The Dimer Interface of Bovine Cytochrome *c* Oxidase is Influenced by Local Post-translational Modifications and Lipid Binding

Idlir Liko[1], Matteo Degiacomi[1], Shabaz Mohammed[1,2], Shinya Yoshikawa[3], Carla Schmidt[1,4*] and Carol V. Robinson[1*]

[1] Department of Chemistry, University of Oxford, South Parks Road, Oxford, United Kingdom.

[2] Department of Biochemistry, University of Oxford, South Parks Road, Oxford, United Kingdom.

[3] Picobiology Institute, Graduate School of Life Science, University of Hyogo, 3-2-1 Koto, Kamigohri Ako, Hyogo 678-1297, Japan.

[4] present address: HALOmem, Martin Luther University Halle-Wittenberg, Kurt-Mothes-Str. 3, 06114 Halle, Germany.

* correspondence:

carla.schmidt@biochemtech.uni-halle.de

or

carol.robinson@chem.ox.ac.uk

Abstract (max 250)

Bovine cytochrome *c* oxidase is an integral membrane protein complex comprising 13 protein subunits and associated lipids. Dimerisation of the complex has been proposed, however definitive evidence for the dimer is lacking. We employed advanced mass spectrometry methods to investigate the oligomeric state of cytochrome *c* oxidase and the potential role of lipids and post-translational modifications in its subunit interfaces. Mass spectrometry of the intact protein complex revealed that both the monomer and the dimer are stabilised by large lipid entities. We identified these lipid species from the purified protein complex thus implying that they interact specifically with the enzyme. We further identified phosphorylation and acetylation sites of cytochrome *c* oxidase, located in the peripheral subunits and in the dimer interface respectively. Comparing our phosphorylation and acetylation sites with those found in previous studies of bovine, mouse, rat and human cytochrome *c* oxidase we found that while some acetylation sites within the dimer interface are conserved, suggesting a role for regulation and stabilisation of the dimer, phosphorylation sites were less conserved and more transient. Our results therefore provide insights into the locations and interactions of lipids with acetylated residues within the dimer interface of this enzyme and thereby contribute to a better understanding of its structure in the natural membrane. Moreover dimeric cytochrome *c* oxidase, comprising 20 transmembrane, six extramembrane subunits and associated lipids, represents the largest integral membrane protein complex that has been transferred via electrospray intact into the gas-phase of a mass spectrometer representing a significant technological advance.

Significance (max 120)

Cytochrome *c* oxidase is the last of the enzymes in the mitochondrial electron transport chain and represents a large integral membrane protein complex coupling conversion of molecular oxygen to water with proton translocation across the membrane. We employed mass spectrometry to study cytochrome *c* oxidase extracted from bovine heart and provide the first evidence that both the monomer and the dimer are present in solution and are stabilised by annular and interfacial lipids interactions. We also identified post-translational modifications and located them in the structure of cytochrome *c* oxidase; the location of the acetylated residues in particular suggests a role for this modification in modulating lipid interactions and hence the interplay of the two monomeric subunits.

\body

Introduction

The electron transport chain in both bacteria and mitochondria consists of a series of membrane-embedded protein complexes that transfer electrons from donors to acceptors to establish a proton gradient across the inner membrane. This proton gradient is then utilised by ATP synthase to produce energy. Cytochrome *c* oxidase (C*c*O) is the last enzyme of the electron transport chain and receives electrons from cytochrome c molecules to catalyse reduction of oxygen to water [1]. C*c*O in bovine heart contains 13 different subunits [2] with various associated proteins with lower binding affinities [3]. The three largest subunits (I-III) are encoded by mitochondrial genes [4] and form the core of the enzyme [5]; subunits I and II contain metal binding sites whereas subunit III stabilises these catalytic centres and harbours the O_2 transfer pathway [6]. The remaining ten subunits are thought to have a stabilising function [5]. C*c*O was found to be dimeric in the crystal lattice [7] making it one of the largest transmembrane protein complexes studied by X-ray crystallography.

A significant quantity of lipids [8] have been identified and located in crystal structures of both bacterial and bovine C*c*O [4, 9, 10]. Two cardiolipins (CL), one phosphatidylcholine (PC), three phosphatidylethanolamines (PE), four phosphatidylglycerols (PG) and three triglycerides (TG) have been identified in bovine crystalline C*c*O [11]. Ether phospholipid choline plasmalogens, typically present in mammalian cell membranes, were also identified, but were found to be identical to the corresponding PCs and likely are derivatives thereof [11]. Several lipids were also identified that bridge the two C*c*O monomers suggesting that they may stabilise a putative dimer interface [11, 12] and a critical role for CL has been proposed following its removal and concomitant inactivation of the enzyme [13, 14].

CcO is regulated by expression of specific isoforms that are dependent on the tissue, species, or developmental stage, by allosteric effectors such as ATP/ADP, palmitate or calcium and by reversible phosphorylation [3]. To date, a total of 18 phosphorylation sites has been identified [3]. One of these 18 sites (Thr11-VIa) was proposed in the x-ray structure of C*c*O although not confirmed at the resolution attained [15]. Recently an interplay between allosteric ATP inhibition and reversible phosphorylation was proposed implying inhibitory phosphorylation sites [16].

Dissociation of tightly bound C*c*O subunits is also thought to regulate the enzyme [3]. Alkaline, urea or DDM treatment leads to selective loss of protein subunits [17, 18] and is linked with inactivation of the enzyme [19]. Thermal denaturation revealed a stabilising role for tightly bound phospholipids in both monomeric and dimeric C*c*O [20]. Subunit VIb has been proposed to increase the stability of dimeric C*c*O as suggested by its position in the crystal structure [3]. In addition to dimeric C*c*O, a mitochondrial supercomplex containing monomeric C*c*O has been detected [21]. As allosteric ATP inhibition requires cooperativity of the two monomers however, a biological relevance for the monomeric form continues to be debated.

We address the question of how lipids and post-translational modifications (PTMs) modulate dimerisation of C*c*O by using mass spectrometry (MS). Transferring intact protein assemblies into the gas phase enabled us to investigate the oligomeric state of C*c*O in solution. Locating the phosphorylation sites in the atomic structure of C*c*O reveals multiple sites on the peripheral domains, exposed to the mitochondrial matrix. By contrast the hotspots for acetylation occur primarily along the proposed dimer interfaces, in close contact with the 'lipid plug', suggesting the role of these acetyl groups in modulating the dimeric enzyme through changes in lipid interactions.

Results

Bovine C*c*O contains 13 individual subunits. C*c*O was isolated from bovine heart and we first used liquid chromatography-coupled tandem-MS (LC-MS/MS) of the peptides after protein hydrolysis to confirm the 13 subunits defined previously. As C*c*O contains many integral membrane protein subunits we performed both tryptic and chymotryptic hydrolysis of the proteins to allow generation of sufficient peptides for confident protein and PTM identification. Subunits I-III, VIIb and VIII are mostly embedded in the membrane and have only few tryptic cleavage sites; their sequence coverage after tryptic digestion was therefore low (6-40 %). After chymotryptic hydrolysis these subunits could be identified with high sequence coverage (20-72 %) (**Supplementary Table S1**). Using this approach we confirmed the presence of all 13 subunits and also identified an associated protein VDAC (uniprot-ID P68002) with moderate sequence coverage (61 %). VDAC was previously reported as associated with C*c*O but is not considered a constituent of the core complex [3, 22].

