiniTrap G-25 desalting column (GE Healthcare), and crosslinked proteins were precipitated with 20% (w/v) trichloroacetic acid (TCA, Sigma) at 4 °C for 1 h. Proteins were pelleted by centrifugation at  $21,000 \times g$  for 15 min and washed with 10% TCA in 0.1 M Tris-HCl and then with acetone (Fisher). The solvent was discarded, and the pellet was air-dried and then stored at  $-80$  °C for analysis by MS.

**MS of crosslinked proteins.** Crosslinked proteins were resuspended in an appropriate volume of solution A (2.5% SDS and 50 mM triethylammonium bicarbonate final concentrations) and reduced with 10 mM DTT (US Biological) for 30 min at 30 °C, followed by alkylation with 50 mM iodoacetamide (Sigma Aldrich) for 30 min at 30 °C. The proteins were processed using an S-Trap<sup>TM</sup> according to the protocol recommended by the supplier (Protifi, C02-mini) and digested with trypsin (Thermo Fisher Scientific) in 1:10 (w/w) enzyme/protein ratio for 1 h at 47 °C. Peptides eluted from this column were vacuum-dried and resuspended with the peptide fractionation-elution buffer for liquid chromatography-MS (LC-MS; (70% (v/v) LC-MS-grade water (Thermo Fisher Scientific), 30% (v/v) ACN (Thermo Fisher Scientific) and 0.1 % (v/v) trifluoroacetic acid (TFA, Thermo Fisher Scientific)). Peptides were fractionated using AKTA Pure 25 with Superdex 30 Increase 3.2/300 (GE Life Science) at a flow rate of 30  $\mu\text{L}/\text{min}$  of the elution buffer, and 100  $\mu\text{L}$  fractions were collected. Based on the elution profile, fractions containing enriched crosslinked peptides of higher molecular masses were vacuum-dried and resuspended with LC-MS-grade water containing 0.1% (v/v) TFA for MS analysis. One half of each fraction was analyzed by a Q-Exactive HF mass spectrometer (Thermo Fisher Scientific) coupled to a Dionex Ultimate 3000 UHPLC system (Thermo Fischer Scientific) equipped with an in-house made 15 cm long fused silica capillary column (75  $\mu\text{m}$  ID), packed with reversed-phase Repro-Sil Pur C18-AQ 2.4  $\mu\text{m}$  resin (Dr. Maisch GmbH, Ammerbuch, Germany). Elution was performed using a gradient from 5 to 45% B (90 min), followed by 90% B (5 min), and re-equilibration from 90 to 5% B (5 min) with a flow rate of 400 nL/min (mobile phase A: water with 0.1% formic acid; mobile phase B: 80% ACN with 0.1% formic acid). Data were acquired in data-dependent MS/MS mode. Full scan MS settings were: mass range 300–1800 *m/z*, resolution 120,000; MS1 AGC target 1E6; MS1 Maximum IT 200. MS/MS settings were: resolution 30,000; AGC target 2E5; MS2 Maximum IT 300 ms; fragmentation was enforced by higher-energy collisional dissociation with stepped collision energy of 25,

27, 30; loop count top 12; isolation window 1.5; fixed first mass 130; MS2 Minimum AGC target 800; charge exclusion: unassigned, 1, 2, 3, 8, and >8; peptide match off; exclude isotope on; dynamic exclusion 45 s<sup>43</sup>. Raw files were converted to mgf format with TurboRawToMGF 2.0.8<sup>77</sup>.

MeroX 2.0.1.4<sup>44,46</sup> (for MS-cleavable and non-cleavable crosslinkers, <http://www.stavrox.com/>) was used as the main search engine for both DMTMM and DSBU. Additionally, FindXL<sup>46,47</sup> (for non-cleavable crosslinkers, <http://biolchem.huji.ac.il/nirka/software.html>) was used to analyze DMTMM spectra, and MassAI 19.07<sup>78</sup> (for MS-cleavable and non-cleavable crosslinkers, <http://www.massai.dk>) was used to analyze DSBU spectra to cross-validate results from MeroX in each case. Only the crosslinked peptides that were cross-validated by two different search engines (i.e., MeroX and FindXL for DMTMM and MeroX and MassAI for DSBU) were used for docking experiments (Supplementary Table 4). MeroX was run in RISEUP mode, with default crosslinker mass and fragmentation parameters for DSBU and in Quadratic mode with default crosslinker mass parameters for DMTMM; precursor mass range, 300–10,000 Da; minimum precursor charge 4; precursor and fragment ion precisions 5.0 and 10.0 ppm, respectively; maximum number of missed cleavages 3; carbamidomethylation of cysteine and oxidation of methionine, as fixed and variable modifications, respectively; results were filtered for score (>10) as well as the MeroX built-in CL spectral matches FDR (<1%). The database consisted of the proteins listed in Supplementary Table 2, supplemented with the assembly factors such as CcoH and the 116 proteins of the cRAP database ([www.thegpm.org](http://www.thegpm.org); contaminant proteins commonly occurring in MS analyses). The latter increased the size of the decoy database to yield more reliable FDRs. FindXL was used to analyze and validate MeroX results for DMTMM CLs. The default FindXL parameters were used as described before<sup>47</sup> with the possible CL amino acids adjustments for K, Y, S, or T on one peptide and E or D on the other peptide. MassAI was used to validate CLs identified by MeroX for DSBU with standard settings, except 5 ppm MS accuracy, 0.05 Da MS/MS accuracy, 3 allowed missed cleavages, and carbamidomethylation of cysteine and oxidation of methionine, as fixed and variable modifications, respectively. Additional filters were: # fragments A  $\geq 4$ , # fragments B  $\geq 4$ , Intensity  $\geq 1000$ , Score  $\geq 10$ . Databases used in FindXL and MassAI consisted of the proteins listed in Supplementary Table 2, supplemented with the assembly factors, such as CcoH. Visualization of the CLs in the structures used Chimera<sup>79</sup> with the Xlink Analyzer plug-in<sup>80</sup>.

**Negative staining and cryo-EM sample preparation.** For negative staining, purified SCs were diluted to 0.01–0.05 mg/mL concentrations in TBS (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.01% DDM, and 1 mM AEBF) buffer, 5  $\mu\text{L}$  applied to glow-discharged (20 s, 25 mA, Pelco easiGlow) carbon-coated Cu grids (CF300-CU, EMS), incubated for 1 min, and stained with 2% uranyl acetate. Grids were imaged on FEI Tecnai 12 transmission electron microscope operating at 120 kV, with a CCD camera (Gatan BM-Ultrascan).

