Supplementary Information for

# Molecular mechanism of cardiolipin-mediated assembly of respiratory chain supercomplexes

C. Arnarez[a,†], S.J. Marrink[a], X. Periole[a,*]

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

[†] Current address: Department of Physics and Astronomy, University of Delaware, 258 Sharp Lab, 19712 Newark, Delaware, United States of America

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

**Content:**

## 1. Extended Methods

**Molecular models.** The models of CIII and CIV are identical to the ones used in our previous studies (1, 2). Briefly, CIII's dimer was built from a combination of four experimental structures (PDB entries: 1l0l (3), 1sqb/1sqq (4) and 2a06 (5)), excluding the six hemes and two iron-sulfur clusters. CIV's (monomer) was built from the PDB entries 1occ and 2occ (6), also excluding two deeply buried hemes. CG conformations of CIII and CIV together with their annular lipid shell were extracted from our recent 100 μs simulations from which CL binding sites on the individual complexes were identified (1, 2) (Fig. 1).

**Molecular systems.** 9 CIII dimers and 27 CIVs and their annular lipid shell were placed on a grid and randomly rotated along the membrane normal (Fig. 1). A 1:15 CL:POPC molar ratio lipid bilayer was used to mimic the mitochondrial membrane, corresponding to ~13% phosphorus content due to CL (6.5 molar %), within the range of experimental values (10 to 20% (7, 8)). In the case of the system devoid of CL in the bulk, two CLs were conserved in sites Ia and Ib of CIII (1) to prevent the otherwise collapse of subunit G. The two membrane environments have identical phosphorus content.

**Simulation details.** Simulations were performed using the GROMACS simulation package version 4.0 (9) and the Martini CG force field 2.0 for biomolecules (10) and its extension 2.1 to proteins (11) together with the *ElNeDyn* approach (12). The Martini model is well suited to study a large variety of membrane related processes (13), including membrane protein self-assembly (14, 15).

Conventional simulation setups associated with the use of the Martini force field were used. The trajectories described in this work were ran using an integration time step of 20 fs for production runs (10 fs for equilibration) and non-bonded interactions cutoff at 1.2 nm with the Lennard-Jones potential shifted from 0.9 to 1.2 nm. The electrostatic potential is shifted from 0.0 to 1.2 nm and a relative dielectric screening constant of 15 is used. 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 pressure was weakly coupled (Berendsen barostat (16)) to an external bath at 1 bar using a relaxation time of 1.2 ps following a semi-isotropic pressure scheme. CG sodium ions were used to neutralize the systems. CL topology and parameters 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 $kJ.mol^{-1}.nm^{-2}$ (12). Trajectory snapshots were saved every 1 ns for the analyses. The simulations proved the complexes and their subunits to be numerically and structurally stable over the time scale of the simulations.

**Supercomplex characterization.**

*Protein contacts.* Buried protein surfaces were used to characterize protein-protein contacts. Buried surfaces or protein burials between complexes $i$ and $j$, $a_{ij}$, were computed as $a_i + a_j - a_{ij}$ where $a_i$ is the accessible surface of the complex $i$ isolated and $a_{ij}$ the accessible surface of complexes $i$ and $j$ assembled. Buried surfaces were computed used the *g_sas* tool available in the GROMACS suite, which uses the double cubic lattice method (18). The probe radius was fixed to 0.26 nm, which corresponds to the van der Waals radius of a Martini bead. Alternatively, the number of contacts between complexes was used as a metric for protein-protein interactions. It is a considerably less computationally demanding approach and reports on the same observable (Fig. S13). In addition it allows a precise quantification of different contributions for protein-protein interactions. Protein-protein contacts were cutoff at 0.8 nm. The differentiation between intramembranous and extramembranous parts of each complex was based on a simple rule: if contacts with lipids are formed at any moment of the previously analyzed 100 μs trajectories (1, 2), the residue is assumed as being part of the intramembranous section.