We used denaturing LC-MS to determine exact masses from the mass spectra of ten of the 13 subunits (**Figure S1** and **Supplementary Table S2**). We were not able to measure masses of subunits I, III and VIIb presumably because these membrane proteins are very hydrophobic and their separation by LC is impeded (see also above). Interestingly the proteins eluted over several minutes, in different chromatographic peaks, indicative of multiple forms and with higher masses than their theoretical values suggesting that they contain PTMs.

PTMs are located on the peripheral subunits and in the dimeric interface of C*c*O. We carried out standard proteomics methods (see **Methods**) and included potential phosphorylation of serine, threonine and tyrosine residues as well as acetylation of lysine residues in database searches. In total, we identified 22 modified residues (8 phosphorylated and 14 acetylated) in seven of the 13 C*c*O subunits (**Figure 1** and **Supplementary Table S3**). Most of the modified residues are located in soluble domains of subunits IV, Va and Vb with at least four PTMs in each of these subunits. Interestingly none of these subunits shows a preference for phosphorylation or acetylation, all three subunits are equally modified by both PTMs. However, all phosphorylated residues are serines or threonines with no phosphorylated tyrosine detected. Representative spectra are shown for both phosphorylation and acetylation (**Supplementary Figure S2**). Comparing our results with previous studies in which 11 acetylation and 3 phosphorylated residues were identified in human, mouse or rat (**Table S4**) we find that the majority of acetylated residues are conserved across species (**Figure 2** and **Supplementary Figure S3**). Locating the PTMs within the crystal structure of the C*c*O dimer (PDB ID 2OCC) shows that phosphorylation occurs primarily on the matrix-side of the soluble subunits (**Figure 1**). By contrast acetylated lysine residues are more widely distributed, often in proximity to phosphorylation sites.

C*c*O monomers and dimers in solution. We next acquired mass spectra of C*c*O under high-energy conditions to strip the detergent micelle and release the intact protein complex in the gas phase [23, 24]. The low *m/z* region of these spectra shows charge state series assigned to single protein subunits (**Figure 3**). We identified these based upon their masses as subunits IV, Va, VIc and VIIb and located them in the crystal structure. They correspond to peripheral membrane and soluble subunits of the complex

We also observed three peak series with a molecular weight of 76 kDa, 98 kDa and 157 kDa corresponding to subcomplexes of C*c*O (**Supplementary Figure S4**). The mass of the

smallest complex, together with the crystal structure, suggests that subunits I and IV form the core of these assemblies. Numerous assignments are possible for the 98 kDa complex, however, it is likely that it contains the core subunits I and IV as well as multiple peripheral subunits. The high mass complex corresponds to the intact C*c*O monomer that has lost several peripheral subunits. Interestingly, the mass of this subcomplex is in accord with the loss of the subunits IV, Va, VIc and VIIb detected in the low *m/z* region of this mass spectrum (see above, **Figure 3**) leaving a stable core containing subunits I, II, III, Vb, VIa, VIb, VIIa, VIIc and VIII. Interestingly C*c*O is not stable under the relatively harsh conditions required to release the protein complex from the DM micelle. We therefore screened various detergents to identify one that requires lower activation energy to release the C*c*O complex intact from the detergent micelle. C_8E_4 is considered a 'harsher' detergent, i.e. its ability to substitute for natural lipids is low compared with DM or DDM, but it requires lower activation energy to release proteins from micelles in the gas phase [25].

Indeed, after exchanging DM with C_8E_4 and recording a mass spectrum we observed two charge state series at high *m/z* (**Figure 4**). Considering the experimental masses of the protein subunits the theoretical masses of the intact C*c*O monomer and dimer correspond to 205.4 and 410.8 kDa, respectively. The two peak series have masses of 215.7 and 435.3 kDa. The broad peaks observed in the mass spectrum are attributed to multiple incomplete PTMs and heterogeneous lipid binding. The difference between calculated and theoretical masses corresponds to an additional 10 and 25 kDa for the monomer and dimer, respectively. These masses suggest that the complexes are associated with a large lipid entity of approximately 10 kDa per monomer, with a further 5 kDa of mass in dimeric C*c*O, ascribed to lipids located in the inner cavity within the complex.

Lipids that associate with CcO. We analysed the lipid content of C*c*O using an established protocol [26] digesting the proteins with trypsin and separating the resulting peptides from the associated lipids by LC. The lipids were eluted directly into a high-resolution MS and analysed by LC-MS/MS. We searched the precursor masses against a virtual lipid database containing TG, PC, phosphatidic acids (PA), phosphatidylserines (PS), PE, PG, phosphatidylinositols (PI) phosphatidylinositol phosphates (PIP) and CL with various acyl chain lengths (www.lipidmaps.com). This precursor ion search does not include the identification of the lipids by tandem-MS. We therefore inspected the fragment spectra of abundant lipids by MS/MS data search (www.lipidmaps.org) to verify the lipids and identify the isomers. Following this workflow we identified four lipid species with different fatty

acids including PG, PI, PS and PE (**Supplementary Table S5**). Manual inspection of the MS/MS spectra also revealed one species that was not included in the lipid database, phosphatidylglycerol phosphate (PGP), characterised by neutral loss of phosphate and co-elution with the corresponding PG (**Supplementary Table S5**). The presence of PGP in our analysis implies that the PG isomers that we identified originate from their related PGP by neutral loss of the phosphate group during activation in the MS. Our lipid analysis further revealed preferences of fatty acids in these lipids: 16:0, 18:0, 18:1, 18:2 and 20:4 being clearly favoured consistent with the fatty acid composition of mammalian heart mitochondria with similar proportions [27].

The most abundant lipid precursors are indicated in a combined mass spectrum (**Supplementary Figure S5A**) which also reveals singly charged species above *m/z* 1300, most likely corresponding to CL isomers (**Supplementary Figure S6**). A precursor search revealed their numbers of carbon atoms and double-bonds in the fatty acids allowing us to conclude possible combinations from the favoured side chains of other lipids (**Supplementary Figure S5A**). The most abundant CL isomer is assigned to 66:5 and is of significantly shorter chain length than the average mitochondrial CL [28]. In total, we identified 35 lipid isomers from six lipid classes. Of these, PE is the most abundant species with 14 isomers and the highest precursor intensities (**Supplementary Figure S5A and Table S5**). Compared to previous studies we also identified lipid species that were not anticipated (e.g. PI and PGP) (**Supplementary Figure S5B**). PI and PS together represent minor constituents of the mitochondrial membrane [27]. Interestingly PGP has not been identified in the inner mitochondrial membrane previously.

CcO dimers are formed with multiple interfacial lipids. The high level of lipids associated with this complex implies not only that they are sufficient to form a protective shield but that they may also reside within subunit interfaces. Indeed the difference between the anticipated mass of the dimer and the measured mass equates to 25 kDa consistent with over 30 phospholipids. The crystal structure of CcO (PDB ID 2OCC) reveals a large cavity between the monomers with few protein-protein contacts. We calculated the volume of this cavity to define the number of lipids that could be aCcOmmodated (**Supplementary Information**) by aligning the crystal structure with the membrane and determining the hydrophobic region of the membrane-spanning subunits as 28.6 Å. The volume was then calculated by inserting the crystal structure into a mesh grid then removing all the grid points that clash with protein atoms in the crystal structure. The volume of the cavity was thus defined by the volume of the

remaining grid points within the hydrophobic region (**Figure 5**). The volume for fatty acid side chains identified in our lipid analysis that could be accommodated in the central cavity equates to 7-8 lipids.

