Supplementary Material

Structural insights into electron transfer in $ca_{a3}$-type cytochrome oxidase

Joseph A. Lyons$^1$, $^3$, David Aragão$^3$, $^5$, Orla Slattery$^1$, Andrei V. Pisliakov$^6$, Tefik Soulimane$^{1,2,*}$ and Martin Caffrey$^{3,4,*}$

Department of Chemical and Environmental Sciences$^1$ & Materials and Surface Science Institute$^2$, University of Limerick, Limerick, Ireland.
School of Biochemistry and Immunology$^3$ & School of Medicine$^4$, Trinity College Dublin, Dublin, Ireland.
Australian Synchrotron$^5$, 800 Blackburn Road, Clayton, VIC 3168, Australia
Theoretical Biochemistry Laboratory$^6$, RIKEN Advanced Science Institute, Wako, Saitama, Japan.

Correspondence and requests for materials should be addressed to T. S. (tewfik.soulimane@ul.ie) or M.C (martin.caffrey@tcd.ie).
Supplementary Discussion

1. Lipid embedded in SU I/III
At the interface between the fused SU I and III domains, density for a glycoglycerophospholipid was found. It has been modeled as a 2′-O-(1,2-diacyl-sn-glycero-3-phospho)-3′-O-(α-N-acetyl-glucosaminyl)-N-glyceroyl alkyl-amide (Supplementary Fig. 6) following an analysis of similar lipids from *Thermus* and *Meiothermus*\(^1\). The hydrocarbon moieties have been modeled as a pair of iso-C17:0 acyl chains and a C19:0 N-alkyl amide based on the most abundant hydrocarbons reported at growth temperatures above 75 °C\(^1\). In Type A1 HCOs, phosphatidylethanolamine (*R. sphaeroides*, PDB ID: 1M56), phosphatidylcholine (*P. denitrificans*, PDB ID: 1QLE) and phosphatidylglycerol (bovine heart, PDB ID: 3AG1) occupied the internal lipid cavity. The three acyl chains of the glycoglycerophospholipid in *caaa*<sub>3</sub>-oxidase fill most of the space occupied by the chains of the two glycerophospholipid molecules reported for the mitochondrial and bacterial *caaa*<sub>3</sub>-oxidases. The helical packing at the cytoplasmic side of the SU III domain in *caaa*<sub>3</sub>-oxidase is tighter in comparison with that seen in the mesophilic Type A HCOs (Supplementary Fig. 7). This may reflect a more extensive interaction between the SU III domain and the native glycoglycerophospholipid headgroup and the thermophilic nature of the protein. Since the oxidase must function naturally at high temperatures, these interactions may contribute to the over stability of the enzyme.

2. Ligand assignments in active site
Given the large X-ray scattering from the metal cofactors in the binuclear centre as well as the fact that metal centers are readily reduced by synchrotron X-radiation, interpretation of electron density attributable to the bridging ligands is tentative. The hydroxide ion and water molecule modeled in the active site (Fig. 2a, b) are consistent with proposed catalytic intermediates as determined by EXAFS and ENDOR measurements\(^2,3\) and have been described in other structures\(^4\). Recent quantum chemical calculations and a 2 Å resolution crystallographic analysis of the bovine heart cytochrome oxidase suggest that the oxygenous species between the metal centers has superoxide character and is derived from a diatomic oxygen species present in the active site\(^5\).
3. Oxygen channel

To identify possible oxygen channels, internal hydrophobic cavities were calculated for \( caa_3 \) (PDB ID: 2YEV), \( ba_3 \) (PDB ID: 1EHK) and \( aa_3 \)-type oxidases from \( R. sphaeroides \) (PDB ID: 2GSM) and \( P. denitrificans \) (PDB ID: 3HB3) using the program Hollow\(^6\). In \( caa_3 \)-oxidase the analysis identified a continuous Y-shaped channel that leads to the active site (Fig. 2d). The bifurcated channel, originating at two points on the apolar protein exterior, converges and terminates at the active site. A comparison of the computed oxygen channel in \( caa_3 \)-oxidase with those in Type A1 and B HCOs reveals that the channel is conserved, with differences in size and geometry highlighted in Supplementary Fig. 9. The presence of two apparent oxygen entry sites in \( ba_3 \) and \( caa_3 \)-oxidase (Supplementary Fig. 9a, b) suggests a common mechanism for oxygen uptake in thermophilic HCOs. However, the hydrophobic region of the channel upstream of the active site is more constricted in Type A than in Type B HCOs (Supplementary Fig. 9a-d). This observation supports a mechanism of side chain mobility in Type A oxidases which regulates oxygen delivery to the binuclear centre, as described previously\(^7,8\).