For cryo-EM, the tripartite SC was used as purified, whereas the purified bipartite SC was mixed with a 5-fold molar excess of purified cyt  $c_7$ <sup>75</sup>. In all, 2.5  $\mu\text{L}$  of the mixture, containing 3–5 mg/mL proteins, were applied to CFlat holey carbon grids (1.2/1.3–400 mesh or 2/2–300 mesh) (EMS), which were glow discharged (2 min, 25 mA, Pelco easiGlow) before sample application. Grids were blotted for 9 s at force 0 (CFlat-1.2/1.3) or for 3 s at force  $-5$  (CFlat-2/2) and flash-frozen in liquid ethane cooled with liquid nitrogen using FEI Vitrobot Mark IV (25 °C, 100% humidity). Plunge freezing conditions were optimized using FEI Tecnai TF20 TEM operating at 200 kV, equipped with an FEI Falcon II camera.

**Cryo-EM data acquisition and processing.** All cryo-EM grids were imaged using FEI Titan Krios electron microscope operating at 300 kV. Images of the tripartite SC (SC- $c_7$ ) were recorded using a Gatan K2 Summit direct electron detector, equipped with an energy quantum filter (20 eV), and operated in super-resolution mode at a nominal magnification of  $\times 105,000$ , resulting in a binned pixel size of 1.32 Å. Images were dose-fractionated to 40 frames with a total exposure time of 10 s, and a total dose of  $40 \text{ e}^-/\text{Å}^2$ . Automated data acquisition was carried out using the Latitude software (Gatan), and nominal underfocus values varied from 1 to 2.5  $\mu\text{m}$ . Movies were motion corrected using MotionCor2<sup>81</sup> and CTF parameters were determined with CTFFIND 4.1<sup>82</sup>. About 10,000 particles were manually picked and subjected to an initial reference-free two-dimensional (2D) classification using Relion 3.0<sup>83</sup>. Representative classes were selected and used as a template for auto-picking. After sorting and two rounds of 2D classification,  $\sim 30,000$  particles were retained from the dataset. For 3D classification, an initial model was created using *R. capsulatus* CIII<sub>2</sub> structure (PDB: 1ZRT) and low-pass filtered to 60 Å using EMAN2<sup>84</sup>. A 3D map with a nominal resolution of  $\sim 11$  Å containing a total of  $\sim 12,000$  particles (40%) was obtained, showing a CIII<sub>2</sub> associated on one side with a single copy of CIV. For further analyses, additional datasets were collected using samples from the same batch and identical imaging conditions. From a total of  $\sim 17,000$  images,  $\sim 1,000,000$  particles were automatically picked using the template obtained from the first dataset, and after sorting and two rounds of 2D classification,  $\sim 500,000$  particles were retained. Various subsets of all images were processed separately and yielded the same overall results with slightly different versions of individual maps. For clarity, only the paths to the best representative of each of the final maps are shown (Supplementary Fig. 2a–c). For 3D classification, the map obtained from the first dataset (Supplementary Fig. 2, Box 3) was low-pass filtered to 60 Å and used as the initial model. Of the five classes

obtained, one showed clear features of a CIII<sub>2</sub> dimer associated with a single monomer of CIV (Supplementary Fig. 2, class005 in a, class002 in b and c). The remaining classes showed the same overall shape but lacked density or resolution in different parts of the structure. Only two of the five maps contained a second monomer of CIV associated with the other monomer of CIII<sub>2</sub> (Supplementary Fig. 2c). These two classes showing this feature were combined, yielding a subset of ~220,000 particles. These particles were further processed by another round of 2D classification and subjected to 3D classification, yielding three classes (Supplementary Fig. 2c). Only one of the maps obtained in this second round of 3D classification clearly showed density for a second monomer of CIV. Aligned particles from this group were subclassified into six classes in another round of 3D classification using a soft mask and no image alignment. The low number of particles (~5,000) per class limited the resolution of the maps (<10 Å) and could not be further improved through 3D refinement.

Refinement of the individual maps (Supplementary Fig. 2a, b) led to relatively low-resolution reconstructions, particularly in the CIV portion of the map, probably due to greater structural heterogeneity in the sample with respect to this portion of the tripartite SC. The highest resolution was obtained with a map containing a single copy of CIV by combining all classes showing well-defined features corresponding to it and subjecting them to a second round of 3D classification, focused on CIV portion of the map. For this purpose, a soft mask around CIV was created by fitting the closely related structure of *P. stutzeri* (PDB: 3MK7) into the map and low-pass filtering it to 10 Å. All information outside of this mask was subtracted from the aligned particles, and the remaining particles were subjected to a masked classification without image alignment, to yield six classes (Supplementary Fig. 2a). Alternatively, a cylindrical mask was wrapped around the CIV portion of the map and used in a similar procedure (Supplementary Fig. 2b). In each case, classes showing the highest level of details were retained, and the entire unsubtracted particle dataset was subjected to 3D auto-refinement followed by per-particle CTF refinement, Bayesian polishing, and post-processing using Relion 3.0. The two extreme conformations of the orientation of CIV relative to CIII<sub>2</sub> were represented by the maps SC-1A and SC-1B (Supplementary Fig. 2a, b), which were refined to 6.1 and 7.2 Å resolutions, respectively (Supplementary Table 3).