*Lipid contacts.* In the determination of the lipid content at the surface of the individual complexes and their interface in a supercomplex, a lipid was counted in contact if its head group (PO4 or GL5 beads for POPC and CL, respectively) was within 1.0 nm of the protein. As in our previous studies (1, 2), CL was considered to occupy a binding site if the center of mass (COM) of its headgroup (glycerol plus phosphatidyl moieties) is within 0.8 nm from the COM of the binding site. Interfacial lipids are lipids that are in contact with at least one of protein interfaces, while shared lipids are in contact with both partners. The protein interface was defined for each complex by the set of residues in contact with the annular lipids (< 1.0 nm) common to both complexes (< 1.0 nm) and averaged over the last μs of the simulation.

*Supercomplex architecture.* The architecture of a SC was defined by the relative orientation of the two complexes following the virtual bond analysis (19) as we did previously in the case of rhodopsin (15). $\Phi_1$ and $\Phi_3$ define the orientation of CIII and CIV relative to CIV and CIII, respectively, in the plane of the bilayer (Fig. S14A). Alternatively the relative orientation of the complexes in a SC can be monitored from the location of the COM of the protein interfaces on both complexes defining $\gamma_1$ and $\gamma_3$ (Fig. S14B). The two definitions provide similar results (see Fig. S14C,D) although slightly different information.
In the case of a membrane bilayer the motion of proteins are in two dimensions, reducing the description of their relative orientation to one distance and two angles. These were monitored using three anchor points on each complex. For CIII, these three reference points were defined as the COM of the sites $I_{CIII}$ and $I'_{CIII}$, the COM of the four subunits A, A', B and B' and the COM of site $I_{CIII}$. For CIV, the COM of the subunits A, F and C were respectively used. The $C_2$ symmetric axis of CIII dimer was taken into account while determining the CIII/CIV relative orientations.

*CL binding sites contact map.* The involvement of CL binding sites in the interface between complexes was quantified by measuring the distances between sites over the simulation. Two sites were assumed to be in contact or shared between the two complexes when the distance between their COM was less than 3.0 nm.

*CL non-binding surface.* We compared lipid-binding affinities on CL binding sites (specifically on $II_{CIV}$) to a non CL-specific surface of the protein using free energy calculations (described in the next section). The non-binding surface was defined as a portion of the protein (CIV) between sites $IV_{CIV}$ and $Vb_{CIV}$. A set of residues was manually selected by visual inspection of the protein surface. This surface was chosen on the matrix side in such a way that while allowing a clear separation from the neighboring sites ($IV_{CIV}$ and $Vb_{CIV}$), it keeps the similarity with site $II_{CIV}$ to which it is compared. The center of mass of the backbone beads of these residues was used as reference point for the pulling calculations (see below). The selected residues cover 4 subunits of CIV and are: A:Gly269, A:Met271, B:Ala58, E:His2, E:Ser4, E:His5, F:Glu63, F:Glu64 and F:Thr68; reported as chain:residue.

**Lipid binding strength.** Potentials of Mean Force (PMFs) were used to quantify the binding affinities or binding strength of lipid molecules to CL binding sites and non CL-specific surface of CIV. PMFs were calculated using the umbrella sampling technique following our

previous protocol (2). The system used was composed of a CIV embedded in a patch of ~360 POPC lipids and one molecule of the lipid of interest. We tested the binding of CL with various charges (from -2 *e* to 0 *e*) and tail contents (mono- and dilyso-CL), and a range of phospholipids common to the inner mitochondrial membrane: POPG (from -2 *e* to 0 *e*), POPS (-1 *e*), POPE and POPC molecules. For each case we ran simulations with a set of umbrella potentials that covered from the lipid-bound situation to free in the bulk membrane; ~2.5 nm from the bound state. This distance assures that the profile reaches a plateau value. The umbrella potentials were applied to the distance between the center-of-mass of the lipid head group and of the set of residues defining the binding site with a 1000 kJ $mol^{-1}$ $nm^{-2}$ force constant. These choices resulted in 20 to 30 umbrella simulations spaced by 0.1 nm. We reduced the positional and rotational motions of the protein by applying a weak position restrain on three backbone beads of the protein. These position restraints were not applied on the z coordinate and used a 200 kJ $mol^{-1}$ $nm^{-2}$ force constant (0.42 kcal $mol^{-1}$ $Å^{-2}$). We applied the distance restraint between the centers-of-masses only along the direction normal to the protein surface. Therefore, an additional weak position restraint (harmonic potential, 100 kJ $mol^{-1}$ $nm^{-2}$ force constant) was applied on one bead of the lipid head group (GL5 for CL and PO4 for regular lipids) to restrict its exploration of the direction parallel to the protein surface (perpendicular to the pulling direction). The direction normal to the membrane plane was let free. Each umbrella simulation was run for 6 μs. The weighted histogram analysis method (WHAM) (20) was used to combine and unbias the simulations to produce the free energy profiles. The Bayesian bootstrapping method was used to estimate the statistical error on the calculation. Each umbrella position was treated independently for the bootstrapping.