4. Proton pathways

The D-pathway consists of a chain of ordered internal waters complete with orienting polar side chains extending from Asp 103 at the cytoplasmic side of the membrane to the active site via the YS “gate” (Fig. 3a, b). The ordered water chain of the D-pathway in \( caa_3 \)-oxidase largely resembles that found in both the \( P. denitrificans \) (PDB ID: 3HB3) and \( R. sphaeroides \) (PDB ID: 2GSM) oxidases (Supplementary Fig. 10a). Electron density for two conformations of Asp 103 was found and the alternative conformations were modeled (Fig. 3a).

The K-pathway of \( caa_3 \)-oxidase (Fig. 3c), starting at residue Glu 84(IIC), is centered on a conserved lysine, Lys 328, and is consistent with the pathway described for other Type A1 oxidases (Supplementary Fig. 10b). However, in \( caa_3 \)-oxidase the \( \varepsilon \)-amine of Lys 328 is hydrogen bonded to Gln 286, replacing a water mediated hydrogen bond between the \( \varepsilon \)-amine of lysine and the hydroxyl of a serine found in Type A1 HCOs (Lys 362 and Ser 299 \( R. sphaeroides \) numbering). The lysine-glutamine interaction does not appear to be a hallmark of Type A2 oxidases because the \( caa_3 \)-oxidase from \( Rh. marinus \) contains a serine at the glutamine locus. Further, the canonical serine involved in water mediated interaction with the
lysine, viz., Ser 299 in *R. sphaeroides*, in the K-pathway of Type A1 oxidases is conserved in *caa3*-oxidase from *Rh. marinus*. The bridging water in Type A1 HCOs is proposed to stabilize the protonated form of the lysine side chain at neutral pH in the oxidized enzyme\(^9\). We propose that the Lys 328 - Gln 286 pairing in *caa3*-oxidase is functionally equivalent to the lysine-water-serine interaction in Type A1 HCOs (Supplementary Table 2). This warrants functional and mutagenesis studies as per Brändén *et al*\(^9\). Of note too is the conformation of the Met 282 side chain in *caa3*-oxidase which points away from the proton entry site. This orientation could provide the Lys 328 side chain access to the entrance of the K-pathway via Asn 331. In the case of Type A1 oxidases, the corresponding methionine appears to extend into this space (Supplementary Fig. 10b). Molecular dynamics simulations (MDS) confirm and extend the experimentally defined water distributions and the proposed proton wires in the D- and K-pathways (Supplementary Fig. 11).

5. Exit pathways
A conserved water cluster, serving as the proton exit site to the periplasmic space, is located above the heme $a_3$ propionates (Fig. 2c). It is connected directly via hydrogen bonds to a central magnesium which is coordinated by three water molecules and by the side chains of Glu 299(Ic), His 377 and Asp 378. The cluster, which consists of seventeen waters, is similar to that observed in Type A1 cytochrome oxidases. From the water pool, several water chains can be identified leading to the periplasm. They mark routes by which pumped protons can exit the protein. MDS identifies Arg 173(IIC) as the most probable exit site on the *caa3*-oxidase surface (Supplementary Fig. 12). A similar conduit has been described in Type A1 HCOs\(^10\).

6. Cytochrome c domain comparision
The crystal structure of the recombinantly expressed soluble cytochrome c domain of the *caa3*-oxidase from *Rh. marinus* (Cyt $c_{R,m}$) revealed a novel cytochrome c fold\(^11\). It includes a two-stranded anti-parallel $\beta$-sheet (residues 41-44 and 49-52, PDB ID: 1W2L) which interacts with the A-propionate of heme $c_{11}$ (Supplementary Fig. 14b). Next to the $\beta$-sheet is a loop (Loop D, residues 64-74, PDB ID: 1W2L) which covers the D-propionate. In this case, the D-propionate is bent back on itself with the carboxyl oriented in the direction of the methionine that ligates the heme iron. Homology modeling suggested that the same fold was
present in the cytochrome $e$ domain of the $T. \text{thermophilus caa}_3$-oxidase (Cyt $c_{T.t}$)\textsuperscript{11}. While the distinguishing features of a $\beta$-sheet and Loop D are recognizable in Cyt $c_{T.t}$, the two cytochrome $c$ domains differ significantly in detail, (Supplementary Fig. 14c, d). By comparison, the structure reported here for Cyt $c_{T.t}$ is of the domain in its native orientation next to its cupredoxin partner in the oxidase complex. Thus, the shorter $\beta$-sheet (residues 274-275(IIc) and 280-281(IIc)) and Loop D (residues 295-302 (IIc)) and their relative dispositions in Cyt $c_{T.t}$ presumably facilitate direct contact between the propionates of heme $c$ and the cupredoxin domain for efficient electron transfer through the enzyme.