Images of the bipartite SC supplemented with cyt *c*<sub>γ</sub> (SC + *c*<sub>γ</sub>) were recorded by a Gatan K3 direct electron detector equipped with an energy quantum filter (20 eV) and operated in counting mode at a nominal magnification of ×64,000, corresponding to a pixel size of 1.36 Å. Images were dose-fractionated to 80 frames with a total exposure time of 3.1 s and a total dose of 40 e<sup>-</sup>/Å<sup>2</sup>. Nominal underfocus values varied from 1 to 2.5 μm. Automated data acquisition was carried out using the Latitude software (Gatan) and image shift (~2 μm, 4 images per stage position) was used for accelerated data collection. For this sample, 5480 images were collected in the first session, motion corrected using MotionCor2<sup>81</sup>, and CTF parameters were determined with CTFFIND 4.1<sup>82</sup>. For auto-picking, the template previously obtained with the tripartite SC was used, with this and all subsequent steps done in Relion 3.0<sup>83</sup>. After sorting and two rounds of 2D classification, ~228,000 particles were retained (Supplementary Fig. 5a). The same initial model as for the tripartite SC was used for 3D classification into five classes. Two classes (001 and 002) resembled more to the overall shape of a CIII<sub>2</sub> without CIV, while the other three classes showed the same overall shape of a CIII<sub>2</sub>CIV SC, as seen with the tripartite sample. In all, 37,460 particles corresponding to the SC class with the highest level of detail were combined with SC particles from the second dataset (Supplementary Fig. 5b) to yield the best final map of the bipartite SC. The second dataset consisted of 12,200 images, which were collected and processed under the same conditions as the first dataset. In all, 1.6 million particles were auto-picked, and after 2 rounds of 2D classification, classes were split into ~340,000 large particles likely representing CIII<sub>2</sub>CIV and ~465,000 particles resembling CIII<sub>2</sub> (Supplementary Fig. 5b, d). The former particles were subjected to 3D classification into three classes, and the class most similar to the overall shape of CIII<sub>2</sub>CIV was identified. It contained ~118,000 particles, but due to its low level of detail, it was subjected to another round of 2D classification. The 34,819 particles thus retained were combined with 37,460 SC particles from the first dataset (Supplementary Fig. 5a). After 3D classification into five classes, the class with the highest level of detail contained 56% of particles and clearly showed the overall shape of CIII<sub>2</sub>CIV. In all, 41,017 aligned particles corresponding to this class were extracted and subjected to a second round of 3D classification without image alignment using a soft mask. Of the six 3D classes, the one with the highest nominal resolution contained 14,978 particles (36%) and was subjected to 3D auto-refinement and post-processing, followed by per-particle CTF refinement and Bayesian polishing. After a second round of 3D auto-refinement and post processing, the final map (SC-2A) of the bipartite CIII<sub>2</sub>CIV was obtained at a nominal resolution of 5.18 Å (Supplementary Table 3). This map was very similar to that (SC-1A) of the tripartite CIII<sub>2</sub>CIV, except for its higher resolution and strongly improved density and resolution of the extra TMHs at the interface. Unlike the tripartite samples, no major class corresponding to the conformation seen in SC-1B was identified.

As the bipartite SC samples contained a significant amount of CIII<sub>2</sub> particles without CIV, subsets of smaller 2D classes consistent with a CIII<sub>2</sub> dimer were selected and processed separately. Of the ~376,000 total particles retained from the dataset 1, ~267,000 were identified as resembling CIII<sub>2</sub>. After a second round of 2D classification (Supplementary Fig. 5c), ~213,000 particles were retained and used to generate an ab initio 3D model. This model was low-pass filtered to 60 Å and used

as initial model for the 3D classification of ~465,000 CIII<sub>2</sub> particles from dataset 2 (Supplementary Fig. 5d). Particles corresponding to the class with the highest resolution and level of detail (~185,000) were combined with the CIII<sub>2</sub> particles from dataset 1 to yield ~400,000 particles and subjected to 3D classification. The ~170,000 particles corresponding to the best class were further processed following two different strategies. CIII<sub>2</sub> being a homodimer based on X-ray structures, C2 symmetry was applied in the next round of 3D classification and in all subsequent steps (Supplementary Fig. 5e). The best class contained ~120,000 particles (70%), which were extracted and, similar to CIII<sub>2</sub>CIV, subjected to another round of 3D classification without image alignment using a soft mask. Of the six classes obtained, some showed the ED of both monomers of the FeS protein close to heme *b*<sub>L</sub> (“b position”, b-b) while other classes showed both EDs close to heme *c*<sub>1</sub> (“c position”, c-c). Individual classes were subjected to 3D auto-refinement and post-processing, followed by two rounds of per-particle CTF refinement, Bayesian polishing, 3D auto-refinement, and post-processing. Only the final map with the highest nominal resolution (3.30 Å) is shown (Supplementary Fig. 5e, CIII<sub>2</sub>; Supplementary Table 3). This map contained 37,997 particles and showed both FeS-EDs close to b position but at a lower local resolution and occupancy than the rest of the map, indicating structural flexibility and conformational heterogeneity. A similar approach was used in Supplementary Fig. 5f, except that no C2 symmetry was applied in the two rounds of 3D classifications. Of the three best classes obtained in the second round, one showed both EDs in the b position, one showed both in the c position, and one showed a heterodimeric conformation with one monomer in the b and the other in the c position. Each of the maps was further processed as in Supplementary Fig. 5e, but C2 symmetry was only applied in case of the homodimeric structures. The final maps were CIII<sub>2</sub> c-c with both EDs in the c position at 3.8 Å, CIII<sub>2</sub> b-b with both EDs in the b position at 3.5 Å, and CIII<sub>2</sub> b-c with one ED in the b and the second in the c position at 4.2 Å (Supplementary Table 3). The nominal resolutions thus obtained were slightly lower than in map CIII<sub>2</sub> (Supplementary Fig. 5e). However, by omitting C2 symmetry application to the 3D classifications, the conformational heterogeneity of the FeS protein EDs was resolved, and a subset of particles showing a heterodimeric conformation was identified. The different locations (i.e., b or c positions) of the EDs were clear in the three maps (Supplementary Fig. 5f), but their occupancy and local resolutions was lower than the rest of the structure.

A third dataset was also collected and yielded ~303,000 particles after the first 2D classification. Adding these particles to the first two datasets did not improve the maps shown in Supplementary Fig. 5a, b, but it turned out to be informative. As shown in Supplementary Fig. 5g, particles from dataset 3 were subjected to two rounds of 3D classification. Interestingly, an extra density near the periplasmic domain of CcoP could be seen in some classes. This was not observed in any subclass obtained from the datasets 1 and 2 and tentatively thought to correspond to the cyt domain of cyt *c*<sub>γ</sub>. Focused classification followed by 3D auto-refinement and post-processing, using wider soft masks around the periplasmic domain of CcoP to avoid cutting off any of the extra density, led to a final map (SC-2B) with limited resolution (10.5 Å) and could not be further improved by sub-classification due to the low number of particles in each subclass.

