university of groningen

University Medical Center Groningen

University of Groningen

Evidence for Cardiolipin Binding Sites on the Membrane-Exposed Surface of the Cytochrome bc(1)

Arnarez, Clement; Mazat, Jean-Pierre; Elezgaray, Juan; Marrink, Siewert; Periole, Xavier

Published in:
Journal of the American Chemical Society

DOI:
10.1021/ja310577u

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2013

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Arnarez, C., Mazat, J-P., Elezgaray, J., Marrink, S-J., & Periole, X. (2013). Evidence for Cardiolipin Binding Sites on the Membrane-Exposed Surface of the Cytochrome bc(1). Journal of the American Chemical Society, 135(8), 3112-3120. DOI: 10.1021/ja310577u

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 11-02-2018

Supporting Information for

Evidence for Cardiolipin Binding Sites on the Membrane-Exposed Surface of the Cytochrome bc_1

Clement Arnarez[a], Jean-Pierre Mazat[b], Juan Elezgaray[c], Siewert-Jan Marrink[a], Xavier Periole[a,1]

[a]Groningen Biomolecular Sciences and Biotechnology Institute and Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands

[b]IBGC-CNRS UMR 5095 and Université Bordeaux Segalen, 1, rue Camille Saint-Saëns, 33077 Bordeaux, France

[c]Chimie et Biologie des Membranes et Nanoobjets, UMR 5248 – CNRS & Université Bordeaux 1, 2 rue Robert Escarpit, 33600 Pessac, France

[a1]To whom correspondence should be addressed. E-mail: x.periole@rug.nl

Content

1) Additional Results

a) CL preferential contact surface on bovine cytochrome *c* oxidase

A detailed description of the CL binding sites found on CIV_b has been published elsewhere.[1] Briefly, the densities of CLs (Fig. S4) obtained from the 100 μs simulation of wild type CIV_b (CIV_b-WT) demonstrate the existence of seven CL binding sites. The sites are labeled I_{CIVb} to VII_{CIVb}. The sites I_{CIVb} to V_{CIVb} are located on the matrix leaflet. The sites VI_{CIVb} and VII_{CIVb} are on the IMS leaflet. Sites V_{CIVb} and VII_{CIVb} are occupied by more than one CL and are therefore divided into Va_{CIV}/Vb_{CIV} and $VIIa_{CIV}$/$VIIb_{CIV}$.

b) Residue content of the $CIII_b$ CL binding sites

The analysis of the residue composition of the CL binding site based on the 100 μs simulation of $CIII_b$-WT revealed a few interesting features (Fig. S3). We first collected all the residues at least once in contact with a CL during the simulation. This set of residues basically covers the entire transmembrane region of the protein surface indicating that during the simulation the CLs explore the complete protein surface. Not surprisingly the distribution of contacts as a function of the residue type is in global agreement with the amino acid distribution in integral membrane proteins[2] keeping in mind that a residue would need to be exposed to the membrane to be in contact with a CL. 12% of these residues participate to the CL binding sites in $CIII_b$. Arginine is the most prominent residues in the CL binding sites (18%). Together with lysine they account for 22% of the CL ligands in $CIII_b$. This large contribution from positively charged residues might be expected since a CL carries a double negative charge. Serine residues also have a strong contribution to the binding sites found in $CIII_b$. The phenylalanine, alanine, leucine, threonine and tyrosine also populate the CL binding sites in $CIII_b$. Interestingly, only a limited fraction of the full set of contacts made by a residue type with CLs is actually part of the binding sites. This is most evident in the case of arginine residues, which have a high contribution to the binding sites but only 28% of arginines in contact with a CL participate to a site (Fig. S3). Overall these observations demonstrate that the CLs explored uniformly the full protein surface exposed to the membrane and that the interactions between CLs and the protein are specific. Notably not only electrostatic interactions are determinant for the binding of a CL, but various non-

charged residues are also important. This is in line with earlier reports focusing on CIV[3] and another study that showed that lipid tails were more resolved than the head groups in binding regions, and that electrostatic interactions (shut down by mutating specific charged residues) were playing a minor role in the definition of these binding sites.[4]

2) Extended Methods

a) Molecular models

Protein models: Models of CIII and CIV were based on their experimentally determined structures from bovine heart mitochondria ($CIII_b$, CIV_b). The *Protein Data Bank* (PDB) entries 1l0l,[5] 1sqb & 1sqq[6] and 2a06[7] were used to build a complete model of $CIII_b$ and 1occ & 2occ[8] for CIV_b. The CIII found in yeast mitochondria ($CIII_y$) was simulated for comparison. Its structure was taken from the PDB entry 3cx5[9] in which case the protein was fully resolved. Structures of bovine heart mitochondria $CIII_b$ available in the PDB contain 11 subunits per monomer: A, B, C (cytochrome *b*), D (cytochrome c_1), E (Rieske complex), F, G, H, I (related to the *Rieske* complex), J and K (see Fig. 1). The $CIII_y$ is composed of 9 subunits: A, B, C (cytochrome *b*), D (cytochrome c_1), E (*Rieske* complex), F, G, H and I. While yeast and bovine subunits A to E are similar, yeast subunits F, G, H and I correspond to bovine subunits H, F, G and I. Bovine subunits I and K are missing in $CIII_y$. CIV_b is composed of 13 subunits labeled A to M (see Fig. 1). Although both CIII and CIV are found in a dimeric form in the crystals only the dimer of CIII is thought to be functionally relevant and was thus simulated as such. CIV_b was simulated as a monomer. Details on the complexes and their subunits are given in Fig. 1.