7. Proposed electron entry mechanism.
Electron transfer between the cytochrome $bc$ complex and $caa_3$-oxidase of $T. \text{thermophilus}$ has been proposed to occur directly without a mobile electron carrier\textsuperscript{12}. However, because the structure of the $\text{Thermus bc}$ oxido-reductase is not known, speculation as to the nature of a $bc$-$caa_3$-oxidase ET complex is not well founded. Structurally, the exposed heme edge in the cytochrome $c$ domain of $caa_3$-oxidase is the most probable region for electron entry into the enzyme. The surface around the heme edge does not have a marked polar character (Fig. 4c) suggesting a hydrophobic interaction with its primary electron donor. Cytochrome $c_{552}$ is a mobile electron carrier in $T. \text{thermophilus}$. It has been shown to efficiently reduce the soluble cytochrome $c$ fragment of $caa_3$-oxidase\textsuperscript{12}. An apolar surface surrounds the exposed heme edge of cytochrome $c_{552}$\textsuperscript{13}. This may provide a complementary site for heme edge-to-heme edge interaction facilitating ET between the cytochrome $c_{552}$ - $caa_3$-oxidase redox pair, in agreement with previous analyses\textsuperscript{14,15}. 
Supplementary Figure 1. Structure based sequence alignment for SU I of Type A cytochrome oxidases; \textit{caa}_3\textsuperscript{-} (\textit{T. thermophilus}, this work, residues 12-516), \textit{aa}_3\textsuperscript{-} (\textit{P. denitrificans}, PDB ID: 3HB3, residues 17-545), \textit{aa}_3\textsuperscript{-} (\textit{R. sphaeroides}, PDB ID: 2GSM, residues 17-551) and \textit{aa}_3\textsuperscript{-} (bovine heart, PDB ID: 1V54, residues 2-514) oxidases. Prosthetic group ligands are indentified below the residues by a boldface \textit{a}, \textit{a}3 and \textit{Cu} for heme \textit{a}, heme \textit{a}3 and Cu\textsubscript{B}, respectively. Residues implicated in the D- and K-proton pathways are labeled, respectively, with boldface D and K. The background colors indicate the sequence conservation among the four structures ranging from black (highly conserved) to white (poor conservation). Residue numbering for the different oxidases is arranged along the right hand border of the figure. The \textit{caa}_3\textsuperscript{-}-oxidase shares a structure based sequence identity of 44, 43 and 38 \% with the \textit{aa}_3\textsuperscript{-}-oxidases of \textit{P. denitrificans}, \textit{R. sphaeroides} and bovine heart, as calculated in Modeller. Percentage sequence identity was calculated from the number of identical residues/minimum length.
Supplementary Figure 2. Crystals of caa₃-oxidase grown by the in meso method using 7.7 MAG as the hosting lipid as viewed under normal (a-c) and cross-polarized light (d). Pictures were taken after 7-10 days at 20 °C. The precipitant contained 14-21 % (v/v) polyethyleneglycol 400, 0.1 M NaCl, 0-0.1 M Li₂SO₄ and 0.1 M sodium citrate pH 4.5-5.0. Full details are described under Methods.
Supplementary Figure 3. Packing arrangement in crystals of caax-oxidase. View perpendicular to a, the ab, b, the bc and c, the ac faces of the unit cell. Symmetry related protein molecules are represented in blue and red. The unit cell boundaries are outlined in black. Packing within the crystal is of Type 1\(^{16}\) and is consistent with other structures solved with crystals grown by the in meso method.
Supplementary Figure 4. A wall-eyed stereo view of a section of the SU I/III helix VI (residues 246-267) with its 2mFo-DFc electron density contoured at 1σ (grey mesh) of the final refined model. The protein is shown in sticks with atomic coloring. Density for the crosslink between Tyr 254 and His 250 is clearly visible at the upper right of the helix.
Supplementary Figure 5. Amino acid sequence of SU IV of caa3-oxidase determined using a combination of N- and C-terminal sequencing and mass spectral analysis. Mass spectrometry was used to confirm the calculated molecular weight of the entire subunit as 7,347 Da. N-terminal sequencing of the first 41 residues was performed on the deformylated polypeptide and C-terminal sequencing of the last 7 residues was carried out on the intact subunit. Cyanogen bromide treatment of SU IV generated three fragments, CNBr-1 (orange), CNBr-2 (blue) and CNBr-3 (green). The fragments were separated by HPLC and individually sequenced. Full details are described under Methods. For reference, the genes encoding for SU IV (locus tag: TTHA1863) and for SU I/III and IIc (locus tags: TTHA0312 and TTHA0311) are at different loci in the T. thermophilus genome.