We therefore propose that the 'lipid shield' of monomeric and dimeric C*c*O accounts for 10 and 20 kDa, respectively, while the 'lipid plug' within the dimeric interface corresponds to 5 kDa, consistent with the additional 25 kDa. (**Figure 4**). Highlighting the 'lipid plug' in the crystal structure of C*c*O and locating the PTMs identified here reveals that these acetylated residues line the inner cavity and point towards several lipid molecules. (**Figure 5 lower panel**). Acetylation will change the charge properties of the amine groups and modulate electrostatic interactions with phosphoplipids providing a plausible fine-tuning mechanism to modulate protein lipid interactions

Discussion

We have characterised C*c*O purified from bovine heart by determining accurate subunit masses and identifying PTMs, lipids and a protein associated with the intact complex. Transferring intact C*c*O into the gas phase of a mass spectrometer from C_8E_4 allowed us to release intact monomeric and dimeric C*c*O containing 13 and 26 subunits, respectively. The masses observed for these complexes revealed a large lipid entity associated with both oligomers, the C*c*O monomer binding an additional 10 kDa corresponding to ~13 phospholipids when assuming an average mass of 775 Da (**Supplementary Table S5**). The lipids associated with dimeric C*c*O constitute 25 kDa. Assuming that both individual monomers are stabilised by lipids, about six phospholipids likely stabilise the dimeric interface, specifically CL is thought to stabilise the dimeric C*c*O interface [11, 14, 29, 30]. Recent molecular dynamics simulations identified seven binding sites for CL in C*c*O [29] most of which are contradictory to sites identified from crystal structures [11, 15].

We identified five lipid classes, with several isomers, and a number of CL. Considering the number of isomers identified (**Supplementary Table S5**) and the peak intensities recorded (**Supplementary Figure S5A**) PE is the most prevalent lipid. Given that PE and PC are most abundant in bovine heart tissue [27], their prevalence in the natural membrane suggests that they are often in contact with complexes and may constitute the annular lipids of the C*c*O complex. PC was not identified in this study but was identified previously [11]. Our lipid analysis also revealed three new lipid classes associated with bovine C*c*O: PI, PS and PGP. Interestingly these lipids are only minor constituents of the mitochondrial inner membrane

[27] and are therefore likely sequestered from exchange with bulk lipids in the membrane (PE and PC) by virtue of their specific binding in the dimer interface. A similar phenomenon has been reported previously for the membrane-spanning ring in V- and F-type ATPases. In both cases the 'lipid plug' is represented by low-abundant lipids in their respective membranes, sequestered from exchange with bulk lipids that make up the surrounding lipid environment [26, 31]. We therefore propose that PI, PS and PGP bind specifically, presumably to the inner cavity of the C*c*O dimer interface to impart stability to dimeric C*c*O by binding in the inner cavity. We also observed CL, a lipid previously suggested to be in close contact to the proteins that form inner cavities [11, 13, 14]. In high-resolution structures reported previously CL is located at the contact sites between the two monomeric complexes which make relatively few protein interactions [6] (see also **Figure 5**). The chain length of the abundant CL isomer identified here (16:2) agrees well with the dimensions of the hydrophobic membrane-spanning domain, calculated for the 'lipid plug' of C*c*O.

Molecular dynamics simulations of monomeric C*c*O revealed two high-affinity and five low-affinity CL binding sites only on the surface of the monomer implying that binding of CL in the dimeric interface only occurs during dimer formation [29]. Interestingly, the two high-affinity CL binding sites were located in close proximity to the proton channels and were associated with the regulation of the enzyme [32, 33] while the low-affinity binding sites were linked to stabilisation of subunits VIa and VIb, important for dimerisation of C*c*O [30, 34]. A loss of C*c*O function in the absence of CL was shown recently through destabilisation of quaternary structure following CL removal [13, 14]. Specific association of CL, and location in the dimeric interface and/or at the regulatory sites of C*c*O, therefore represents an important issue to resolve to understand fully C*c*O function. We propose based on these results that CL likely mediates the equilibrium between monomeric and dimeric forms of CcO.

Comparing the previous X-ray structure of bovine heart C*c*O, wherein only the symmetric 72:8 CL was identified [11], with the lipids found here which include 66:5, 68:8, 72:10, 76:13 and 78:8 a greater dispersity of chain length and a higher degree of unsaturation is revealed. This observation could be consistent with diverse locations for CL in different subunit interfaces requiring different chain lengths. We propose that with their two phosphatidic acid moieties CLs form a bridge between subunits, their tailored side chains acting as a 'hydrophobic glue', to stablise different subunit interactions as well as the C*c*O dimer interface.

To gain more insights into the lipid mediated stabilisation of subunit interfaces as well as possible regulatory mechanisms in C*c*O we also identified acetylation and phosphorylation sites. Interestingly, subunits IV, Va and Vb were modified to the greatest extent. Locating these PTMs in the crystal structure of C*c*O, and taking into account the flexibility of their locations on unstructured loops, reveals that the phosphorylation sites are present predominantly in soluble domains, accessible to the matrix side of the inner membrane, accessible for kinases and phosphatases. Strikingly, phosphorylation was detected primarily on serine, and to a lower extent on threonine residues. In agreement with other studies [35] this indicates that these sites have a regulatory role rather than a stabilising one which is often ascribed to phosphorylation on tyrosine residues [35].

Comparing our PTMs with those reported in databases (www.uniprot.org) reveals that only 5 of our modified sites coincide with those identified previously (**Supplementary Table S3**). This relatively low overlap might have several origins. Critically, phosphorylation and acetylation are highly dynamic and reversible, as such they are often lost during purification or extraction either by dephosphorylation/deacetylation due to the presence of the respective enzymes [36]. In addition, PTMs are often of low abundance in natural sources (<10%) due to their functional roles in signaling cascades [37]. Low levels of PTMs have also been shown to regulate subunit interfaces [31, 38] with subtle differences in phosphorylation status having major impact on complex stability [31]. Even though developments of higher resolution and sensitivity have increased the numbers of PTMs identified by MS [39] often only subsets are captured and therefore do not necessarily correspond in different studies. Further modified sites have been proposed for C*c*O [4] in addition to the new sites described here, implying that the number of PTMs in this complex is likely to increase with future MS developments and greater understanding of purification/extraction procedures. A comparison over different mammalian species presented here revealed that the acetylation sites identified, and those deposited in databases, are highly conserved over bovine, human, mouse and rat in C*c*O subunits IV, Va and Vb. While only five PTMs were consistently identified in bovine C*c*O, nine acetylation sites and three phosphorylation sites are conserved in subunits IV, Va and Vb over these four species. Aligning their amino acid sequences shows a high sequence homology (>82%) suggesting that the novel PTMs identified here, or previously, will likely be present in other mammalian species.