In addition to the wild type of the protein complexes a couple of mutants of the bovine and yeast CIII dimer were simulated. A version of the $CIII_b$ with the subunit G truncated to remove its transmembrane section (residue G:Q23 to residue G:A75) was used to test the hypothesis that this particular subunit might protect bound CLs from exchanges with the bulk. Similarly in the $CIII_y$ the H subunit was truncated (residue H:K34 to residue H:V93). $CIII_y$ was also modeled with lysine to leucine mutations at positions 288, 289 and 296 of subunit D (corresponding to D:K227, D:K228 and D:K235 in the PDB entry 3cx5). These three lysine residues have been shown to be determinant for the function of $CIII_y$ due to their involvement in the bind of catalytic CLs.[10] We tested their potential role in the definition of the CL binding sites found in the WT simulations.

CLs found in the crystal structures: Two CLs per monomer are consistently found in the crystal structures of $CIII_b$. They are located within a cavity buried at the surface of the protein (CL1a and CL1b in Fig. 1). In contrast a single CL is found at the same location in $CIII_y$. These CLs are fundamental for a correct function of CIII and have been associated with the

proton transfer pathway.[10] In the manuscript they will be referred to as the catalytic CLs. Another CL is systematically found in the structures of $CIII_y$ (CL2 in Fig. 1). CL2 is deeply embedded in the membrane at the interface between the two monomers of the protein dimer. An additional CL is found in the PDB entry 1sqp of $CIII_b$ (CL3 in Fig. 1). CL3 is also embedded in the large transmembrane cavity network of the complex but seems more accessible to the bulk membrane.[6,11] This might rationalize its rare appearance in the structure files. All buried CLs (CL1a,b and CL2) were included in the simulations unless otherwise indicated, and are referred to as co-crystalized CLs.

In CIV_b two CLs per monomer co-crystallize with the protein (Fig. S4). The consideration of the symmetry of the dimer arrangement of the protein in the crystals results in three CL binding sites per monomer. One CL binding site (CL1 in Fig. S4) is located on a membrane-exposed protein surface, with its headgroup facing the matrix. Two more are located at the interface of the dimer of the complex (CL2 and CL3 in Fig. S4), with head groups facing the IMS. Since we simulate CIV_b in its monomeric form, the three CL positions described above are exposed to the membrane bulk and were therefore not included in the initial conformation of the system.

b) System set-up

Composition: A list of the systems simulated is given in Table S1 together with some key characteristics. Each system was placed in a rectangular box containing a solvated and pre-equilibrated POPC/CL membrane patch. A typical system (*e.g.* $CIII_b$-WT, Fig. 1) contains the protein (4230 residues; 9358 beads), a POPC bilayer (878 lipids; 11414 beads) including CLs (52 CLs; 1404 beads) and the aqueous phase (48835 water beads and 118 sodium ions), totalizing slightly more than 70000 CG beads, which typically runs at the speed of ~1 μs/day on a machine with 120×2.6 MHz CPUs.

Preparation: Starting from the protein structures described above, the models were converted to CG models and embedded into pre-equilibrated lipid bilayers using local scripts. CLs present in the crystal structures were used to fit CG CLs into the models. The systems were energy minimized, and relaxation periods were successively used for the solvent and side-chains. Each system was first simulated for 10 ns with positional restrains on backbone beads of the protein using a 10 fs integration time step before starting production runs.

Supercomplex formation: The initial conformations for the self-assembly simulations of complexes CIII and CIV into a supercomplex were built as follows: the individual complexes surrounded by their crown of lipids (CLs and POPCs) were extracted from the final conformations of the simulations of independent bovine complexes III and IV ($CIII_b$-WT and CIV_b-WT). The complexes were placed at ~10 nm from each other (surface to surface distance). Due to the dimeric structure of CIII, more statistics could be obtained by having two CIVs, one on each side of CIII. The three complexes were then embedded in a solvated and pre-equilibrated POPC bilayer and simulated for 5 μs. Ten repeat simulations were performed starting from different, randomized, initial orientations.

c) Simulation details

All simulations were performed using the GROMACS simulation package version 4.0[12]. The systems were described with the MARTINI CG force field for biomolecules (version 2.0),[13] its extension to proteins (version 2.1)[14] together with the ElNeDyn approach.[15] Conventional simulation setups associated with the use of the MARTINI force field were used. That includes an integration time step of 20 fs for production runs and non-bonded interactions cutoff at a distance r_{cut} = 1.2 nm. The Lennard-Jones potential is shifted from r_{shift} = 0.9 nm to r_{cut}. The electrostatic potential is shifted from r_{shift} = 0.0 nm to r_{cut}. The protein, membrane bilayer (POPC and CL) and aqueous phase (water and Na^+ ions) were coupled independently to an external temperature bath at 300 K using a Berendsen thermostat[16] with a relaxation time of 0.5 ps; the self-assembly simulations were carried out at a temperature of 350K to accelerate the association/dissociation events. The pressure was weakly coupled (Berendsen barostat)[16] to an external bath at 1 bar using a relaxation time of 1.2 ps and a semi-isotropic pressure scheme. CG sodium ions (solvated) were used to neutralize the systems. Parameters for CLs were taken from the work of Dahlberg *et al.*[17] The ElNeDyn approach, which defines an elastic network between the backbone beads to control the conformation of a protein, was used on each subunit of the complexes separately, *i.e.* springs are not present between the subunits. The integrity of the complex is only dependent on non-bonded interactions. Extend of the network was 0.9 nm and the force constant of the springs was set to 500 kJ mol^{-1} nm^{-2}.[15]

The complexes ($CIII_b$, $CIII_y$ and CIV_b) and their respective subunits were both numerically and structurally stable over time scales up to 100 μs. The root mean square deviation (RMSD) of the various versions of the respiratory chain complexes simulated and with

respect to the experimental structures typically reached values of 0.3-0.4 nm, indicative of structural stability. The definition of elastic networks on individual subunits of the complexes (ElNeDyn) was mandatory to maintain the integrity of the subunits although their dynamics and flexibility were conserved.

d) Atomistic simulation

To test whether the protein-lipid interactions might have been affected by the coarse-grained representation of the molecules, we also performed an MD simulation using an atomistic force field.