Supplementary Figure 6. A novel native thermophilic glycoglycerophospholipid identified in SU I/III of caa3-oxidase. a, The electron density, contoured at 1σ in blue, attributed to the lipid is located at the interface of the SU I and SU III domains in SU I/III and has been modeled as a 2'-O-(1,2-diacyl-sn-glycero-3-phospho)-3’-O-(α-N-acetyl-glucosaminyl)-N-glyceroyl alkyl-amide with two iso-C17:0 acyl chains and a C19:0 N-alkyl amide. b, An expanded view of the headgroup region of the lipid and its coordination by residues in SU I/III and by structured waters. c, Chemical structure of the lipid headgroup with truncated chains is shown with the molecule in approximately the same orientation as that in a and b.
Supplementary Figure 7. A comparison of the transmembrane helices surrounding the native glycerolipid in \textit{caa}_{3}\text{-oxidase} with equivalent helices in \textit{aa}_{3}\text{-oxidase} from \textit{P. denitrificans}. \textbf{a}, View from the cytoplasm. \textbf{b}, View from the periplasm. Helices are displayed as rods with the loops removed for clarity. The transmembrane helices in the SU I domain of SU I/III in \textit{caa}_{3}\text{-oxidase} have a very similar structure and arrangement to those in SU I of \textit{aa}_{3}\text{-oxidase} with a r.m.s.d of 0.99, 0.92 and 1.09 Å for \textit{P. denitrificans}, \textit{R. sphaeroides} and bovine heart, respectively. For simplicity, only the two helices (TM3, TM4) in SU I contacting the lipid (grey spheres) are shown here. A comparison of the helices in the SU III domain of \textit{caa}_{3}\text{-oxidase} (blue rods, TM13-19) with those in SU III of \textit{aa}_{3}\text{-oxidase} (orange rods) reveals distinct differences (double-headed arrows in \textbf{a}) in the positions of the cytoplasmic ends of a cluster of four helices (TM16-TM19). Collectively, these four helices are displaced relative to those in \textit{aa}_{3}\text{-oxidase} by an average of ~4.8 Å (measured helix center to helix center) and, as a result provide for tighter packing against the lipid. Helix numbering is that ascribed to SU I/III of \textit{caa}_{3}\text{-oxidase}. Helix numbering for SU III of \textit{aa}_{3}\text{-oxidase} can be obtained by subtracting 12 from the helix number in the figure.
Supplementary Figure 8. The heme \( a_1 \) and binuclear center of \( caa_3 \)-oxidase highlighting ligand geometry and distances. \( \textbf{a} \), Heme \( a_1 \). \( \textbf{b} \), Binuclear centre. The binuclear center includes heme \( a_{13} \) and CuB. The bridging water and hydroxide are shown as coral spheres. Hydrogen and coordinate bonds are shown as red and black dashed lines, respectively. The water - hydroxide hydrogen bond length is 2.3 Å. For comparison, the shortest hydrogen bond recorded in the Cambridge Structural Database is 2.4 Å. The covalent bond length in molecular oxygen is 1.3 Å.
Supplementary Figure 9. A comparison of calculated oxygen channels in Type A and B cytochrome oxidases. Hydrophobic cavities corresponding to oxygen channels were identified for a, \textit{ca}a$_3$- (\textit{T. thermophilus}), b, \textit{ba}$_3$- (\textit{T. thermophilus}), c, \textit{aa}$_3$- (\textit{R. sphaeroides}) and d, \textit{aa}$_3$- (\textit{P. denitrificans}) oxidases using the program HOLLOW, as described under Methods. In (b-d) the oxygen channel for \textit{ca}a$_3$-oxidase (red) was superposed with the corresponding oxygen channels calculated for Type A and B HCOs (cream). Hydrophobic residues lining the cavities in the different oxidases are identified by residue number and are shown in ball and stick representation. Those from \textit{ca}a$_3$-oxidase are colored blue and are included in all four panels for reference. Relevant residues for b, \textit{ba}$_3$- (\textit{T. thermophilus}), c, \textit{aa}$_3$- (\textit{R. sphaeroides}) and d, \textit{aa}$_3$- (\textit{P. denitrificans}) oxidases are identified by residue number and are shown in ball and stick representation in yellow and grey. The active site heme and Cu$_{B}$ of the different oxidase types are included in the corresponding panels.