While phosphorylation has been characterized extensively over the last decade, acetylation was initially only considered to play a role in histone modification. However recent large-

scale proteomic studies have revealed that acetylation represents a major PTM in eukaryotes [40] and prokaryotes [41]. Interestingly, 63% of mitochondrial proteins contain acetylation sites, most of which have been assigned an inhibitory role [42]. Interestingly the acetylation sites identified here, and those reported previously, cluster with phosphorylation sites in close proximity to the dimeric interface (**Figure 5 lower panel**).

Given that there are very few protein-protein interactions in the C*c*O dimer interface, representing a rather weak interface, it is likely that lipids could play a role in stabilisation (**Figure 5**). Assuming that the low abundance lipids PS and PI are sequestered in the inner cavity of C*c*O, lysine residues in this region might undergo electrostatic interactions with the headgroups of the lipids. Acetylation of these amino groups would thus modulate the stability of dimeric C*c*O in line with previous reports of lysine acetylation changing lipid binding properties [43]. PTM clusters at protein interfaces and in unstructured regions have been described previously, moreover their importance for stabilising and controlling protein interactions has been proposed [44]. Interactions between PTMs of proteins and associated lipids is however a relatively new concept that is sure to have major implication for the organisation of membrane protein complexes.

In summary this protein-lipid assembly is the largest integral membrane complex analysed intact by electrospray MS to date. Although previous studies have demonstrated the survival of rotary ATPases [26, 31] these complexes have large soluble head regions facilitating their survival outside of the membrane. Moreover early MS studies delivered only approximate subunit masses and lipid composition due to the low resolution available at that time [45]. The fact that we could preserve the intact C*c*O dimer with an assembled lipid plug and transfer this from solution to gas phase, with both monomeric and dimeric species present, provides the first compelling evidence for the existence of this lipid mediated dimer in solution. The functional role of this dimer is not yet clear however. Similarly other complexes of the electron transport chain, such as the bc_1-complex, have been found to be dimeric [46]. Although C*c*O and bc_1 are functionally independent, an interplay between the monomeric complexes is often discussed (for example see [47]). The role of lipids and PTMs in mediating interactions between C*c*O and bc_1 is often debated but not yet defined. By distinguishing lipids that bind within an annular belt, from those in the central lipid plug, and by locating PTMs within the structure, we propose that the interplay of the monomeric complexes is fine-tuned by multiple acetylation and tailored by selective lipid-binding.

Methods

Further details of all methods are given in **Supplementary Methods.**

C*c*O purification. C*c*O was purified as crystalline preparation in DM from bovine heart as described in the supplementary methods. Each sample was prepared from one heart.

Protein and PTM identification. Proteins were separated by SDS-PAGE and digested with Trypsin. Peptides were separated by nano-LC (EASY nLC 1000 system, Thermo Scientific) and directly eluted into a mass spectrometer (Q Exactive Orbitrap, Thermo Scientific). Proteins and PTMs were identified by database searching.

Denaturing LC-MS. Proteins were separated by nano-LC (DionexUltiMate 3000 RSLC, Thermo Scientific) and directly eluted into a QSTAR XL mass spectrometer (ABSciex).

MS of intact CcO. Mass spectra were acquired in ammonium acetate on a Q-ToF II or Synapt1 mass spectrometer (Waters) modified for high masses [48] using in-house prepared gold-coated glass capillaries [49].

Detergent exchange. The detergent was exchanged as described previously [50].

Lipid identification. CcO proteins were digested with Trypsin and the peptide/lipid mixture was separated by nano-LC (DionexUltiMate 3000 RSLC, Thermo Scientific). Peptides and lipids directly eluted into an LTQ-Orbitrap XL mass spectrometer (Thermo Scientific).

Acknowledgements

We thank Emily Barlow for preliminary MS experiments and Kyoko Shinzawa-Ito for purification of bovine C*c*O. This work was funded by an ERC grant (IMPRESS, ERC 268851) to CVR. CVR further acknowledges funding from the MRC (98101) and the Wellcome Trust (WT008150 and WT099141). CVR is a Royal Society Professor. MTD acknowledges funding from the Swiss National Science Foundation (grant nb. P2ELP3_155339). SY is funded by a Grant-in-Aid for the Global Center of Excellence Program and for Scientific Research (A) 2247012 and (B) 26291033, by CREST and by Targeted Proteins Research Program, each provided by the Japanese Ministry of Education, Culture, Sports, Science and Technology. SY is a Senior Visiting Scientist in the RIKEN Harima Institute.

Figure Legends

Figure 1: PTMs are located in the peripheral subunits and in the dimer interface of C*c*O. Phosphorylated (yellow) and acetylated (orange) residues are highlighted on the crystal structure of C*c*O (PDB ID 2OCC). Residue numbers and the type of PTM (*p*, phosphorylation or *ac*, acetylation) are given. Phosphosites are located mainly in peripheral subunits on the matrix side. Acetylated residues are located at protein interfaces and in the C*c*O dimer interface. . **(A)** side view of C*c*O, **(B)** bottom view of C*c*O.

Figure 2: PTMs cluster in subunits IV, Va and Vb. PTMs found in bovine CcO (above the bar) in this study (filled) with those reported previously (empty symbols) are indicated by squares. PTMs identified in both this and previous studies are transparent. Phosphorylation (red) and acetylation (blue) sites are shown. A comparison between the bovine sites (squares, on top) and those in other mammalia (rat/human/mouse, circles, below the bar) shows that mostly acetylation sites are conserved across the species (filled, transparent circles). The mitochondrial transit peptide is shown in grey.

Figure 3: MS of CcO reveals peripheral subunits. The low m/z region (<2600 m/z) of the mass spectrum shows four peak series corresponding to peripheral subunits IV (blue), Va (red), VIc (orange), and VIIb (cyan). Charge states for the most intense peaks of each series are given. The peripheral subunits are highlighted on the crystal structure of CcO (PDB ID 2OCC). Intact CcO was not observed in DM detergent presumably due to the high activation energy required to release the proteins from the detergent micelle.

Figure 4: MS of intact C*c*O reveals stabilisation of monomeric/dimeric C*c*O by associated lipids. After exchanging DM detergent for C_8E_4 detergent (top) the C*c*O complex remains intact and the MS reveals peak series for monomeric (blue circles) and dimeric (grey circles) C*c*O. Charge states for the most intense peak of each series are given. Observed masses of the monomer (215 kDa) and dimer (435 kDa) are higher than the theoretical masses calculated from subunit masses consistent with the association of lipids in monomeric and dimeric C*c*O. The broad peak width of the charge states is assigned to multiple lipid bound states and heterogeneous PTMs.

Figure 5: The 'lipid plug' of C*c*O. Side view (top panel) and bottom view (lower panel) of the C*c*O dimer (PDB ID 2OCC). The membrane spanning hydrophobic region (28.6 Å) is

indicated (top panel). The volume of the inner cavity (11'346 $Å^3$, red) allows incorporation of 7 to 8 phospholipids or 3 to 4 CL, respectively. PTMs (phosphorylation, yellow and acetylation, orange) in the dimeric interface of C*c*O were located in the crystal structure. *p*S30-Vb is not included in the crystal structure and *p*S33-Vb is shown instead (yellow). The bottom view of the complex reveals at least six acetylated lysine residues at the dimer interface suggesting interactions with lipids. Protein-protein interactions in the dimer interface are shown as backbone and cartoon representation for the two monomers (magnifications). Amino acids in direct contact with other amino acids are shown as licorice revealing a weak dimer interface stabilised by phospholipids.