Methods. In the atomistic simulation, the peptides were described with the 54A7 parameter set of the GROMOS force field.[18] For the phosphatidylcholine lipids (POPC and CL) we used an in-house version of a new Gromos53A6-based lipid force field.[19] The SPC water model[20] was used. To obtain the starting structure for the atomistic simulations, the last frame of the 100 μs CG simulation of the $CIII_b$-WT system was transformed to the underlying atomistic representation using our recent resolution transformation algorithm.[21] During the resolution transformation, the system was cooled down from an initial temperature of 1000 K to the desired target temperature of 300 K by 30 ps of simulated annealing, during which the atomistic particles were coupled to their corresponding CG beads through harmonic restraints. Subsequently, the coupling was gradually removed within a time span of 28 ps. These annealing simulations were carried out in the NVT ensemble. No constraints were applied, and an integration time step of 1 fs was used. To control the temperature, stochastic coupling with an inverse friction constant τ_T = 0.1 ps was applied. The other parameters for the resolution transformation were set to the standard values.[21] Temperature and pressure coupling was applied similar to the CG simulations, with time constants τ_T = 0.1 ps and τ_p = 1 ps, respectively. Non-bonded interactions within 0.9 nm were updated at every time step, and interactions between 0.9 and 1.4 nm every 10 steps. The long-range electrostatic interactions were computed with the PME algorithm.[22]

System details. The snapshot after t = 100 μs from the CGMD simulation of the $CIII_b$-WT system was transformed to the atomistic representation, followed by a 100 ps simulation with position restraints applied to the backbone, and a free 100 ns atomistic MD simulation. Including protein, lipids, and water molecules, the system comprised of about 700,000

particles.

d) Analysis details

The first μs of the simulations is considered as an equilibration period and is thus not included in the analysis. Repeat simulations are first analyzed separately to judge the convergence of the observables and when appropriate repeat simulations are concatenated to obtain better statistics. Trajectory snapshots were saved every 500 ps for analysis purpose unless otherwise indicated.

Binding site definition. The densities of the CLs were used to localize the potential binding site at the protein surface. The densities of presence of the CLs in the systems were computed by means of the "occupancy" option of the VolMap tool of VMD[23] with a 0.2 nm resolution grid. The densities depict an occupancy level at least 5 times higher than the values corresponding to the bulk region. CLs binding sites on the protein were defined by the following protocol. For each frame of the 100 μs simulations of the WT complexes we define time series of groups of residues simultaneously in contact with a unique CL. A contact is counted when a residue found within 0.7 nm of the CL head group (phosphatidyl and glycerol CG beads). The groups obtained are then clustered using a GROMOS-like algorithm:[24] groups with at least 66% of common residues are clustered together and the most populated cluster (*i.e.* the groups composing it) is removed from the pool of groups; the procedure is repeated on the remaining groups until the pool is empty. The residues present in at least 80% of the groups composing a cluster are selected as the representative group of the cluster. The symmetry of the dimer was used as additional refinements to define the final clusters for $CIII_b$. This was only applicable to the sites II_{CIIIb}, III_{CIIIb} and IV_{CIIIb}; the instability of site I_{CIIIb} in one monomer and the absence of symmetry in the cases of sites V_{CIIIb} and VI_{CIIIb} made it difficult to define unique sites. Note that several other protocols were tested for the definition of the binding sites. The final definition of the sites was only slightly sensitive to the protocol used.

Binding site averaged occupancy and residence times: To further characterize the binding sites, their CL occupancy and residence time are determined. The occupancy of a site *j* by a CL L at a time t, noted $\xi_L^j(t)$, is determined as:

$$\xi_L^j(t) = \begin{cases} 1, & d_{jL}(t) \leq \delta_L \\ 0, & d_{jL}(t) > \delta_L \end{cases}$$

where $d_{jL}(t)$ is the distance between the center of mass of the CL head group and the center of mass of the binding site *j*; δ_L is a cutoff value set to 0.8 nm. The slight difference between δ_L and the cutoff used to define a contact in the definition of the binding sites (0.7 nm) accounts for the relative large size of a CL head group. The instantaneous values of the distance are smoothed out over windows of 50 ns. The time series of the site occupancies are additionally smoothed out by averaging over 2.5 ns time windows, which results in an effective averaging of ~55 ns. These precautions were necessary to discard the fast leave and return of CLs from the site considered ("rattling in a cage" type of motion). The error on the occupation values was estimated by splitting the trajectory in 10 slices of 10 μs, computing the average occupation for each slice and the standard deviation to the mean occupation over 100 μs. The values presented in Tables 1 are thus reported with a maximum error of ±0.02.