Supplementary Figure 10. Comparison of the proton pathways in caa₃- and aa₃-type cytochrome oxidases. a, D-pathway showing the superposition of relevant residues and structured waters in caa₃-oxidase of *T. thermophilus* (grey with atomic coloring and red waters, this work) and aa₃-oxidase of *P. denitrificans* (orange, PDB ID: 3HB3) and *R. sphaeroides* (magenta, PDB ID: 1M56). Hydrogen bonds are shown as dashed lines and color coded according to oxidase type. The contrasting configuration of waters at the hydrophobic end of the pathway reflects the different compositions of the gate site of Type A1 and A2 HCOs. b, K-pathway. Details as in a. A hydrogen bond between Lys 328 and Gln 286 in caa₃-oxidase is proposed to replace a water mediated serine – lysine hydrogen bonding interaction found in Type A1 HCOs. Residue numbering is color coded according to oxidase type.
Supplementary Figure 11. Water distribution and representative configurations of the H-bonded water chains in \textit{cca}1-oxidase proton pathways obtained by MDS. \textbf{a, b}, The D-pathway (MD snapshots after 16 and 28 ns, respectively). \textbf{c, d}, The K-pathway (MD snapshots after 74 ns in simulation 1 (c) and after 158 ns in simulation 2 (d)). Variations in the Lys 328 side chain conformation correspond to different water distributions in the K-pathway. This is apparent when the distributions with the side chain in the 'up' (c) and 'down' conformation (d) are compared. Hydrogen bonds are shown as dashed lines. Transparent green surfaces represent the calculated water densities (averaged over MD trajectories and shown at a 30\% occupancy level). Cyan spheres identify the positions of water molecules resolved in the X-ray structure.
**Supplementary Figure 12.** A representative configuration of the H-bonded water networks which form in the water cluster and exit pathway region during MDS (MD snapshot after 62 ns in simulation 1; see legend to Supplementary Figure 11). A plausible proton exit path from the water cluster is indicated by the black arrow and goes via Arg 173(IIC) which is located on the protein surface in contact with bulk solvent. Hydrogen bonds are shown as dashed lines. Transparent green surfaces represent the calculated water densities (averaged over MD trajectories and shown at a 30% occupancy level).
Supplementary Figure 13. Detailed view of the gating region in Type A1 and Type A2 oxidases. a, Type A1 – *P. denitrificans* *aa*$_3$-oxidase glutamate gate (PDB ID: 3HB3), b, the YS motif, the structural equivalent in *caa*$_3$-oxidase of the classical Type A1 gating glutamate. Interaction of Wat 2074 with Tyr 248 connects it to the D-pathway via Wat 2058. The structure of *caa*$_3$-oxidase reveals a hydrogen bond between Asn 205 and Tyr 248 that could influence the protonation state of the tyrosine during the catalytic cycle. Protein is displayed as dark and light blue ribbons for the Type A1 and A2 oxidases, respectively.
Supplementary Figure 14. Structure comparison of the cytochrome c domain of *T. thermophilus* *caat* oxidase (Cyt c<sub>T.t</sub>) with the recombinant soluble cytochrome c domain of *Rh. marinus* *caat* oxidase (Cyt c<sub>R.m</sub>, PDB ID: 1W2L). a, cytochrome c domain of *T. thermophilus* *caat*-oxidase, b, soluble cytochrome c of *Rh. marinus* *caat*-oxidase c, superposition of Cyt c<sub>T.t</sub> and Cyt c<sub>R.m</sub>, and d, an expanded view of the loop and β-sheet region in the vicinity of the heme propionates. The β-sheet region is color coded green (a, c and d) and orange (b, c and d), respectively. The β-sheet region in Cyt c<sub>R.m</sub> is more extended than it is in Cyt c<sub>T.t</sub>, possibly reflecting the lack of an interaction with the cupredoxin domain and/or the fact that it is involved in crystal contacts in the recombinant domain structure. Loop D is defined as residues 295-302 (IIc) (blue, PDB ID: 2YEV) and residues 64-74, (grey, PDB ID: 1W2L) for a and b, respectively. Heme c A- and D-propionates are indicated by boldface A and D, respectively.
Supplementary Figure 15. A comparison of caa3-oxidase from *T. thermophilus* and *cbb3*-oxidase from *Pseudomonas stutzeri* (PDB ID: 3MK7)\(^{17}\) to highlight structural differences in the periplasmic domains. **a**, Overview of integral and extramembrane domains. **b**, Expanded view of periplasmic domains. Subunits are color coded as follows. *cbb3*-Oxidase: CcoN (light grey), CcoP (orange), CcoO (blue) with red hemes. *caa3*-Oxidase: SU I residues 12-504 (dark grey), SU IIc (green), heme c (black) and coppers (grey). Individual models for *cbb3*-oxidase and *caa3*-oxidase are shown in (i) and (iii), respectively. Superposed models are shown in (ii). The SU III domain of SU I/III and SU IV in *caa3*-oxidase are omitted for clarity. Proteins were aligned by secondary structure matching of the conserved SU I, shown in **a**.
**Supplementary Table 1.** Data collection and refinement statistics