**Refinement of *R. capsulatus* CIII<sub>2</sub> in the cryo-EM maps.** The X-ray based structure of *R. capsulatus* CIII<sub>2</sub> (PDB: 1ZRT) was fitted into map CIII<sub>2</sub> (EMD-22189), which had the highest nominal resolution of all maps obtained (Supplementary Fig. 5e), and refined using Phenix1.16<sup>85</sup>. The real space refinement approach included four rounds of global minimization, local grid search, morphing, and simulated annealing, with the final round also including ADP (*B*-factor) refinement. Each round included five cycles using default settings, and morphing and annealing were performed in each cycle. Secondary structures were determined by Phenix1.16, using default search settings, and restrictions were applied during real space refinement. Due to the low occupancy and limited local resolution, the FeS-ED proteins (residues 50–191) were only subjected to rigid body fitting followed by two rounds of global minimization and local grid search (5 cycles each) but not to morphing and simulated annealing. To ensure the correct cofactor geometry, hemes and [2Fe-2S] clusters including their coordinating residues were copied from the high-resolution X-ray structure of the homologous CIII<sub>2</sub> (cyt *bc*<sub>1</sub>) from *R. sphaeroides* (PDB: 6NHH). Validation was performed using MolProbity<sup>86</sup> (<http://molprobity.biochem.duke.edu/>), and outliers (Ramachandran, rotamer, bonds, angles) were manually corrected in Coot<sup>87</sup>, using real space refinement and regularization. The model that was refined in map CIII<sub>2</sub> (EMD-22189) was subsequently used for rigid body fitting into the maps CIII<sub>2</sub> b-b, CIII<sub>2</sub> c-c, and CIII<sub>2</sub> b-c (Supplementary Fig. 5f and Supplementary Table 3). Each chain was treated as one separate body, except the FeS protein, which was split into its TMH (11–49) and ED (50–191) residues. After the procedure, the linker region between the ED and TMH (residues 40–50) was remodeled in Coot using real space refinement and regularization.

**Modeling of *R. capsulatus* cbb<sub>3</sub>-type CIV subunits.** *R. capsulatus* cbb<sub>3</sub>-type CIV was modeled using *P. stutzeri* cbb<sub>3</sub>-type CIV structure (PDB: 5DJQ or 3MK7) (sequence identities for *R. capsulatus* CcoN: 68%, CcoO: 55%, CcoP: 34%) as a template, and comparative models were computed using MODELLER v9.18<sup>88</sup> with defined secondary structure and crosslinking restraints. The secondary structures for the regions without any template coverage (CcoP residues 104–112, 176–178, 184–197, and 204–215; CcoO residues 172–185 and 192–199), due to the insertion

sequences that are present only in *R. capsulatus* CIV, were estimated using PsiPred Protein Sequence Analysis Workbench<sup>89,90</sup>, and added as restraints to MODELLER. Cross-link distances were added as Gaussian restraints to MODELLER with a mean of 18.0 Å and a standard deviation of 1.0 Å. Problematic loop regions (a total of ten loops longer than four amino acids) were detected by MolProbity<sup>86,91</sup> and remodeled using MODELLER “slow” refinement method (Supplementary Table 3). As *R. capsulatus* CcoN is longer than that of *P. stutzeri* with a predicted extra N-ter TMH, whereas CcoP is shorter than only one predicted N-ter TMH instead of two, these regions were not included into the model. The regions CcoO 179–214 as well as CcoP 1–12, 53–59, 161–173, and 272–280 without template coverage that were not supported by the cryo-EM maps were omitted.

**Modeling of *R. capsulatus* CcoH and cyt  $c_{\gamma}$ .** For CcoH (residues Met 1 to Thr35), five ab initio models were obtained using I-TASSER server<sup>92</sup>. All models were almost identical, and one with the lowest energy score was retained. For *R. capsulatus* cyt  $c_{\gamma}$ , its homologs with known structures were detected using HHpred<sup>93</sup>, and the structure of cyt domain of cyt  $c_{552}$  from *P. denitrificans* (PDB: 3M97; sequence identity: 61%) was used as a template for the soluble cyt domain of  $c_{\gamma}$ , and comparative models were computed using MODELLER<sup>88</sup>.

**Modeling of the extra N-ter TMH of CcoN (TMHO).** A 29-residue model consisting of CcoN residues Arg25 to Asp53, which includes the predicted transmembrane region (Met30-Leu48, UniProtKB, D5ARP4), was obtained from the I-TASSER server<sup>94</sup>. The alpha helical region included residues Leu27 to Thr50, and the model including the residues Arg25 to Leu48 was manually fitted into the density map for visualization purpose in Fig. 2. Related coordinates were not deposited in the PDB because the registration could not be determined due to the lack of side chain density.

**Modeling of cyt  $c_{\gamma}$  and CcoH TMH interactions.** The 30-residue-long cyt  $c_{\gamma}$  TMH was manually docked into the density map by positioning Phe15 and Tyr21 into the corresponding densities and refined by real-space refinement and regularization using Coot. The I-TASSER model of CcoH was docked into the map by moving it along the corresponding density, retaining its close association with cyt  $c_{\gamma}$  TMH. The best fit was found when Ala23 and Val24 of CcoH were located at the interface with cyt  $c_{\gamma}$  TMH. The interaction of the two helices was optimized by GalaxyWEB<sup>95</sup>, and the model with the lowest energy profile was refined in map SC-2A using Phenix 1.17. The refinement strategy included five cycles of global minimization, rigid body fitting (where cyt  $c_{\gamma}$  and CcoH chains were treated as separate bodies), local grid search, and ADP (*B*-factor) refinement with default settings for all steps. Secondary structure restraints were applied to the  $\alpha$ -helical parts of CcoH as predicted by I-TASSER (residues 12–34) and to cyt  $c_{\gamma}$  as predicted by GalaxyWEB (residues 2–9 and 12–29). CIII<sub>2</sub> and CIV models were present during the refinement to keep the rest of the map occupied, but changes made during this procedure were discarded as these models were refined separately.

**Integrative modeling of SC.** The entire SC was assembled by an integrative modeling using the cryo-EM map, XL-MS, co-evolutionary analysis, and subunit models described above. CIII<sub>2</sub> and CIV models were fitted into the cryo-EM maps in UCSF Chimera<sup>79</sup>.

**Docking of CcoH, cyt  $c_{\gamma}$ , and cyt  $c_2$  to bipartite SC.** Additional data about the interaction interfaces between the different subunits were obtained using RaptorX-ComplexContact<sup>33</sup> for each pair of the SC subunits. Based on co-evolution and machine learning, this method predicted the pairs of residues that are in contact. The PatchDock, which is an efficient rigid docking method that maximizes geometric shape complementarity<sup>34</sup>, was used to generate docked configurations of CcoH, cyt  $c_2$ , and cyt  $c_{\gamma}$  to CIII<sub>2</sub> and CIV, as appropriate, and in all docking analyses the top 50 models were retained. The different subunits were docked in parallel independently from each other.