The residence time (also referred to as the survival or lifetime,[25,26] θ_L^j, of a CL L in a binding site *j* is obtained from the normalized time-correlation, $\sigma_L^j(t)$, which can be expressed as:

$$\sigma_L^j(t) = \frac{1}{T-t}\sum_{\upsilon=0}^{T-t} \rho_L^j(\upsilon, \upsilon + t)$$

where T is the total time of simulation and $\rho_L^j(\upsilon, \upsilon + t)$ a function taking the value of 1 if the CL L has occupied the site *j* for a time t after coming in contact with the site at time υ, and 0 otherwise. ρ_L^j is obtained from the time series of the site's occupancy by a CL described above. In practice the CLs are not differentiated since they occupy a site one at the time and as before a 2.5 ns window averaging was used to discard the fast leave and return movements of CLs. The residence time of CLs in a specific site *j*, θ^j, is obtained by fitting the time-correlation function relative to this site using:

$$\sigma^j(t) \approx \exp\left(-\frac{t}{\theta^j}\right)$$

Note that the values of the CL's residence time were quite sensitive to the exact protocol used. Notably the smoothing of the raw data to remove the quick binding/unbinding events was found important to capture the relevant behavior of the CL. Alternatively, when the number of CLs exchanging in a site was not sufficient to obtain a statistically relevant $\sigma^j(t)$, the residence time of CLs for a specific site was estimated by collecting the stretches of time during which that site is occupied.

3) Supporting Tables

Table S1: Summary of the simulations performed. The name, number (#), length, CL/POPC molar ratio and some specifications are given.

Name	#	Length [μs]	CL/POPC ratio	Specifications
$CIII_b$-WT	1	100	1/17	Co-crystallized CL included
$CIII_b$-EmpCav	2	10	1/18	Catalytic CLs removed from site I_{CIII}
$CIII_b$-ByLf	2	10	1/37	CLs in only one leaflet at once Catalytic CLs removed from site I_{CIII}
$CIII_b$-woG	2	10	1/17	Partial removal of subunit G
$CIII_y$-WT	1	100	1/19	Co-crystallized CL included
$CIII_y$-EmpCav	2	10	1/19	Catalytic CLs removed from site I_{CIII}
$CIII_y$-woH	2	10	1/19	Partial removal of subunit H
$CIII_y$-Mut	2	10	1/19	Mutated complex Catalytic CLs removed from site I_{CIII}
$CIII_y$-mut-woH	2	10	1/19	Mutated complex Partial removal of subunit H
CIV_b-WT	1	100	1/18	Monomeric complex
$CIII_b$-WT/CIV_b-WT	10	5	1/30	Sites already populated

Table S2: Occupation of the sites I_{CIII} by CLs in bovine heart and yeast mitochondria. The average error on the values reported is ±0.01 and was estimated on the 100 μs simulations (see Methods section, Table S1). Note that the subsites a and b of site I_{CIII} are not considered in this analysis. Occupations are given for site I_{CIII} as a single cavity. The labels of the simulations are given in Table S1. $^{\#}$reflects the entrance of an additional CL in site I as depicted by the CL densities in Fig. S1.

Simulation		$CIII_b$-WT	$CIII_b$-woG		$CIII_y$-WT	$CIII_y$-woH		$CIII_y$-mut-woH	
			sim 1	sim 2		sim 1	sim 2	sim 1	sim 2
Ξ_L	mer 1	1.99	1.07	0.98	1.00	0.85	0.83	0.07	0.03
	mer 2	2.00	0.76	1.25	$1.66^{\#}$	0.91	0.64	0.23	0.02
	aver	2.00	1.02		1.33	0.81		0.09	

Table S3: Statistics on the dynamical properties of $CIII_b$ sites. For each site, occupation (Ξ), residence time (θ, μs), the number of CL involved and binding/unbinding events are provided for each CL binding to a site.

Site		Ia	Ib	II	III	IVa	IVb
Ξ	mer 1	0.99	1.01	0.91	0.69	0.63	0.73
	mer 2	0.95	1.71[#]	0.96	0.66	0.91	0.71
θ	mer 1	> 100	> 100	7.4	0.5	2.6	0.8
	mer 2	> 100	> 100	3.5	2.3	1.4	0.9
#CLs	mer 1	1	2	10	11	7	17
	mer 2	1	2	5	12	9	15
#events	mer 1	12	19	40	189	151	246
	mer 2	65	167	30	90	147	196

Table S4: List of residues forming the CL's binding sites. They correspond to the residues shown in Figure 2 in the main manuscript. Residues carrying a net charge on the side chain are highlighted in blue (positive) and red (negative).

Site	Subunit	Residues
Ia	C	S30, W31, T42, C45, R48, V49, Y104, Y105, P209, M317, Q323
	F	S71, M72, Q74
	G	R41
Ib	C	S30, L231, L235
	D	Y304, K307, R308, W311, K315
	G	Y30, F31, G34, N37, V38, R41
II	C	P320
	F	A25, G27, K30
	G	R43, C44, C45, L47, R48
III	E	R110, F113
	G	A26, F27, H29, Y30
IVa	A	Y381, E440, R442, S446, Y450, F480
	J	S12, L13, R16, R17, T20
	K	P19
IVb	A	S382, A383, T384, A438, R439, R442
	K	R11, R15