## 2. Limitations of our approach

The CGMD simulations presented in this work provide an unprecedented view of protein supramolecular formation and organization in biological membranes. Most notably, we revealed the subtle role of a particular lipid, CL, in steering the self-assembly of proteins, CIII and CIV, and provided a detailed view of the lipid content of the protein-protein interface.

It is however important to point at a few limitations to our approach. Despite the simplification brought by the removal of some degrees of freedom in our CG approach, the systems simulated remain extremely large and could be studied at only a fraction of the

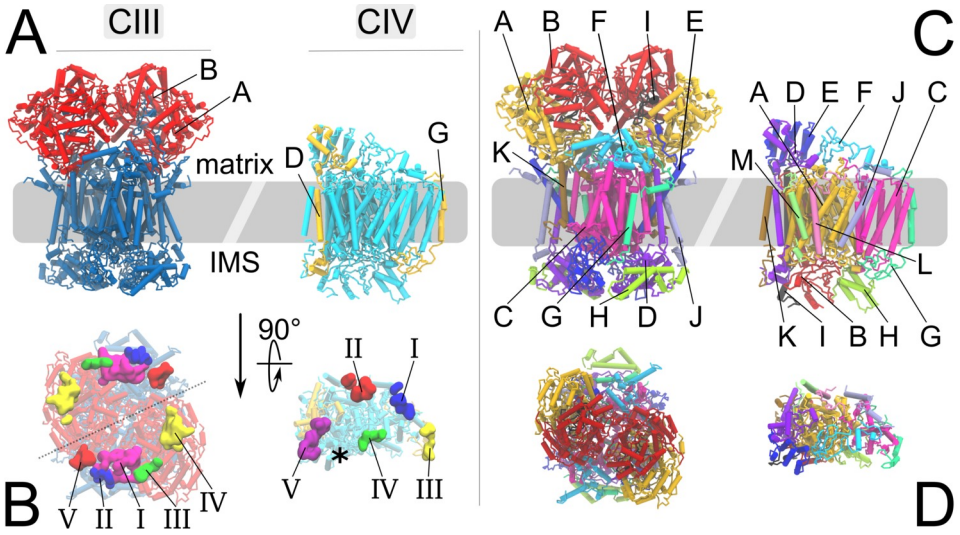
biological time scales. Therefore, our 20 μs trajectories unravel only parts of the effective mechanism of supercomplex formation, and an equilibrium state has not been reached. In addition, though our system represents one of the largest membrane patches simulated till date and includes 36 individual protein complexes, the statistics on protein-protein contacts remains necessarily limited. The number of CIII/CIV and CIV/CIV interfaces the simulations provided amount to 21 and 9 in presence of CL, and 19 and 16 without CL, respectively. Although self-assembly simulations do allow for identifying relevant protein-protein interfaces, the lack of binding/unbinding events prevents assessing their relative strength even with a large set of complexes (15). More sophisticated approaches are needed to assess the relative stability of proteins interfaces reliably (21), but these techniques are unfortunately not currently practicable on systems of size studied in this work.