References

1. Babcock, G.T. and M. Wikstrom, *Oxygen activation and the conservation of energy in cell respiration.* Nature, 1992. **356**(6367): p. 301-9.
2. Kadenbach, B., et al., *The complexity of respiratory complexes.* Trends Biochem. Sci., 1983. **8**: p. 398-400.
3. Kadenbach, B. and M. Huttemann, *The subunit composition and function of mammalian cytochrome c oxidase.* Mitochondrion, 2015. **24**: p. 64-76.
4. Tsukihara, T., et al., *The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 A.* Science, 1996. **272**(5265): p. 1136-44.
5. Yoshikawa, S., K. Muramoto, and K. Shinzawa-Itoh, *Proton-pumping mechanism of cytochrome C oxidase.* Annu Rev Biophys, 2011. **40**: p. 205-23.
6. Yoshikawa, S., et al., *Structural studies on bovine heart cytochrome c oxidase.* Biochim Biophys Acta, 2012. **1817**(4): p. 579-89.
7. Tsukihara, T., et al., *Structures of metal sites of oxidized bovine heart cytochrome c oxidase at 2.8 A.* Science, 1995. **269**(5227): p. 1069-74.
8. Robinson, N.C., *Specificity and binding affinity of phospholipids to the high-affinity cardiolipin sites of beef heart cytochrome c oxidase.* Biochemistry, 1982. **21**(1): p. 184-8.
9. Harrenga, A. and H. Michel, *The cytochrome c oxidase from Paracoccus denitrificans does not change the metal center ligation upon reduction.* J Biol Chem, 1999. **274**(47): p. 33296-9.
10. Svensson-Ek, M., et al., *The X-ray crystal structures of wild-type and EQ(I-286) mutant cytochrome c oxidases from Rhodobacter sphaeroides.* J Mol Biol, 2002. **321**(2): p. 329-39.
11. Shinzawa-Itoh, K., et al., *Structures and physiological roles of 13 integral lipids of bovine heart cytochrome c oxidase.* EMBO J, 2007. **26**(6): p. 1713-25.
12. Yoshikawa, S. and A. Shimada, *Reaction mechanism of cytochrome c oxidase.* Chem Rev, 2015. **115**(4): p. 1936-89.
13. Musatov, A. and N.C. Robinson, *Bound cardiolipin is essential for cytochrome c oxidase proton translocation.* Biochimie, 2014. **105**: p. 159-64.
14. Sedlak, E. and N.C. Robinson, *Destabilization of the Quaternary Structure of Bovine Heart Cytochrome c Oxidase upon Removal of Tightly Bound Cardiolipin.* Biochemistry, 2015. **54**(36): p. 5569-5577.
15. Tsukihara, T., et al., *The low-spin heme of cytochrome c oxidase as the driving element of the proton-pumping process.* Proc Natl Acad Sci USA, 2003. **100**(26): p. 15304-15309.

16. Helling, S., et al., *Multiple phosphorylations of cytochrome c oxidase and their functions.* Proteomics, 2012. **12**(7): p. 950-9.
17. Saraste, M., T. Penttila, and M. Wikstrom, *Quaternary structure of bovine cytochrome oxidase.* Eur J Biochem, 1981. **115**(2): p. 261-8.
18. Thompson, D.A., L. Gregory, and S. Ferguson-Miller, *Cytochrome c oxidase depleted of subunit III: proton-pumping, respiratory control, and pH dependence of the midpoint potential of cytochrome a.* J Inorg Biochem, 1985. **23**(3-4): p. 357-64.
19. Sedlak, E. and N.C. Robinson, *Sequential dissociation of subunits from bovine heart cytochrome C oxidase by urea.* Biochemistry, 2009. **48**(34): p. 8143-50.
20. Sedlak, E., et al., *The kinetic stability of cytochrome C oxidase: effect of bound phospholipid and dimerization.* Biophys J, 2014. **107**(12): p. 2941-9.
21. Wittig, I. and H. Schagger, *Advantages and limitations of clear-native PAGE.* Proteomics, 2005. **5**(17): p. 4338-4346.
22. Roman, I., et al., *In vitro interactions between the two mitochondrial membrane proteins VDAC and cytochrome c oxidase.* Biochemistry, 2005. **44**(39): p. 13192-13201.
23. Barrera, N.P., et al., *Micelles protect membrane complexes from solution to vacuum.* Science, 2008. **321**(5886): p. 243-6.
24. Barrera, N.P., et al., *Mass spectrometry of membrane transporters reveals subunit stoichiometry and interactions.* Nat Methods, 2009. **6**(8): p. 585-7.
25. Reading, E., et al., *The role of the detergent micelle in preserving the structure of membrane proteins in the gas phase.* Angew Chem Int Ed Engl, 2015. **54**(15): p. 4577-81.
26. Zhou, M., et al., *Mass spectrometry of intact V-type ATPases reveals bound lipids and the effects of nucleotide binding.* Science, 2011. **334**(6054): p. 380-5.
27. Daum, G., *Lipids of mitochondria.* Biochim Biophys Acta, 1985. **822**(1): p. 1-42.
28. Schlame, M., S. Brody, and K.Y. Hostetler, *Mitochondrial cardiolipin in diverse eukaryotes. Comparison of biosynthetic reactions and molecular acyl species.* Eur J Biochem, 1993. **212**(3): p. 727-35.
29. Arnarez, C., S.J. Marrink, and X. Periole, *Identification of cardiolipin binding sites on cytochrome c oxidase at the entrance of proton channels.* Sci Rep, 2013. **3**: p. 1263.
30. Musatov, A. and N.C. Robinson, *Cholate-induced dimerization of detergent- or phospholipid-solubilized bovine cytochrome C oxidase.* Biochemistry, 2002. **41**(13): p. 4371-6.
31. Schmidt, C., et al., *Comparative cross-linking and mass spectrometry of an intact F-type ATPase suggest a role for phosphorylation.* Nat Commun, 2013. **4**: p. 1985.
32. Robinson, N.C., *Functional binding of cardiolipin to cytochrome c oxidase.* J Bioenerg Biomembr, 1993. **25**(2): p. 153-63.
33. Robinson, N.C., F. Strey, and L. Talbert, *Investigation of the essential boundary layer phospholipids of cytochrome c oxidase using Triton X-100 delipidation.* Biochemistry, 1980. **19**(16): p. 3656-61.
34. Sedlak, E. and N.C. Robinson, *Phospholipase A(2) digestion of cardiolipin bound to bovine cytochrome c oxidase alters both activity and quaternary structure.* Biochemistry, 1999. **38**(45): p. 14966-72.
35. Ly, T., et al., *A proteomic chronology of gene expression through the cell cycle in human myeloid leukemia cells.* Elife, 2014. **3**: p. e01630.
36. Zhao, Y. and O.N. Jensen, *Modification-specific proteomics: strategies for characterization of post-translational modifications using enrichment techniques.* Proteomics, 2009. **9**(20): p. 4632-41.
37. Baeza, J., et al., *Stoichiometry of Site-specific Lysine Acetylation in an Entire Proteome.* J Biol Chem, 2014. **289**(31): p. 21326-21338.
38. Morgner, N., et al., *Hsp70 Forms Antiparallel Dimers Stabilized by Post-translational Modifications to Position Clients for Transfer to Hsp90.* Cell Reports, 2015. **11**(5): p. 759-769.