4) Supporting Figures

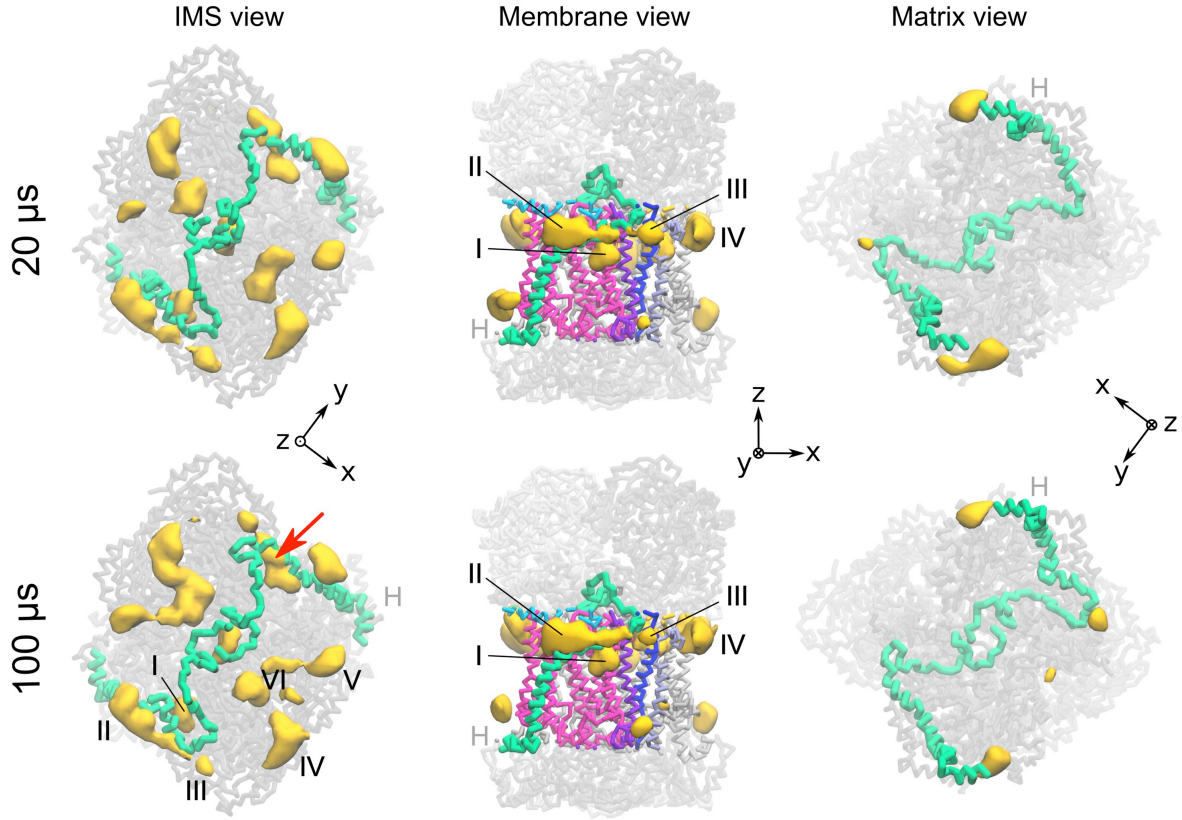

Figure S1: CL densities on the $CIII_y$ shown from the matrix (left), membrane (middle) and inter membrane space (IMS, right) views. The top and bottom rows show the CL densities obtained from an analysis of the first 20 µs and 100 µs, respectively, of the simulation $CIII_y$-WT simulation (Table S1). The comparison of the first 20 µs and 100 µs shows the evolution of the CL densities in time. Of particular interest is the site I_{CIIIy} in which case a CL penetrates the cavity at ~25 µs (red arrow).

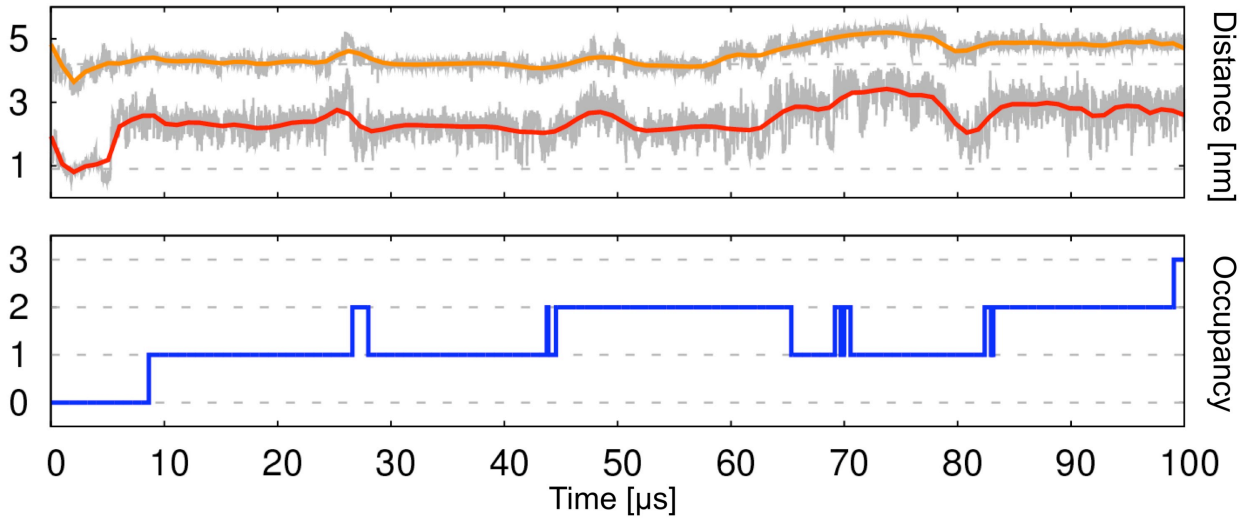

Figure S2: Dynamic of the second site VI of $CIII_b$. It is complementary of Figure 3A,B in the main manuscript. The top panel depicts a pair of distances representative of the motion of two subsections of subunit K (the color code follows the one in the snapshots of Fig. 3A). The bottom panel shows the occupancy of the inner cavity.