Furthermore, the accuracy of the representation of protein-lipid and protein-protein interactions is potentially limited by the use of a CG model. However, the Martini model has been shown to reproduce experimental lipid binding sites on an increasing number of membrane proteins (1, 2, 22, 23). Another important issue of our models is the potential over-stabilization of protein-protein interactions in aqueous solution, reported for both atomistic (24) and CG approaches (25). This might have slowed down the reorientation of the complexes after their encounter in our simulation. But note that protein interaction strengths are relatively well modeled when in the bilayer. We have reported studies of the association of GpA (26) and WALP peptide (27) in model membranes using the Martini CG model and found that the free energy profile of the GpA peptide was essentially identical to one reported earlier using an atomistic force field (28) and that the estimated dimerization free energy of the WALP peptides agreed with the value obtained from fluorescence resonance energy transfer experiments (29).

A further limitation is the simplified composition of the mitochondrial membrane in our simulations. Experiments show that the association and function of the respiratory chain supercomplexes depends on the complexes and cofactors available in the membrane (30). We did not include complex I (CI), for which the structure was resolved at atomic resolution only recently (31), and not for bovine heart mitochondria but *Thermus thermophilus*. Our simulations also do not include the Rcf1 complexes that seem to be involved in the formation of the SC, as recently observed in yeast (32-34). It is not clear how these omissions might have affected the interfaces observed in the present study. Furthermore, we have assumed

POPC to be a good representation of the mitochondrial lipid matrix, thus omitting lipids with a PE head group that are found in equal amount as PC lipids (7, 8). PE lipids appear not to be required for SCs stability (35). *In vitro*, although CL-dependent reconstitution of yeast SCs does not require PE, no larger megacomplexes are formed in the absence of PE (W. Dowhan, personal communication). Preliminary analysis of a system with a membrane composed of POPC:POPE:CL at a 15:15:2 molar ratio indicates that large scale SC formation is also observed in the presence of PE (Fig. S8). Furthermore, the PMFs of CL and POPG reporting on their binding strength to site II of CIV (Fig. 4A) were not significantly affected by the presence of PE lipids at a 1:1 molar ratio with PC (Fig. S9), suggesting that POPC is a good model for our study. A more detailed analysis of the effect of PE will be presented elsewhere.

## Supplementary Figures




| A | B | C | D | E | F | G | H | I | J | K |
|---|---|---|---|---|---|---|---|---|---|---|
| I | II | cyt *b* | cyt $c_1$ | Rieske | VI | VII | VIII | IX | X | XI |
| UQCRC1 | UQCRC2 | CYTb | CYC1 | UQCRFS1 | UQCRH | UQCRB | UQCRQ | UQCRFS1 | UQCR10 | UQCR10 |

| A | B | C | D | E | F | G | H | I | J | K | L | M |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| I | II | III | IV | Va | Vb | VIa2 | VIb1 | VIc | VIIa1 | VIIb | VIIc | VIIIb |
| COX1 | COX2 | COX3 | COX4-1 | COX5a | COX5b | COX6a2 | COX6b1 | COX6c | COX7a1 | COX7b | COX7c | COX8b |

**Figure S1:** Molecular system. **(A)** Lateral view of the bovine cytochrome $bc_1$ (complexe III; CIII) and cytochrome *c* oxidase (complex IV; CIV), as reported in Figure 1. **(B)** Top view of the same complexes, with binding sites as described in our previous studies (1, 2). **(C)** and **(D)** Details of the subunits composing CIII and CIV, lateral and top view respectively. Bottom tables reports the various nomenclatures (36) used for protein subunits for CIII (top table) and CIV (bottom table).

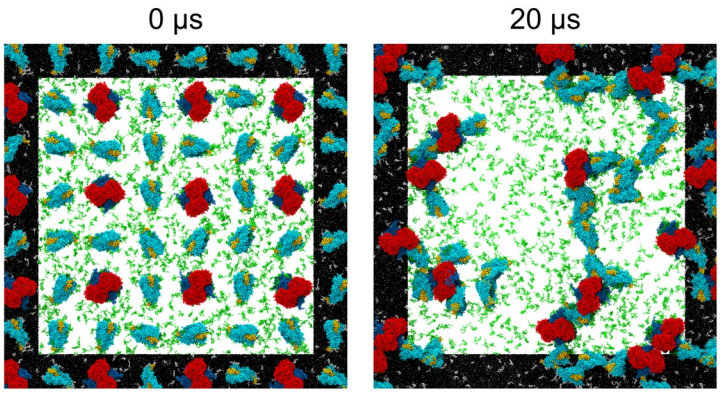



**Figure S2:** Snapshots of the a POPC:POPE:CL (15:15:2) system viewed from the matrix side of the membrane. The color code is as in Figure 1. To ease the visualization solvent molecules were discarded and POPC/POPE lipids were removed from the unit cell so the periodic boundary conditions are emphasized. Similar behavior as observed for the POPC:CL system was observed.