|                   | Native                          | Anomalous                     |
|-------------------|--------------------------------|-------------------------------|
| **Data collection** |                                |                               |
| Space Group       | 12                             | 12                            |
| Cell Dimensions   |                                |                               |
| $a$, $b$, $c$ (Å) | 127.3, 76.0, 300.3              | 126.7, 75.9, 299.6            |
| $\beta$ (°)      | 92.2                           | 92.2                          |
| Resolution        | 118.7 – 2.4 (2.49 - 2.36)       | 77.0 – 3.5 (3.69 - 3.50)       |
| Wavelength (Å)    | 0.98                           | 1.74                          |
| $R_{\text{merge}}$ | 16.0 (92.1)                    | 23.0 (79.3)                   |
| $R_{\text{pim}}$  | 5.8 (34.2)                     | 21.9 (75.0)                   |
| $I / \sigma(I)$   | 9.3 (2.2)                      | 7.2 (2.7)                     |
| Completeness (%)  | 100.0 (100.0)                  | 99.6 (99.8)                   |
| Redundancy        | 8.4 (8.2)                      | 3.7 (3.7)                     |
| **Refinement statistics** |                          |                               |
| Resolution (Å)    | 77.18 – 2.36                   |                               |
| Number of reflections | 118,243                       |                               |
| Completeness (%)  | 100.0                          |                               |
| $R_{\text{work}} / R_{\text{free}}$ (%) | 17.1 / 21.8               |                               |
| Number of atoms   | 19,587                         |                               |
| Number of chains  | 6                              |                               |
| Number of residues | 2,325                         |                               |
| Number of ligand/ion molecules | 21                        |                               |
| Number of waters  | 392                            |                               |
| Average B-factors (Å²) |                                |                               |
| Macromolecule     | 58.2                           |                               |
| Lipids            | 73.6                           |                               |
| Solvent           | 52.7                           |                               |
| R.m.s deviations  |                                |                               |
| Bond lengths (Å)  | 0.006                          |                               |
| Bond angles (°)   | 1.084                          |                               |
| Ramachandran plot statistics (%) |                |                               |
| Favored region    | 97.9                           |                               |
| Allowed region    | 2.1                            |                               |
| Outlier region    | 0.0                            |                               |