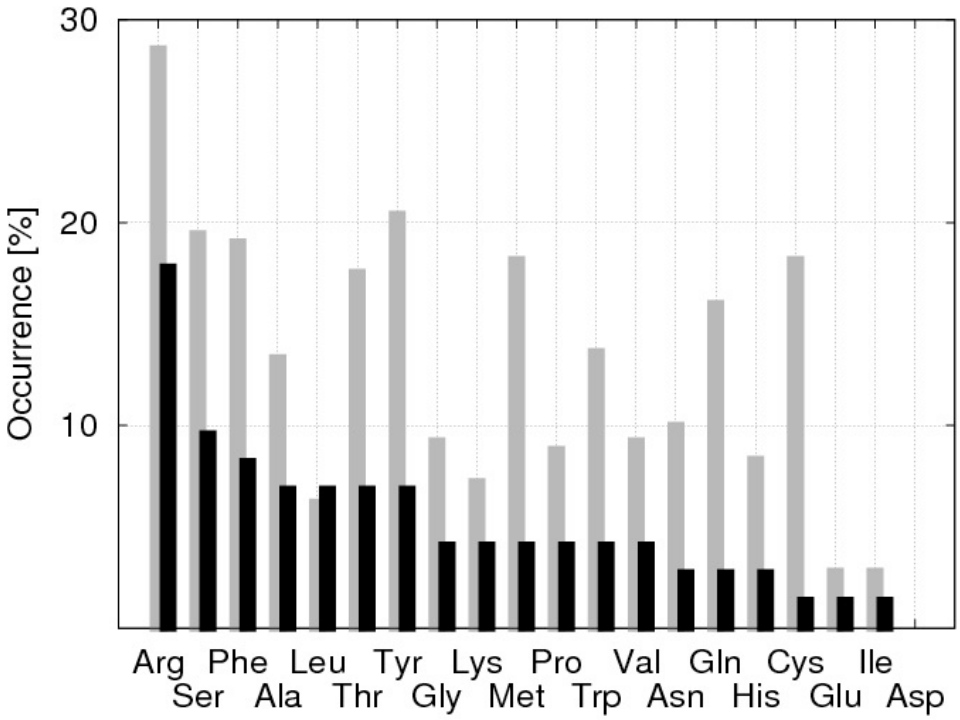

Figure S3: Residue content of the CL binding sites of $CIII_b$. The gray sticks indicate for each residue type the percentage from all the residues of that type that makes a contact with a CL during the simulation. In other words the proportion of this type of residue that is found in the section of CIII visited by CLs: the transmembrane section of CIII exposed to the membrane. The black sticks give for each residue type the percentage of its participation to the CL binding sites. All the residues of the protein were used for the analysis.