39. Olsen, J.V. and M. Mann, *Status of Large-scale Analysis of Post-translational Modifications by Mass Spectrometry.* Mol Cell Prot, 2013. **12**(12): p. 3444-3452.
40. Mischerikow, N. and A.J. Heck, *Targeted large-scale analysis of protein acetylation.* Proteomics, 2011. **11**(4): p. 571-89.
41. Soufi, B., et al., *Proteomics reveals evidence of cross-talk between protein modifications in bacteria: focus on acetylation and phosphorylation.* Curr Opin Microbiol, 2012. **15**(3): p. 357-63.
42. Baeza, J., M.J. Smallegan, and J.M. Denu, *Mechanisms and Dynamics of Protein Acetylation in Mitochondria.* Trends Biochem Sci, 2016. **41**(3): p. 231-44.
43. Maltsev, A.S., J. Ying, and A. Bax, *Impact of N-terminal acetylation of alpha-synuclein on its random coil and lipid binding properties.* Biochemistry, 2012. **51**(25): p. 5004-13.
44. Beilsten-Edmands, V., et al., *eIF2 interactions with initiator tRNA and eIF2B are regulated by post-translational modifications and conformational dynamics.* Cell Discovery, 2015. **15020**.
45. Morgner, N., et al., *A novel approach to analyze membrane proteins by laser mass spectrometry: From protein subunits to the integral complex.* J Am Soc for Mass Spectrom, 2007. **18**(8): p. 1429-1438.
46. Castellani, M., et al., *Direct Demonstration of Half-of-the-sites Reactivity in the Dimeric Cytochrome bc(1) 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 2010. **285**(1): p. 502-510.
47. 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, 2008. **1777**(7-8): p. 1001-19.
48. Sobott, F., et al., *A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies.* Anal Chem, 2002. **74**(6): p. 1402-7.
49. Hernandez, H. and C.V. Robinson, *Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry.* Nat Protoc, 2007. **2**(3): p. 715-26.
50. Laganowsky, A., et al., *Mass spectrometry of intact membrane protein complexes.* Nat Protoc, 2013. **8**(4): p. 639-51.
51. Yoshikawa, S., et al., *An infrared study of CO binding to heart cytochrome c oxidase and hemoglobin A. Implications re O2 reactions.* J Biol Chem, 1977. **252**(15): p. 5498-508.
52. Mochizuki, M., et al., *Quantitative reevaluation of the redox active sites of crystalline bovine heart cytochrome c oxidase.* J Biol Chem, 1999. **274**(47): p. 33403-11.
53. Lomize, M.A., et al., *OPM: Orientations of proteins in membranes database.* Bioinformatics, 2006. **22**(5): p. 623-625.

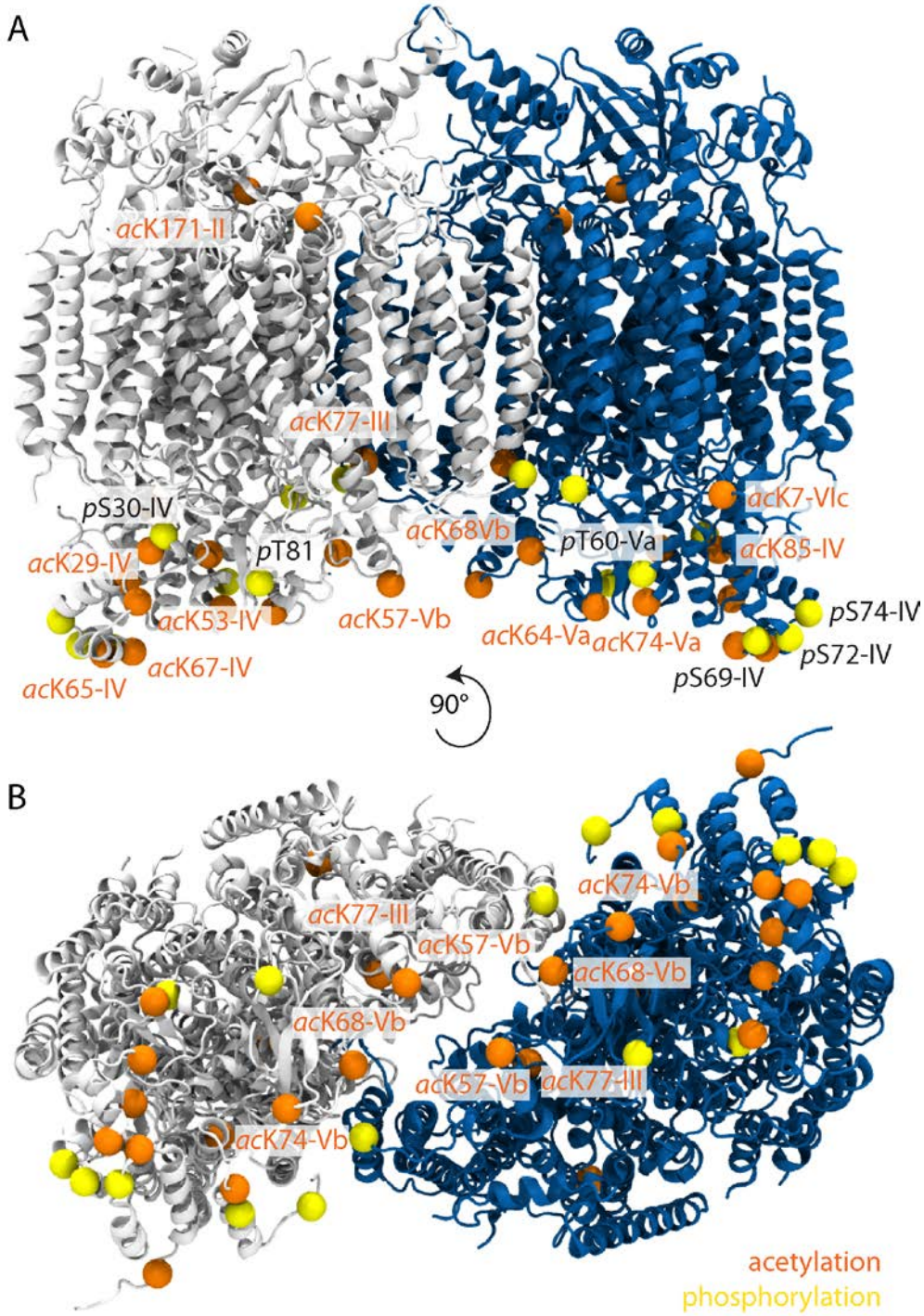
A
acK171-II
acK77-III
pS30-IV
acK29-IV
pT81
acK53-IV
acK57-Vb
acK67-IV
acK65-IV
acK68Vb
pT60-Va
acK64-Va
acK74-Va
acK7-VIc
acK85-IV
pS74-IV
pS72-IV
pS69-IV
90°
B
acK74-Vb
acK77-III
acK57-Vb
acK68-Vb
acK68-Vb
acK57-Vb
acK77-III
acK74-Vb
acetylation
phosphorylation