**Docking of CcoH to CIV.** A total of 16 contacts predicted by RaptorX-ComplexContact, with probabilities >0.5 were used as distance restraints for protein–protein docking<sup>34</sup>. The models satisfied 13 of the contacts, and a single cluster of docked models evidenced by the convergence of the top 50 results was obtained (Supplementary Fig. 7). This cluster coincided with the additional feature seen at the SC interface, and CcoH TMH was modeled into this feature.

**Docking of cyt  $c_2$  to CIII<sub>2</sub> and CIV.** *R. capsulatus* cyt  $c_2$  of known structure (PDB: 1C2N) was docked to CIII<sub>2</sub> using the three distance restraints between cyt  $c_2$  and cyt  $c_1$  derived from the protein crosslinking data (Supplementary Table 4). Similarly, cyt  $c_2$  was docked to CIV using nine distance restraints provided by XLS (1 CL to CcoP and 8 to CcoO, Supplementary Table 4). The models yielded one main cluster in each case and covered 100 and 89% of the data for CIII<sub>2</sub> and CIV, respectively.

**Docking of cyt  $c_{\gamma}$  to CIII<sub>2</sub>CIV SC.** The model of cyt domain of  $c_{\gamma}$  was docked by PatchDock to bipartite CIII<sub>2</sub>CIV, which has both FeS-EDs of CIII<sub>2</sub> in the **b** position. Ten distance restraints derived from the XL-MS data (6 with DMTMM with max limit of 30 Å and 4 with DSBU with max limit of 35 Å, Supplementary Table 4) were used, top models yielded two main clusters (1 and 2) of the docked models on the *p* side of CIII<sub>2</sub> (Supplementary Fig. 4), and satisfied 100% of the restraints. As the cryo-EM data revealed that CIII<sub>2</sub> particles could have their FeS-EDs in the **c** position, the corresponding CIII<sub>2</sub> c-c, CIII<sub>2</sub> b-b, and CIII<sub>2</sub> b-c models were used for docking via PatchDock the cyt domain of  $c_{\gamma}$  onto CIII<sub>2</sub>.

All the models were ranked using statistically optimized atomic potentials (SOAP)<sup>96</sup>, and those that have low SOAP scores were retained.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

*R. capsulatus* database used is: <https://www.uniprot.org>, last modified 01/15/2020; *R. capsulatus* strain ATCC BAA-309/NBRC 16581/SB1003. All structural data and corresponding cryo-EM maps that support the findings of this study have been deposited to PDB and EMDB under the accession codes: CIII<sub>2</sub> (PDB: 6XIO; EMD-22189); CIII<sub>2</sub> c-c (PDB: 6XKT; EMD-22224); CIII<sub>2</sub> b-c (PDB: 6XKU; EMD-22225); CIII<sub>2</sub> b-b (PDB: 6XKV; EMD-22226); SC-2A (PDB: 6XKW; EMD-22227); SC-1A (PDB: 6XKX; EMD-22228); SC-1B (PDB: 6XKZ; EMD-22230). XL-MS data have been deposited to ProteomeXchange PRIDE repository with the dataset identifier PXD020038. All other relevant data are available upon request from the corresponding authors. Source data are provided with this paper.

Received: 23 July 2020; Accepted: 6 January 2021;

Published online: 10 February 2021

## References

- Nicholls, D. G. & Ferguson, S. J. *Bioenergetics 4* (Elsevier, 2013).
- Darrouzet, E., Moser, C. C., Dutton, P. L. & Daldal, F. Large scale domain movement in cytochrome *bc*<sub>1</sub>: a new device for electron transfer in proteins. *Trends Biochem. Sci.* **26**, 445–451 (2001).
- Esser, L. et al. Surface-modulated motion switch: capture and release of iron-sulfur protein in the cytochrome *bc*<sub>1</sub> complex. *Proc. Natl Acad. Sci. USA* **103**, 13045–13050 (2006).
- Khalfauoi-Hassani, B. et al. In *Cytochrome Complexes: Evolution, Structures, Energy Transduction, and Signaling* (eds Cramer, W. & Kallas, T.) 527–555 (Springer, 2016).
- Smith, M. A., Finel, M., Korolik, V. & Mendz, G. L. Characteristics of the aerobic respiratory chains of the microaerophiles *Campylobacter jejuni* and *Helicobacter pylori*. *Arch. Microbiol.* **174**, 1–10 (2000).
- Aspholm, M. et al. Structural alterations in a component of cytochrome *c* oxidase and molecular evolution of pathogenic *Neisseria* in humans. *PLoS Pathog.* **6**, e1001055 (2010).
- Ekici, S., Pawlik, G., Lohmeyer, E., Koch, H. G. & Daldal, F. Biogenesis of cbb (3)-type cytochrome *c* oxidase in *Rhodobacter capsulatus*. *Biochim. Biophys. Acta* **1817**, 898–910 (2012).
- Jenney, F. E. Jr. & Daldal, F. A novel membrane-associated *c*-type cytochrome, cyt  $c_{\gamma}$ , can mediate the photosynthetic growth of *Rhodobacter capsulatus* and *Rhodobacter sphaeroides*. *EMBO J.* **12**, 1283–1292 (1993).
- Turba, A., Jetzek, M. & Ludwig, B. Purification of *Paracoccus denitrificans* cytochrome  $c_{552}$  and sequence analysis of the gene. *Eur. J. Biochem.* **231**, 259–265 (1995).
- Albert, I., Rutherford, A. W., Grav, H., Kellermann, J. & Michel, H. The 18 kDa cytochrome  $c_{553}$  from *Helicobacterium gestii*: gene sequence and characterization of the mature protein. *Biochemistry* **37**, 9001–9008 (1998).
- Kim, M. S. et al. Isolation and characterization of a hybrid respiratory supercomplex consisting of *Mycobacterium tuberculosis* cytochrome *bcc* and *Mycobacterium smegmatis* cytochrome *aa*<sub>3</sub>. *J. Biol. Chem.* **290**, 14350–14360 (2015).
- Winstedt, L. & von Wachenfeldt, C. Terminal oxidases of *Bacillus subtilis* strain 168: one quinol oxidase, cytochrome *aa*<sub>3</sub> or cytochrome *bd*, is required for aerobic growth. *J. Bacteriol.* **182**, 6557–6564 (2000).
- Hochkoeppler, A., Jenney, F. E. Jr., Lang, S. E., Zannoni, D. & Daldal, F. Membrane-associated cytochrome  $c_{\gamma}$  of *Rhodobacter capsulatus* is an electron carrier from the cytochrome *bc*<sub>1</sub> complex to the cytochrome *c* oxidase during respiration. *J. Bacteriol.* **177**, 608–613 (1995).
- Enriquez, J. A. Supramolecular organization of respiratory complexes. *Annu. Rev. Physiol.* **78**, 533–561 (2016).