*a* Values in parentheses are for the highest resolution shell

*b* $R_{\text{merge}} = \Sigma_h |I(h)| - <|I(h)>| / \Sigma_h |I(h)|$, where $<I(h)>$ is the mean intensity of equivalent reflections.
\[ R_{\text{pim}} = \Sigma_n \left( \frac{1}{(N-1)^{0.5}} \sum_i |I(h)_i - \langle I(h) \rangle| / \sum_i I(h)_i \right), \]
where \( \langle I(h) \rangle \) is the mean intensity of equivalent reflections and \( N \) is the redundancy of the data.

\[ R_{\text{work}} = \frac{\Sigma |Fo-Fc|}{\Sigma |Fo|}, \]
where \( Fo \) and \( Fc \) are the observed and calculated structure factor amplitudes, respectively.

\[ R_{\text{free}} = \frac{\Sigma |Fo-Fc|}{\Sigma |Fo|}, \]
calculated using a test-set of randomly sampled reflections (5 % of total observed reflections) omitted in the refinements.

Ramachandran plot statistics were calculated using MolProbity.
**Supplementary Table 2.** Protonation states of proton active residues in *caao* oxidase from *T. thermophilus* at neutral pH.

| Residue | Protonation state$^a$ |
|---------|-----------------------|
| **SU I/III** | |
| Glu 64  | D                     |
| Asp 103 | D                     |
| Tyr 142 | P                     |
| Tyr 143 | P                     |
| Lys 190 | P                     |
| Tyr 194 | P                     |
| Glu 222 | D                     |
| Asp 237 | D                     |
| Tyr 248 | P                     |
| Tyr 254 | P                     |
| Glu 266 | D                     |
| Lys 274 | P                     |
| Tyr 279 | P                     |
| Arg 280 | P                     |
| Lys 328 | D ($\varepsilon = 4$)$^b$; P ($\varepsilon = 10$)$^c$ |
| Asp 373 | P                     |
| Tyr 374 | P                     |
| Asp 378 | D                     |
| Tyr 404 | P                     |
| Arg 447 | P                     |
| Arg 448 | P                     |
| Glu 528 | D                     |
| Arg 529 | P                     |
| **SU IIc** | |
| Asp 81  | D                     |
| Arg 82  | D ($\varepsilon = 4$)$^b$; P ($\varepsilon = 10$)$^c$ |
| Glu 84  | D ($\varepsilon = 4$)$^b$; P ($\varepsilon = 10$)$^c$ |
| Lys 172 | P                     |
| Arg 173 | P                     |
| Asp 174 | D                     |
| Glu 199 | D                     |

$^a$ D = deprotonated ($pK_a < 0$); P = protonated ($pK_a > 14$)

$^b$ Without explicit crystallographic and modelled waters in $pK_a$ calculations

$^c$ With and without explicit crystallographic and modelled waters in $pK_a$ calculations
**Table 3.** Contacts between the cupredoxin and cytochrome c domains of caa₃-oxidase from *T. thermophilus* as identified using PISA

| Hydrogen bonds | Salt Bridges |
|----------------|--------------|
| Arg:138:N      | Arg:138:NH1  |
| Arg:138:NH2    | Arg:207:O    |
| Arg:138:NH1    | Glu:126:N    |
| Arg:211:NH1    | Glu:142:OE2  |
| Arg:211:NH2    | Glu:142:OE2  |
| Ser:204:OG     | Glu:142:OE2  |
| Gln:125:OE1    | Glu:142:OE2  |
| Phe:126:O      | Glu:142:OE2  |
| Asn:139:OD1    | Glu:142:OE2  |
| Ser:140:OG     | Glu:142:OE2  |
| Glu:142:OE2    | Glu:142:OE2  |
| Glu:142:OE2    | Glu:142:OE2  |
| Arg:207:O      | Glu:142:OE2  |
| Phe:126:N      | HEC**:587:O1D|

| Cupredoxin domain | Cytochrome c domain | Distance (Å) |
|-------------------|---------------------|--------------|
| Arg:138:N         | Arg:255:O          | 2.96         |
| Arg:138:NH2       | Gly:263:O          | 3.00         |
| Arg:138:NH1       | Glu:265:OE1        | 2.99         |
| Arg:211:NH1       | Gly:278:O          | 3.01         |
| Arg:211:NH2       | Gly:278:O          | 2.90         |
| Ser:204:OG        | Pro:299:O          | 2.65         |
| Gln:125:OE1       | Lys:302:NZ         | 2.96         |
| Phe:126:O         | Lys:298:NZ         | 2.77         |
| Asn:139:OD1       | Arg:272:NH1        | 3.12         |
| Ser:140:OG        | Glu:265:N          | 2.95         |
| Glu:142:OE2       | Thr:273:N          | 2.81         |
| Glu:142:OE2       | Ser:274:N          | 2.89         |
| Glu:142:OE2       | Thr:273:OG1        | 2.80         |
| Glu:142:OE2       | Ser:274:OG         | 2.94         |
| Arg:207:O         | Lys:298:NZ         | 2.81         |
| Phe:126:N         | HEC**:587:O1D      | 2.95         |