Figure 1

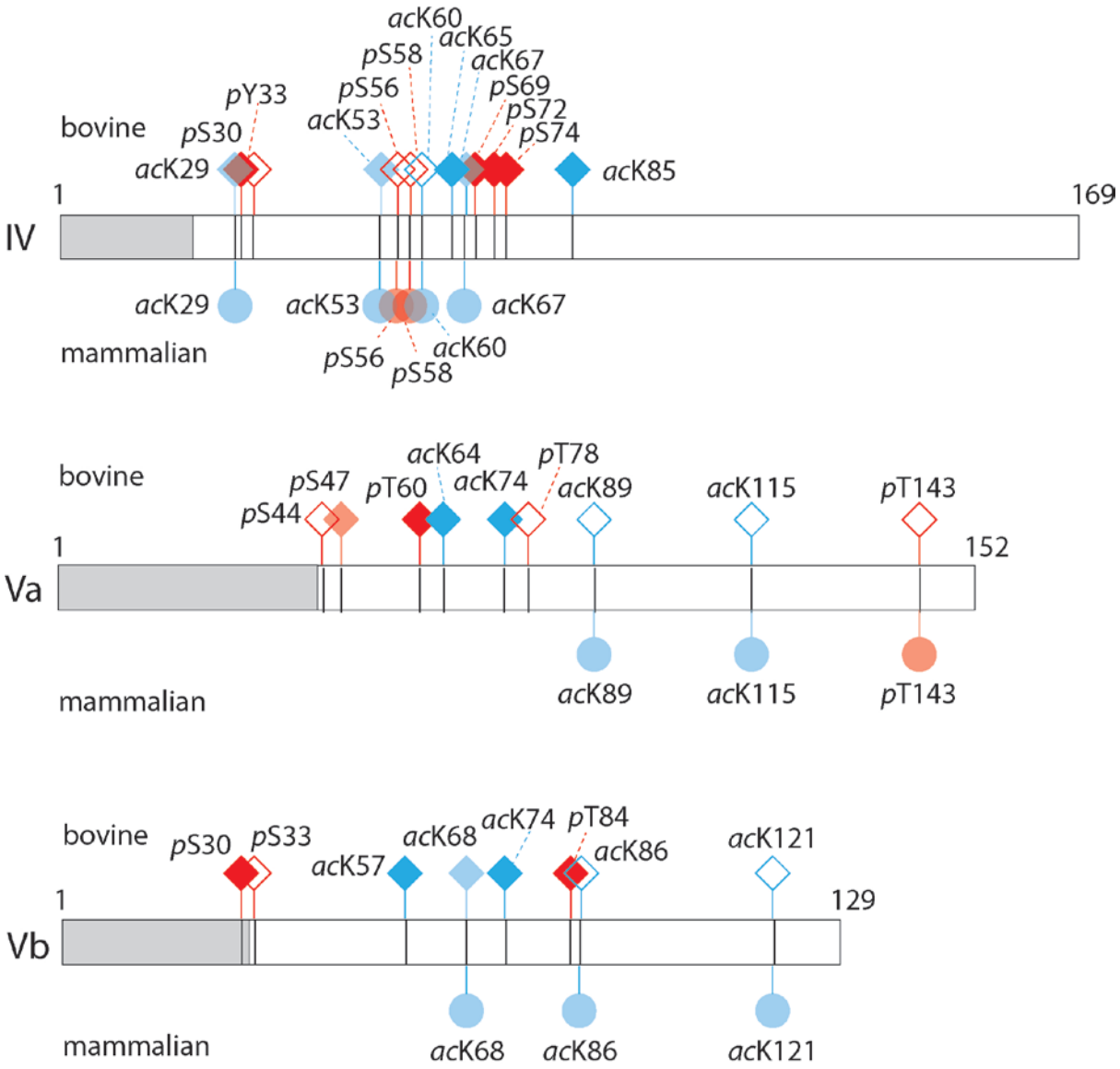
bovine
pY33
pS30
acK29
acK60
acK65
pS58
acK67
pS56
pS69
acK53
pS72
pS74
acK85
1
169
IV
acK29
acK53
acK67
mammalian
pS56
acK60
pS58
bovine
acK64
pT78
pS47
pT60
acK74
pS44
acK89
acK115
pT143
1
152
Va
mammalian
acK89
acK115
pT143
bovine
pS30
pS33
acK68
acK74
pT84
acK57
acK86
acK121
1
129
Vb
mammalian
acK68
acK86
acK121