15. Melo, A. M. & Teixeira, M. Supramolecular organization of bacterial aerobic respiratory chains: from cells and back. *Biochim. Biophys. Acta* **1857**, 190–197 (2016).
16. Brzezinski, P. New structures reveal interaction dynamics in respiratory supercomplexes. *Trends Biochem. Sci.* <https://doi.org/10.1016/j.tibs.2019.10.011> (2019).
17. Letts, J. A., Fiedorczuk, K., Degliesposti, G., Skehel, M. & Sazanov, L. A. Structures of respiratory supercomplex I+III2 reveal functional and conformational crosstalk. *Mol. Cell* **75**, 1131–1146 e1136 (2019).
18. Quintana-Cabrera, R. & Soriano, M. E. ER stress priming of mitochondrial respiratory supercomplex assembly. *Trends Endocrinol. Metab.* **30**, 685–687 (2019).
19. Letts, J. A., Fiedorczuk, K. & Sazanov, L. A. The architecture of respiratory supercomplexes. *Nature* **537**, 644–648 (2016).
20. Gu, J. et al. The architecture of the mammalian respirasome. *Nature* **537**, 639–643 (2016).
21. Hartley, A. M. et al. Structure of yeast cytochrome *c* oxidase in a supercomplex with cytochrome *bc*<sub>1</sub>. *Nat. Struct. Mol. Biol.* **26**, 78–83 (2019).
22. Berry, E. A. & Trumpower, B. L. Isolation of ubiquinol oxidase from *Paracoccus denitrificans* and resolution into cytochrome *bc*<sub>1</sub> and cytochrome *c-aa*<sub>3</sub> complexes. *J. Biol. Chem.* **260**, 2458–2467 (1985).
23. Kao, W. C. et al. The obligate respiratory supercomplex from *Actinobacteria*. *Biochim. Biophys. Acta* **1857**, 1705–1714 (2016).
24. Gong, H. et al. An electron transfer path connects subunits of a mycobacterial respiratory supercomplex. *Science*. <https://doi.org/10.1126/science.aat8923> (2018).
25. Wiseman, B. et al. Structure of a functional obligate complex III<sub>2</sub>IV<sub>2</sub> respiratory supercomplex from *Mycobacterium smegmatis*. *Nat. Struct. Mol. Biol.* **25**, 1128–1136 (2018).
26. Ducluzeau, A. L., Ouchane, S. & Nitschke, W. The *cbb*<sub>3</sub> oxidases are an ancient innovation of the domain bacteria. *Mol. Biol. Evol.* **25**, 1158–1166 (2008).
27. Myllykallio, H., Drepper, F., Mathis, P. & Daldal, F. Electron-transfer supercomplexes in photosynthesis and respiration. *Trends Microbiol.* **8**, 493–494 (2000).
28. Davidson, E., Ohnishi, T., Tokito, M. & Daldal, F. *Rhodobacter capsulatus* mutants lacking the Rieske FeS protein form a stable cytochrome *bc*<sub>1</sub> subcomplex with an intact quinone reduction site. *Biochemistry* **31**, 3351–3358 (1992).
29. Kulajta, C., Thumfart, J. O., Haid, S., Daldal, F. & Koch, H. G. Multi-step assembly pathway of the *cbb*<sub>3</sub>-type cytochrome *c* oxidase complex. *J. Mol. Biol.* **355**, 989–1004 (2006).
30. Valkova-Valchanova, M. B., Saribas, A. S., Gibney, B. R., Dutton, P. L. & Daldal, F. Isolation and characterization of a two-subunit cytochrome *b-c*<sub>1</sub> subcomplex from *Rhodobacter capsulatus* and reconstitution of its ubihydroquinone oxidation (Q<sub>o</sub>) site with purified Fe-S protein subunit. *Biochemistry* **37**, 16242–16251 (1998).
31. Gray, K. A. et al. *Rhodobacter capsulatus* contains a novel *cb*-type cytochrome *c* oxidase without a Cu<sub>A</sub> center. *Biochemistry* **33**, 3120–3127 (1994).
32. Ozturk, Y. et al. Soluble variants of *Rhodobacter capsulatus* membrane-anchored cytochrome *c*<sub>7</sub> are efficient photosynthetic electron carriers. *J. Biol. Chem.* **283**, 13964–13972 (2008).
33. Zeng, H. et al. ComplexContact: a web server for inter-protein contact prediction using deep learning. *Nucleic Acids Res.* **46**, W432–W437 (2018).
34. Schneidman-Duhovny, D., Inbar, Y., Nussinov, R. & Wolfson, H. J. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.* **33**, W363–W367 (2005).
35. Pawlik, G. et al. The putative assembly factor CcoH is stably associated with the *cbb*<sub>3</sub>-type cytochrome oxidase. *J. Bacteriol.* **192**, 6378–6389 (2010).
36. Covian, R. & Trumpower, B. L. Rapid electron transfer between monomers when the cytochrome *bc*<sub>1</sub> complex dimer is reduced through center N. *J. Biol. Chem.* **280**, 22732–22740 (2005).
37. Castellani, M. et al. Direct demonstration of half-of-the-sites reactivity in the dimeric cytochrome *bc*<sub>1</sub> complex: enzyme with one inactive monomer is fully active but unable to activate the second ubiquinol oxidation site in response to ligand binding at the ubiquinone reduction site. *J. Biol. Chem.* **285**, 502–510 (2010).
38. Cooley, J. W., Lee, D. W. & Daldal, F. Across membrane communication between the Q<sub>o</sub> and Q<sub>i</sub> active sites of cytochrome *bc*<sub>1</sub>. *Biochemistry* **48**, 1888–1899 (2009).
39. Chavez, J. D. et al. Chemical crosslinking mass spectrometry analysis of protein conformations and supercomplexes in heart tissue. *Cell Syst.* **6**, 136.e5–141.e5 (2018).
40. Chavez, J. D. & Bruce, J. E. Chemical cross-linking with mass spectrometry: a tool for systems structural biology. *Curr. Opin. Chem. Biol.* **48**, 8–18 (2019).
41. Larance, M. et al. Global membrane protein interactome analysis using in vivo crosslinking and mass spectrometry-based protein correlation profiling. *Mol. Cell. Proteomics* **15**, 2476–2490 (2016).
42. Leitner, A., Faini, M., Stengel, F. & Aebersold, R. Crosslinking and mass spectrometry: an integrated technology to understand the structure and function of molecular machines. *Trends Biochem. Sci.* **41**, 20–32 (2016).
43. Slavin, M. & Kalisman, N. Structural analysis of protein complexes by cross-linking and mass spectrometry. *Methods Mol. Biol.* **1764**, 173–183 (2018).
44. Gotze, M. et al. Automated assignment of MS/MS cleavable cross-links in protein 3D-structure analysis. *J. Am. Soc. Mass Spectrom.* **26**, 83–97 (2015).
45. Solmaz, S. R. & Hunte, C. Structure of complex III with bound cytochrome *c* in reduced state and definition of a minimal core interface for electron transfer. *J. Biol. Chem.* **283**, 17542–17549 (2008).
46. Iacobucci, C. et al. A cross-linking/mass spectrometry workflow based on MS-cleavable cross-linkers and the MeroX software for studying protein structures and protein-protein interactions. *Nat. Protoc.* **13**, 2864–2889 (2018).
47. Kalisman, N., Adams, C. M. & Levitt, M. Subunit order of eukaryotic TrIC/CCT chaperonin by cross-linking, mass spectrometry, and combinatorial homology modeling. *Proc. Natl Acad. Sci. USA* **109**, 2884–2889 (2012).
48. Osyczka, A., Moser, C. C. & Dutton, P. L. Fixing the Q cycle. *Trends Biochem. Sci.* **30**, 176–182 (2005).
49. Crofts, A. R. et al. The Q-cycle reviewed: how well does a monomeric mechanism of the bc(1) complex account for the function of a dimeric complex? *Biochim. Biophys. Acta* **1777**, 1001–1019 (2008).
50. Buschmann, S. et al. The structure of *cbb*<sub>3</sub> cytochrome oxidase provides insights into proton pumping. *Science* **329**, 327–330 (2010).
51. Myllykallio, H., Drepper, F., Mathis, P. & Daldal, F. Membrane-anchored cytochrome *c*<sub>7</sub> mediated microsecond time range electron transfer from the cytochrome *bc*<sub>1</sub> complex to the reaction center in *Rhodobacter capsulatus*. *Biochemistry* **37**, 5501–5510 (1998).
52. Lee, D. W., Ozturk, Y., Osyczka, A., Cooley, J. W. & Daldal, F. Cytochrome *bc*<sub>1</sub>-*c*<sub>7</sub> fusion complexes reveal the distance constraints for functional electron transfer between photosynthesis components. *J. Biol. Chem.* **283**, 13973–13982 (2008).
53. Berry, E. A. et al. X-ray structure of *Rhodobacter capsulatus* Cytochrome *bc*<sub>1</sub>: comparison with its mitochondrial and chloroplast counterparts. *Photosynth. Res.* **81**, 251–275 (2004).
54. Xia, D. et al. The road to the crystal structure of the cytochrome *bc*<sub>1</sub> complex from the anoxygenic, photosynthetic bacterium *Rhodobacter sphaeroides*. *J. Bioenerg. Biomembr.* **40**, 485–492 (2008).
55. Esser, L. et al. Inhibitor-complexed structures of the cytochrome *bc*<sub>1</sub> from the photosynthetic bacterium *Rhodobacter sphaeroides*. *J. Biol. Chem.* **283**, 2846–2857 (2008).
56. Sousa, J. S., Mills, D. J., Vonck, J. & Kuhlbrandt, W. Functional asymmetry and electron flow in the bovine respirasome. *Elife*. <https://doi.org/10.7554/eLife.21290> (2016).
57. Swierczek, M. et al. An electronic bus bar lies in the core of cytochrome *bc*<sub>1</sub>. *Science* **329**, 451–454 (2010).
58. Lanciano, P., Lee, D. W., Yang, H., Darrouzet, E. & Daldal, F. Intermonomer electron transfer between the low-potential *b* hemes of cytochrome *bc*<sub>1</sub>. *Biochemistry* **50**, 1651–1663 (2011).
59. Moser, C. C., Keske, J. M., Warncke, K., Farid, R. S. & Dutton, P. L. Nature of biological electron transfer. *Nature* **355**, 796–802 (1992).
60. Wikstrom, M., Krab, K. & Sharma, V. Oxygen activation and energy conservation by cytochrome *c* oxidase. *Chem. Rev.* **118**, 2469–2490 (2018).
61. Brzezinski, P. & Gennis, R. B. Cytochrome *c* oxidase: exciting progress and remaining mysteries. *J. Bioenerg. Biomembr.* **40**, 521–531 (2008).
62. Verissimo, A. F., Sousa, F. L., Baptista, A. M., Teixeira, M. & Pereira, M. M. Thermodynamic redox behavior of the heme centers of *cbb*<sub>3</sub> heme-copper oxygen reductase from *Bradyrhizobium japonicum*. *Biochemistry* **46**, 13245–13253 (2007).
63. Otten, M. F. et al. Cytochromes *c*<sub>550</sub>, *c*<sub>552</sub>, and *c*<sub>1</sub> in the electron transport network of *Paracoccus denitrificans*: redundant or subtly different in function? *J. Bacteriol.* **183**, 7017–7026 (2001).
64. Kappler, U. & Dahl, C. Enzymology and molecular biology of prokaryotic sulfite oxidation. *FEMS Microbiol. Lett.* **203**, 1–9 (2001).
65. Khalfaoui-Hassani, B. et al. Widespread distribution and functional specificity of the copper importer CcoA: distinct Cu uptake routes for bacterial cytochrome *c* oxidases. *MBio*. <https://doi.org/10.1128/mBio.00065-18> (2018).
66. Wittig, I., Braun, H. P. & Schagger, H. Blue native PAGE. *Nat. Protoc.* **1**, 418–428 (2006).
67. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685 (1970).
68. Bartsch, H., Arndt, C., Koristka, S., Cartellieri, M. & Bachmann, M. Silver staining techniques of polyacrylamide gels. *Methods Mol. Biol.* **869**, 481–486 (2012).
69. Yan, L. J. & Forster, M. J. Resolving mitochondrial protein complexes using nongradient blue native polyacrylamide gel electrophoresis. *Anal. Biochem.* **389**, 143–149 (2009).
70. Thomas, P. E., Ryan, D. & Levin, W. An improved staining procedure for the detection of the peroxidase activity of cytochrome P-450 on sodium dodecyl sulfate polyacrylamide gels. *Anal. Biochem.* **75**, 168–176 (1976).
71. Gill, S. C. & von Hippel, P. H. Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319–326 (1989).
72. Atta-Asafo-Adjei, E. & Daldal, F. Size of the amino acid side chain at position 158 of cytochrome *b* is critical for an active cytochrome *bc*<sub>1</sub> complex and for