**Figure S3** (external file) | Properties of each CIII/CIV and CIV/CIV interface formed during our self-assembly simulations. The data is presented according to the following order: CIII/CIV interfaces in CL-containing system (+CL); CIII/CIV in CL-depleted system (–CL); CIV/CIV in +CL; CIV/CIV in –CL. For each case, below the code corresponding to the internal numbering given to the complexes (A to I for CIII, 1 to 27 for CIV), two lines of snapshots and six lines of plots are presented. In that order:

- The first line of snapshot is depicting a top view of the overall structure of the interface formed. In the case of CIII/CIV structures, the position of CIV (orange) has been symmetrized around CIII (red) and all structure have been fitted to keep the orientation of the central CIII fixed to ease the comparison between all interfaces. In the case of CIV/CIV structures, the first CIV was used as reference for the fit (red).
- The second line reports the location of the CL binding sites (as described in (1, 2) and reported in Fig. S1) on each complex involved. In the case of CIII/CIV structures, the sites of CIII are indicated in nuances of blue and cyan, and the sites of CIV in nuances of red and pink. In the case of CIV/CIV structures, the sites of the first CIV in nuances of blue and cyan, and the sites of the second in nuances of red and pink.
- The first line of plot reports the time evolution of the interface burial (red curve); an aggregated average (over complexes) is presented in Figure 2E,F. In black is reported the distribution of the burial for each microsecond of simulation.
- The second line of plots reports the time evolution of the intra- and extra-membraneous components of the burial. And aggregated set of plots is presented in Figure S4.
- The third line of plots reports the time evolution of the relative angle between the two complexes involved. In the case of CIII/CIV structures, $\Phi_1$ is relative to CIII, $\Phi_3$ to CIV. In the case of CIV/CIV structures, the first and second CIV complexes are denoted by the subscript.
- The fourth line of plots reports the time evolution of the number of lipids at the interface that will form between the involved complexes during the simulation.
- The fifth line of plots reports the time evolution of the number of common lipids shared by the two complexes at the interface.
- The last line of plots reports the contact between binding sites in the interface formed between the two complexes.