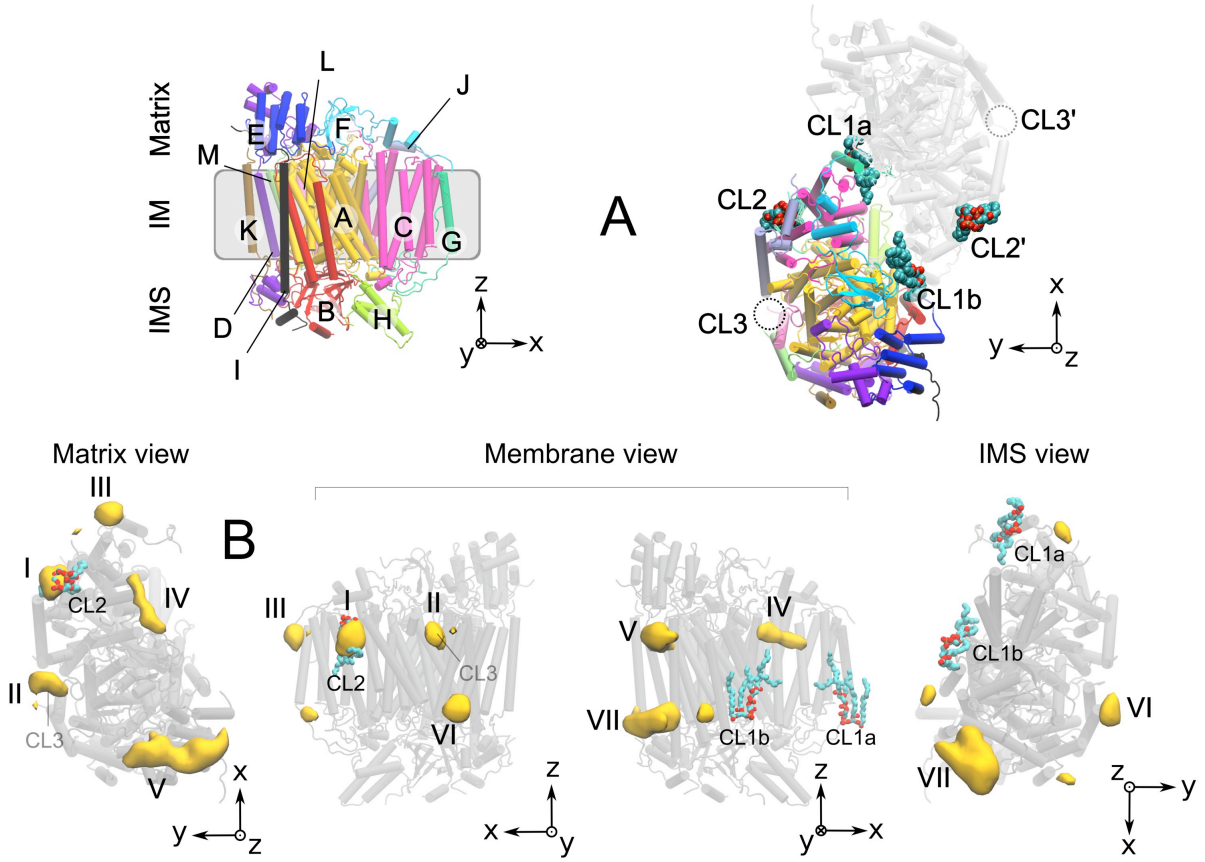

Figure S4: Structural characteristics of the wild type cytochrome *c* oxidase (CIV_b) and its CL binding sites. **(A)** left side: structure of CIV_b with its thirteen subunits (A-M) shown in a cartoon representation and a chain-based color-code; right side: the dimer structure of CIV as found in the X-ray structure is depicted with the CL attached to it. There are three CL binding sites per monomer (CL1, CL2 and CL3). **(B)** Binding sites of the CLs identified by their densities are shown in yellow surfaces at a level of density at least five times higher than in the bulk. Four orientations are given.

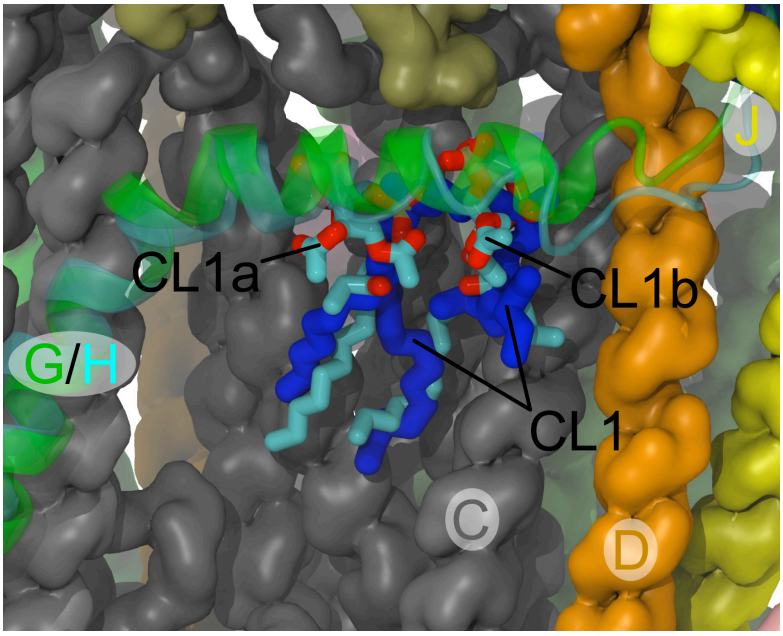

Figure S5: Comparison of the CL content of the site I_{CIII} in the complex III from bovine heart and yeast. The yeast CIII structure (PDB-ID 3cx5) is used for the main subunits involved, which are labeled with their names (C, D and J) and shown in a gross tube representation. The subunit H in yeast (cyan) and its homologue G in bovine heart (PDB-ID 1sqb) are shown using a cartoon representation. The subunit in the two spices has the same orientation. The subunit C of both species was used for the overlay the two structures. The single CL1 (shown in blue sticks) found in the yeast structure occupies the space as CL1a and CL1b (cyan carbon and red oxygen sticks) in $CIII_b$. Note that the orientations of the CL's head are different.