- photosynthetic growth of *Rhodobacter capsulatus*. *Proc. Natl Acad. Sci. USA* **88**, 492–496 (1991).
73. Peters, A., Kulajta, C., Pawlik, G., Daldal, F. & Koch, H. G. Stability of the *cbb<sub>3</sub>*-type cytochrome oxidase requires specific CcoQ-CcoP interactions. *J. Bacteriol.* **190**, 5576–5586 (2008).
  74. Selamoglu, N. et al. Comparative differential cuproproteomes of *Rhodobacter capsulatus* reveal novel copper homeostasis related proteins. *Metallomics* **12**, 572–591 (2020).
  75. Myllykallio, H., Jenney, F. E. Jr., Moomaw, C. R., Slaughter, C. A. & Daldal, F. Cytochrome *c<sub>v</sub>* of *Rhodobacter capsulatus* is attached to the cytoplasmic membrane by an uncleaved signal sequence-like anchor. *J. Bacteriol.* **179**, 2623–2631 (1997).
  76. Holden, H. M., Meyer, T. E., Cusanovich, M. A., Daldal, F. & Rayment, I. Crystallization and preliminary analysis of crystals of cytochrome *c<sub>2</sub>* from *Rhodospseudomonas capsulata*. *J. Mol. Biol.* **195**, 229–231 (1987).
  77. Sheng, Q. et al. Preprocessing significantly improves the peptide/protein identification sensitivity of high-resolution isobarically labeled tandem mass spectrometry data. *Mol. Cell. Proteomics* **14**, 405–417 (2015).
  78. Rutsdottir, G. et al. Structural model of dodecameric heat-shock protein Hsp21: Flexible N-terminal arms interact with client proteins while C-terminal tails maintain the dodecamer and chaperone activity. *J. Biol. Chem.* **292**, 8103–8121 (2017).
  79. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
  80. Kosinski, J. et al. Xlink Analyzer: software for analysis and visualization of cross-linking data in the context of three-dimensional structures. *J. Struct. Biol.* **189**, 177–183 (2015).
  81. Zheng, S. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
  82. Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
  83. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
  84. Tang, G. et al. EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
  85. 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).
  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. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
  88. Webb, B. & Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinformatics* **47**, 5 6 1–32 (2014).
  89. Buchan, D. W. A. & Jones, D. T. The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Res.* **47**, W402–W407 (2019).
  90. Jones, D. T., Tress, M., Bryson, K. & Hadley, C. Successful recognition of protein folds using threading methods biased by sequence similarity and predicted secondary structure. *Proteins (Suppl 3)*, 104–111 (1999).
  91. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).
  92. Yang, J. et al. The I-TASSER Suite: protein structure and function prediction. *Nat. Methods* **12**, 7–8 (2015).
  93. Soding, J. Protein homology detection by HMM-HMM comparison. *Bioinformatics* **21**, 951–960 (2005).
  94. Zhang, W. et al. Integration of QUARK and I-TASSER for ab initio protein structure prediction in CASP11. *Proteins* **84**(Suppl 1), 76–86 (2016).
  95. Ko, J., Park, H., Heo, L. & Seok, C. GalaxyWEB server for protein structure prediction and refinement. *Nucleic Acids Res.* **40**, W294–W297 (2012).
  96. Dong, G. Q., Fan, H., Schneidman-Duhovny, D., Webb, B. & Sali, A. Optimized atomic statistical potentials: assessment of protein interfaces and loops. *Bioinformatics* **29**, 3158–3166 (2013).