* PISA calculations ([www.ebi.ac.uk/msd-srv/prot_int/pistart.html](http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html)) were carried out with the cupredoxin domain (SU IIc, residues 117-216) and the cytochrome c domain (SU IIc, residues 236-325).

** HEC, heme c
1 Yang, Y. L. et al. Structural elucidation of phosphoglycolipids from strains of the bacterial thermophiles Thermus and Meiothermus. J. Lipid Res. 47, 1823-1832, (2006).
2 Michel, H. Cytochrome c oxidase: catalytic cycle and mechanisms of proton pumping—a discussion. Biochemistry 38, 15129-15140, (1999).
3 Fann, Y. C. et al. Structure of CuB in the binuclear heme-copper center of the cytochrome aa3-type quinol oxidase from Bacillus subtilis: an ENDOR and EXAFS study. Biochemistry 34, 10245-10255, (1995).
4 Qin, L., Hiser, C., Mulichak, A., Garavito, R. M. & Ferguson-Miller, S. Identification of conserved lipid/detergent-binding sites in a high-resolution structure of the membrane protein cytochrome c oxidase. Proc. Natl Acad. Sci. USA 103, 16117-16122, (2006).
5 Kaila, V. R. et al. A combined quantum chemical and crystallographic study on the oxidized binuclear center of cytochrome c oxidase. Biochim Biophys Acta 1807, 769-778, (2011).
6 Ho, B. K. & Gruswitz, F. HOLLOW: generating accurate representations of channel and interior surfaces in molecular structures. BMC. Struct. Biol. 8, 49, (2008).
7 Luna, V. M., Chen, Y., Fee, J. A. & Stout, C. D. Crystallographic studies of Xe and Kr binding within the large internal cavity of cytochrome ba3 from Thermus thermophilus: structural analysis and role of oxygen transport channels in the heme-Cu oxidases. Biochemistry 47, 4657-4665, (2008).
8 Riistama, S. et al. Channelling of dioxygen into the respiratory enzyme. Biochim. Biophys. Acta 1275, 1-4, (1996).
9 Brändén, M. et al. On the role of the K-proton transfer pathway in cytochrome c oxidase. Proc. Natl Acad. Sci. USA 98, 5013-5018, (2001).
10 Koepke, J. et al. High resolution crystal structure of Paracoccus denitrificans cytochrome c oxidase: new insights into the active site and the proton transfer pathways. Biochim. Biophys. Acta 1787, 635-645, (2009).
11 Srinivasan, V. et al. Structure at 1.3 Å resolution of Rhodothermus marinus caa3(3) cytochrome c domain. J. Mol. Biol. 345, 1047-1057, (2005).
12 Janzon, J., Ludwig, B. & Malatesta, F. Electron transfer kinetics of soluble fragments indicate a direct interaction between complex III and the caa3 oxidase in Thermus thermophilus. IUBMB Life 59, 563-569, (2007).
13 Than, M. E. et al. Thermus thermophilus cytochrome-c552: a new highly thermostable cytochrome-c structure obtained by MAD phasing. J. Mol. Biol. 271, 629-644, (1997).
14 Crowley, P. B. & Carrondo, M. A. The architecture of the binding site in redox protein complexes: Implications for fast dissociation. Proteins 55, 603-612, (2004).
Molinas, M. F. *et al.* Electron transfer dynamics of *Rhodothermus marinus* caa3 cytochrome c domains on biomimetic films. *Phys Chem Chem Phys* **13**, 18088-18098, (2011).

Michel, H. Crystallization of membrane proteins. *Trends Biochem. Sci.* **8**, 56-59, (1983).

Buschmann, S. *et al.* The structure of cbb3 cytochrome oxidase provides insights into proton pumping. *Science* **329**, 327-330, (2010).