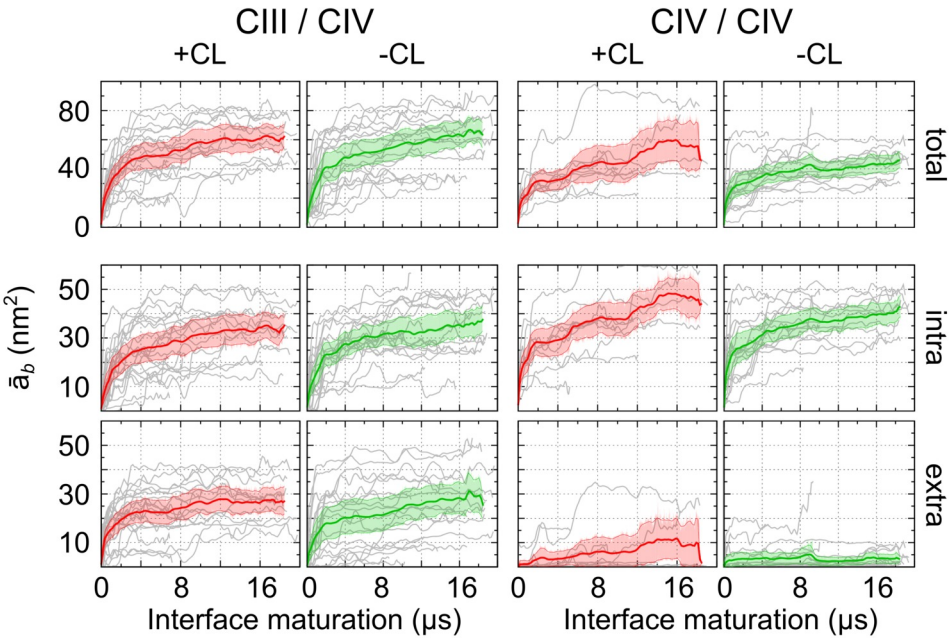



**Figure S4.** Evolution of the protein burial per interface type (*total*) with the contribution from the inside (*intra*) and outside (*extra*) of the bilayer. Each interface is indicated in light gray, the average value in plain colored line and the standard deviation by a shaded area. The time reported corresponds to the interface maturation, $t = 0$ corresponding to the time of formation of the first contact of an interface.

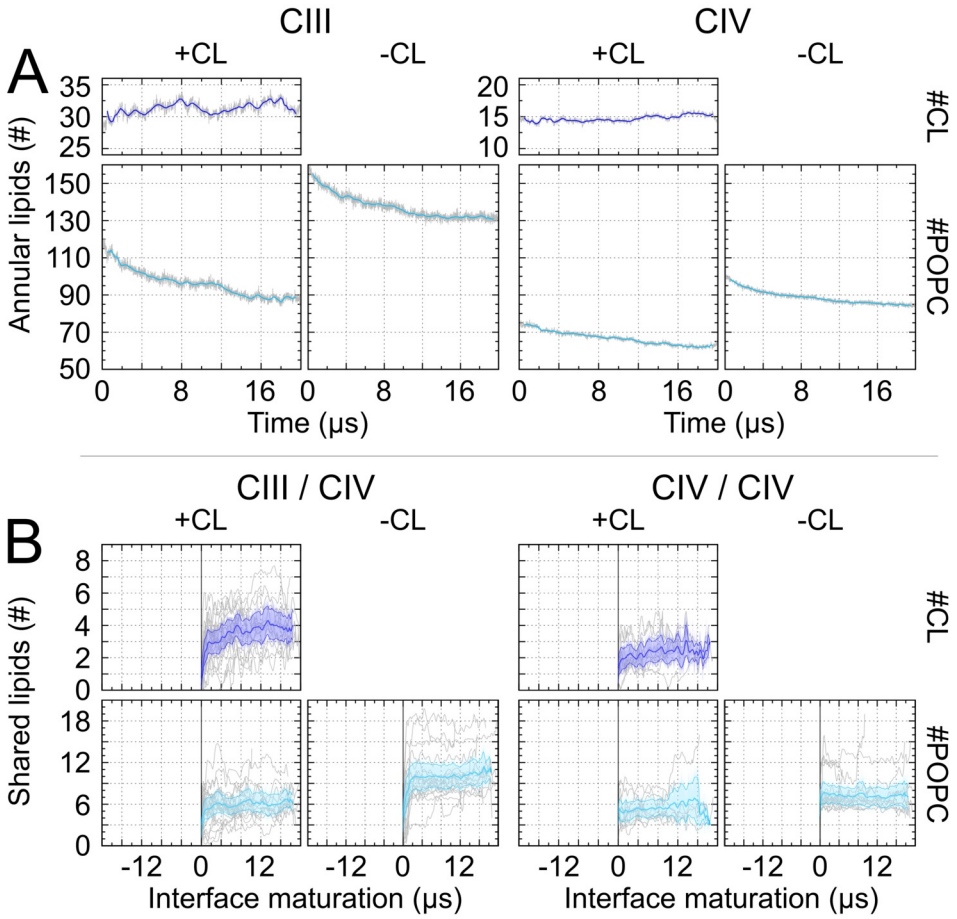



**Figure S5.** Lipid content of protein surroundings and at interfaces between complexes. **(A)** The number of annular lipids (within 1.0 nm of the protein) per complex CIII and CIV is reported in presence (+CL) and absence (–CL) of cardiolipins. **(B)** Number of shared lipids for CIII/CIV and CIV/CIV interfaces. Each interface is indicated in light gray, the average value in plain colored line and the standard deviation by a shaded area. The time reported corresponds to the interface maturation, $t = 0$ corresponding to the time of formation of the first contact of an interface.