## Acknowledgements

This work was supported partly by the NIH grants GM 38237 to F.D., GM123233 to K.M., GM110174 and AI118891 to B.A.G., T32-GM008275 to T.V., T32-GM071339 to H.J.K., and partly by the Division of Chemical Sciences, Geosciences and Biosciences, Office of Basic Energy Sciences of Department of Energy grant DE-FG02-91ER20052 to F.D., by ISF 1466/18, BSF 2016070, and Ministry of Science and Technology 80802 grants to D.S.-D. Y. O. was supported by the grant GRK2202-23577276/RTG from DFG, Germany. Data analysis was partly supported by the NIH grant S10OD023592. This research was supported in part by the NCI, National Cryo-EM Facility at the Frederick National Laboratory for Cancer Research under contract HSSN261200800001E. The authors thank Ulrich Baxa, Thomas Edwards, and Adam Wier for their support and helpful discussions. Some cryo-EM data were also obtained at the University of Massachusetts Cryo-EM Core Facility, and we thank Dr. Chen Xu, Dr. KangKang Song, and Dr. Kyoungwan Lee for their support. Cryo-EM sample screening and optimization was performed at the Electron Microscopy Resource Laboratory at the Perelman School of Medicine, University of Pennsylvania, and we thank Dr. Sudheer Molugu for his support. We also thank Dr. S. Saif Hasan, Dr. Brian G. Pierce, and Dr. Christian Presley at the Institute for Bioscience and Biotechnology Research, University of Maryland for insightful discussions and invaluable help they provided during this study. S.S. and F.D. thank Vivian Kitainda for her assistance with protein purification and O<sub>2</sub> consumption measurements.

## Author contributions

All authors have given approval to the final version of the manuscript. S.S., T.V., Y.O., H.J.K., M.B., N.S., B.A.G., D.S.-D., K.M., and F.D. all contributed to aspects of the experiments, analyzed data, and wrote and edited the manuscript. S.S. purified, characterized, and prepared samples; S.S. and T.V. processed data and built models; Y.O. performed genetic constructs; M.B. and D.S.-D. conducted computer modeling; H.J.K., N.S., and B.A.G. obtained MS data; and K.M. and F.D. managed the overall project and supervised the study.

## 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/s41467-021-21051-4>.

**Correspondence** and requests for materials should be addressed to K.M. or F.D.

**Peer review information** *Nature Communications* thanks the anonymous reviewers for their contributions to the peer review of this work. Peer review 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) 2021