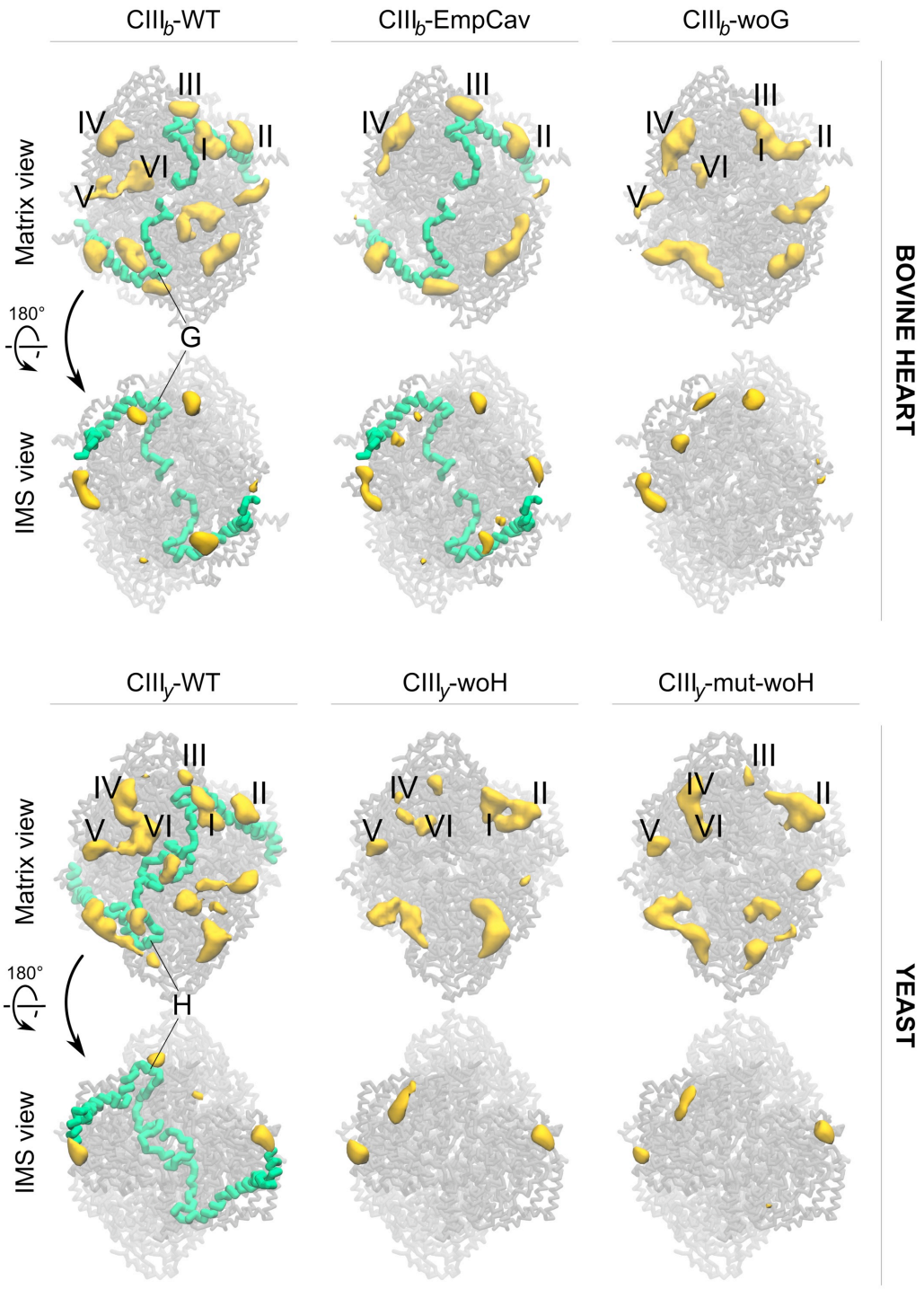

Figure S6: CL densities in simulations of the CIII$_b$ (top) and CIII$_y$ (bottom) of different systems simulated for this work. Two views are reported for each case: from the matrix (first and third rows) and from the inter membrane space (IMS, second and fourth rows). The first column reminds the CL densities obtained from the 100 μs simulations (main Fig 2), while the other two were extracted from a combination of two independent simulations of 10 μs (see Table S1). Similar densities were also observed in the simulations CIII$_b$-ByLf, CIII$_y$-EmpCav, CIII$_y$-Mut (data not shown).

5) Supporting References

(1) Arnarez, C.; Marrink, S.-J.; Periole, X. *Sci. Rep. in press*
(2) Ulmschneider, M.; Sansom, M. S. P. *Biochim. Biophys. Acta* **2001**, *1512*, 1–14.
(3) Qin, L.; Hiser, C.; Mulichak, A.; Garavito, R. M.; Ferguson-Miller, S. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 16117–16122.
(4) Aponte-Santamaria, C.; Briones, R.; Schenk, A. D.; Walz, T.; de Groot, B. L. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 9887–9892.
(5) Gao, X.; Wen, X.; Yu, C.-A.; Esser, L.; Tsao, S.; Quinn, B.; Zhang, L.; Yu, L.; Xia, D. *Biochemistry* **2002**, *41*, 11692–11702.
(6) Esser, L.; Quinn, B.; Li, Y.-F.; Zhang, M.; Elberry, M.; Yu, L.; Yu, C.-A.; Xia, D. *J. Mol. Biol.* **2004**, *341*, 281–302.
(7) Huang, L.-S.; Cobessi, D.; Tung, E. Y.; Berry, E. A. *J. Mol. Biol.* **2005**, *351*, 573–597.
(8) Tsukihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Shinzawa-Itoh, K.; Nakashima, R.; Yaono, R.; Yoshikawa, S. *Science* **1996**, *272*, 1136–1144.
(9) Solmaz, S. R. N.; Hunte, C. *J. Biol. Chem.* **2008**, *283*, 17542–17549.
(10) Lange, C.; Nett, J. H.; Trumpower, B. L.; Hunte, C. *EMBO J.* **2001**, *20*, 6591–6600.
(11) Palsdottir, H.; Hunte, C. *Biochim. Biophys. Acta* **2004**, *1666*, 2–18.
(12) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
(13) Marrink, S.-J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. *J. Phys. Chem. B* **2007**, *111*, 7812–7824.
(14) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S.-J. *J. Chem. Theory Comput.* **2008**, *4*, 819–834.
(15) Periole, X.; Cavalli, M.; Marrink, S.-J.; Ceruso, M. A. *J. Chem. Theory Comput.* **2009**, *5*, 2531–2543.
(16) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
(17) Dahlberg, M. *J. Phys. Chem. B* **2007**, *111*, 7194–7200.
(18) Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. *Eur Biophys J* **2011**, *40*, 843–856.
(19) Schäfer, L. V.; de Jong, D. H.; Holt, A.; Rzepiela, A. J.; de Vries, A. H.; Poolman, B.; Killian, J. A.; Marrink, S.-J. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 1343–1348.
(20) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. *Intermolecular Forces*; Reidel Publishing Company, Dordrecht, The Netherlands, 1981.
(21) Rzepiela, A. J.; Schäfer, L. V.; Goga, N.; Risselada, H. J.; de Vries, A. H.; Marrink, S.-J. *J. Comput. Chem.* **2010**.
(22) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
(23) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.
(24) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. *Angew. Chem. Int. Edit.* **1999**, *38*, 236–240.
(25) Garcia, A. E.; Stiller, L. *J. Comput. Chem.* **1993**, *14*, 1396–1406.
(26) Periole, X.; Rampioni, A.; Vendruscolo, M.; Mark, A. E. *J. Phys. Chem. B* **2009**, *113*, 1728–1737.