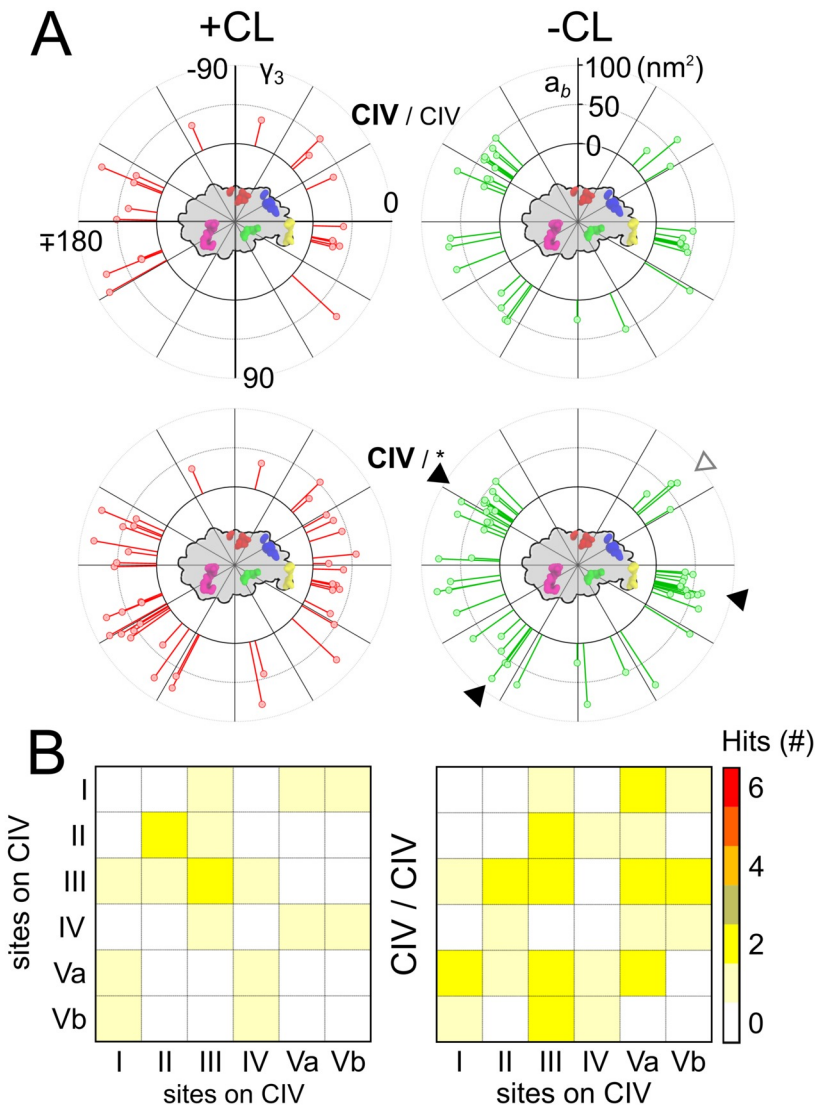



**Figure S6:** Location of supercomplex (SC) contact interfaces on CIV surfaces using the $\gamma_1/\gamma_3$ formalism (see Fig. S10 and Methods for details). **(A)** A stick represents the projection of an interface onto a circle surrounding the protein placed at its center and its length reports the corresponding protein burial, $a_b$. The location of the CIV/CIV interfaces projected onto CIV are shown for both lipid environments (+CL and –CL, respectively). In the bottom subpanels are reported the location of contacts on CIV from CIV/CIII and CIV/CIV interfaces put together. **(B)** Contact map of CL binding sites in CIV/CIV SCs from the simulations +CL and –CL. Sites are assumed to form a contact when distant by < 3.0 nm (see Methods).