figure 2

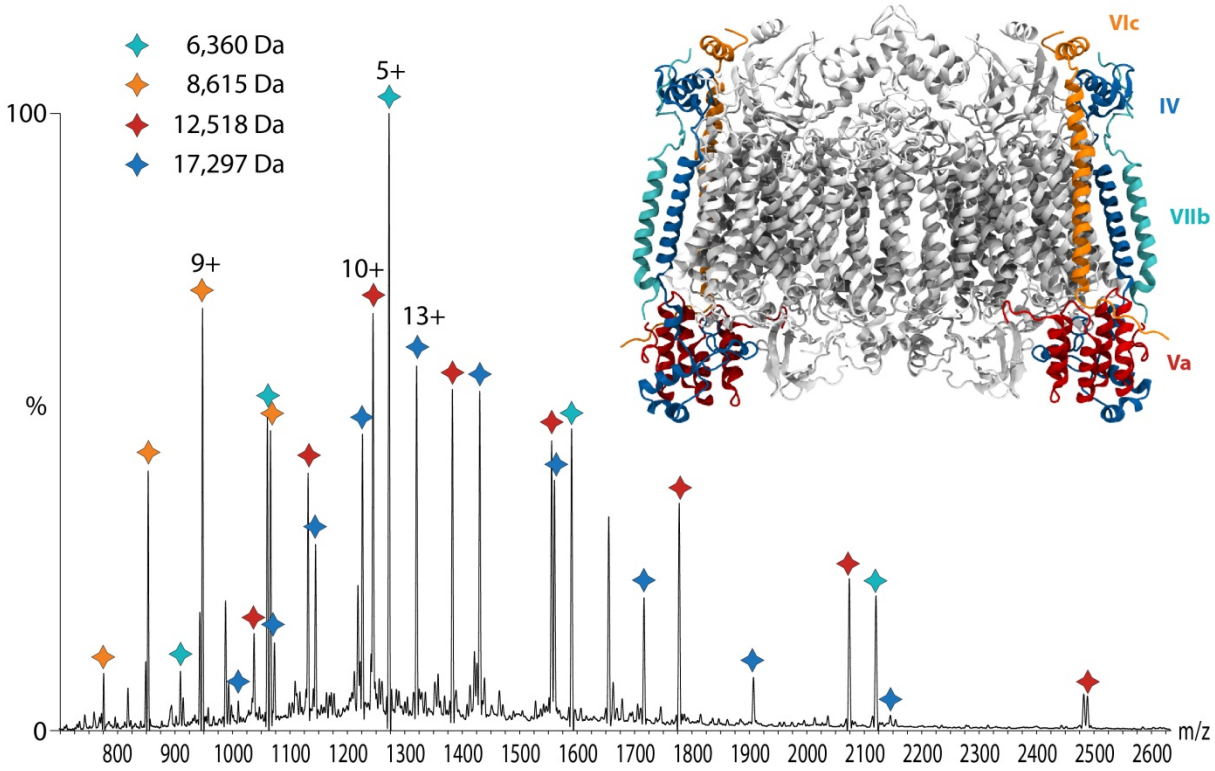

Figure 3

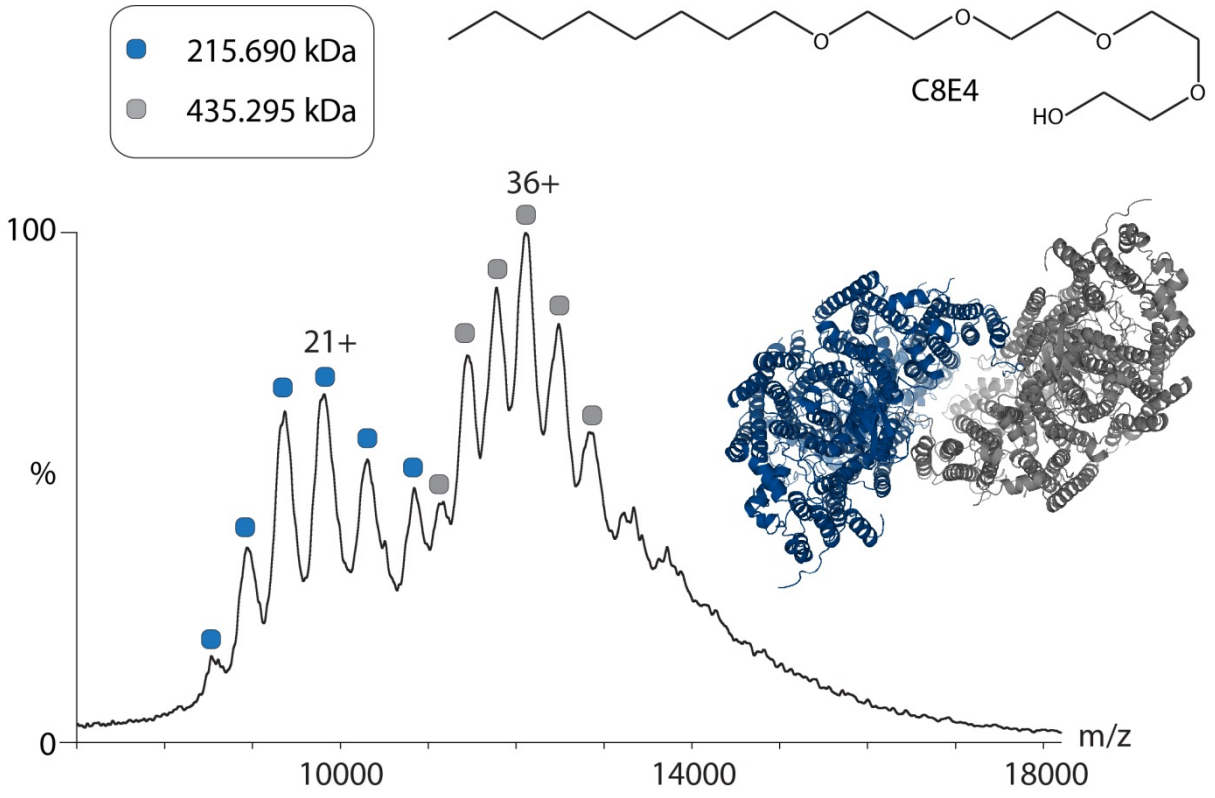

figure 4

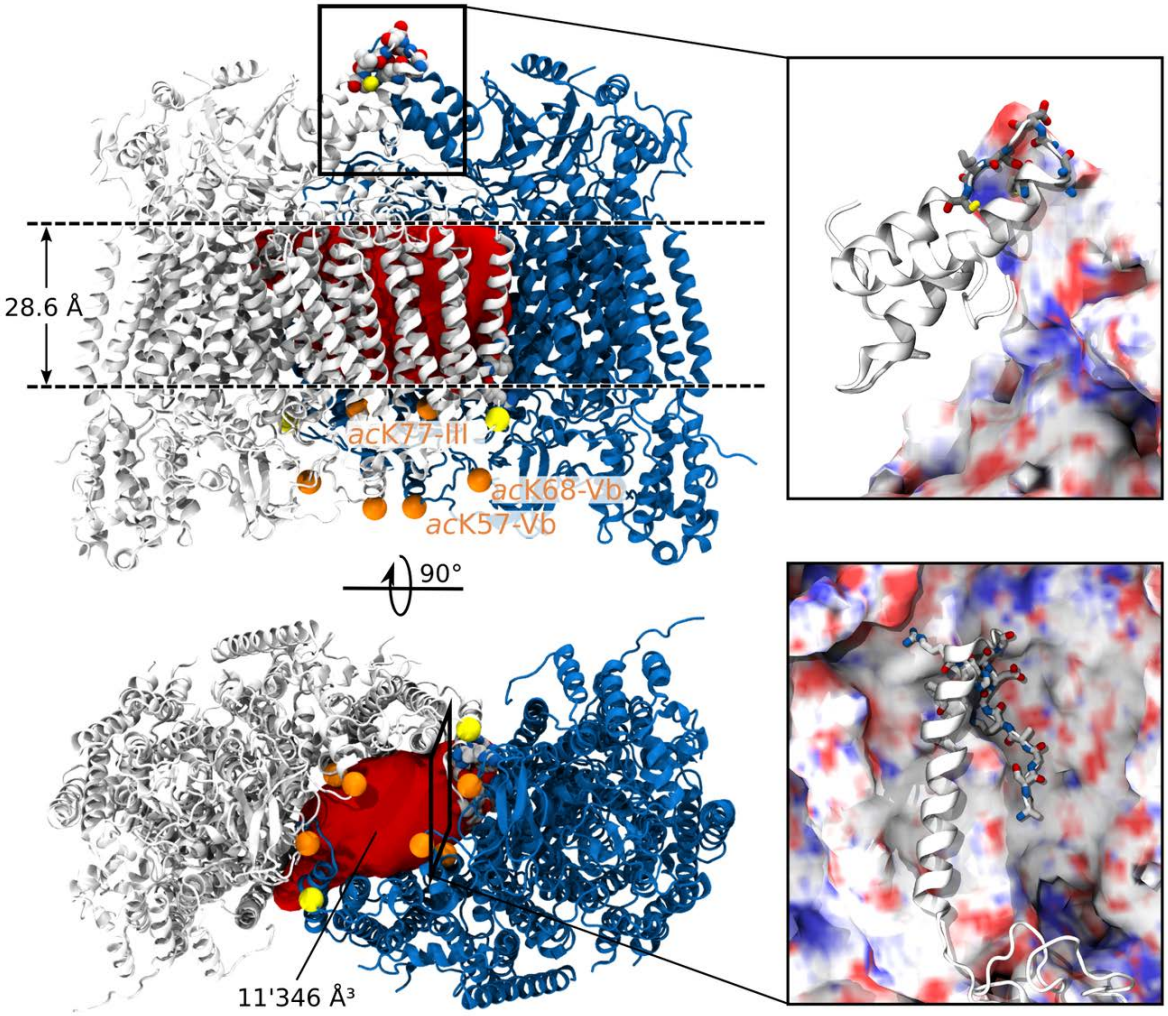

Figure 5