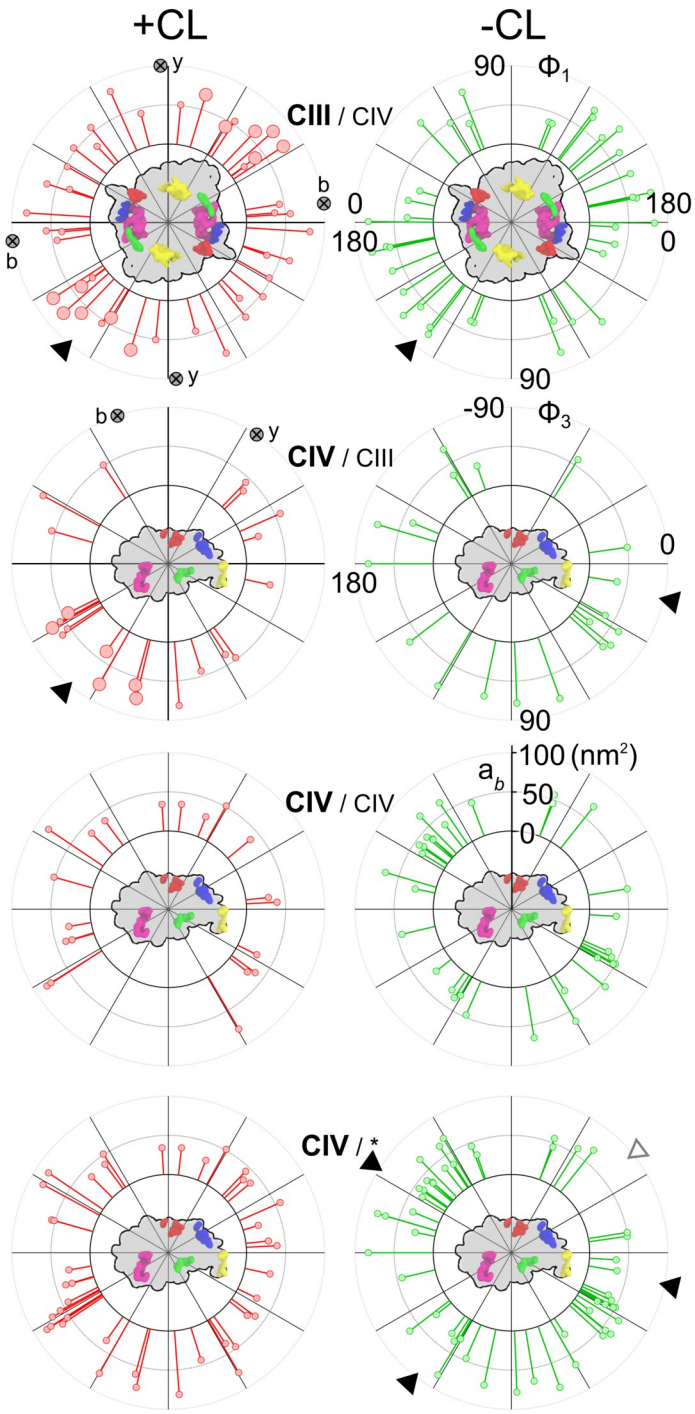



**Figure S7.** Location of supercomplex (SC) contact interfaces on CIII and CIV surfaces using the $\Phi_1/\Phi_3$ formalism (see Fig. S11 and Methods for details). A stick represents the projection of an interface onto a circle surrounding the protein placed at its center and its length reports the corresponding protein burial, $a_b$. The location of the CIII/CIV and CIV/CIV interfaces projected onto CIII and CIV are shown for both lipid environments (+CL and –CL, respectively). In the bottom subpanels are reported the location of contacts on CIV from CIV/CIII and CIV/CIV interfaces put together. The positions of conformations defining the

favored interface found in our simulations (Fig. 2H) are shown with enlarger circles. The dark crosses the position of the bovine and yeast interfaces as defined by the experimental models (37-40). The less defined zones of contacts when compared to $\gamma_1/\gamma_3$ maps (Fig. 2G, S5) reflects the fact that protein contacts mostly do not form in between the centers of masses of the two proteins so $\Phi_1/\Phi_3$ maps are a slightly off.

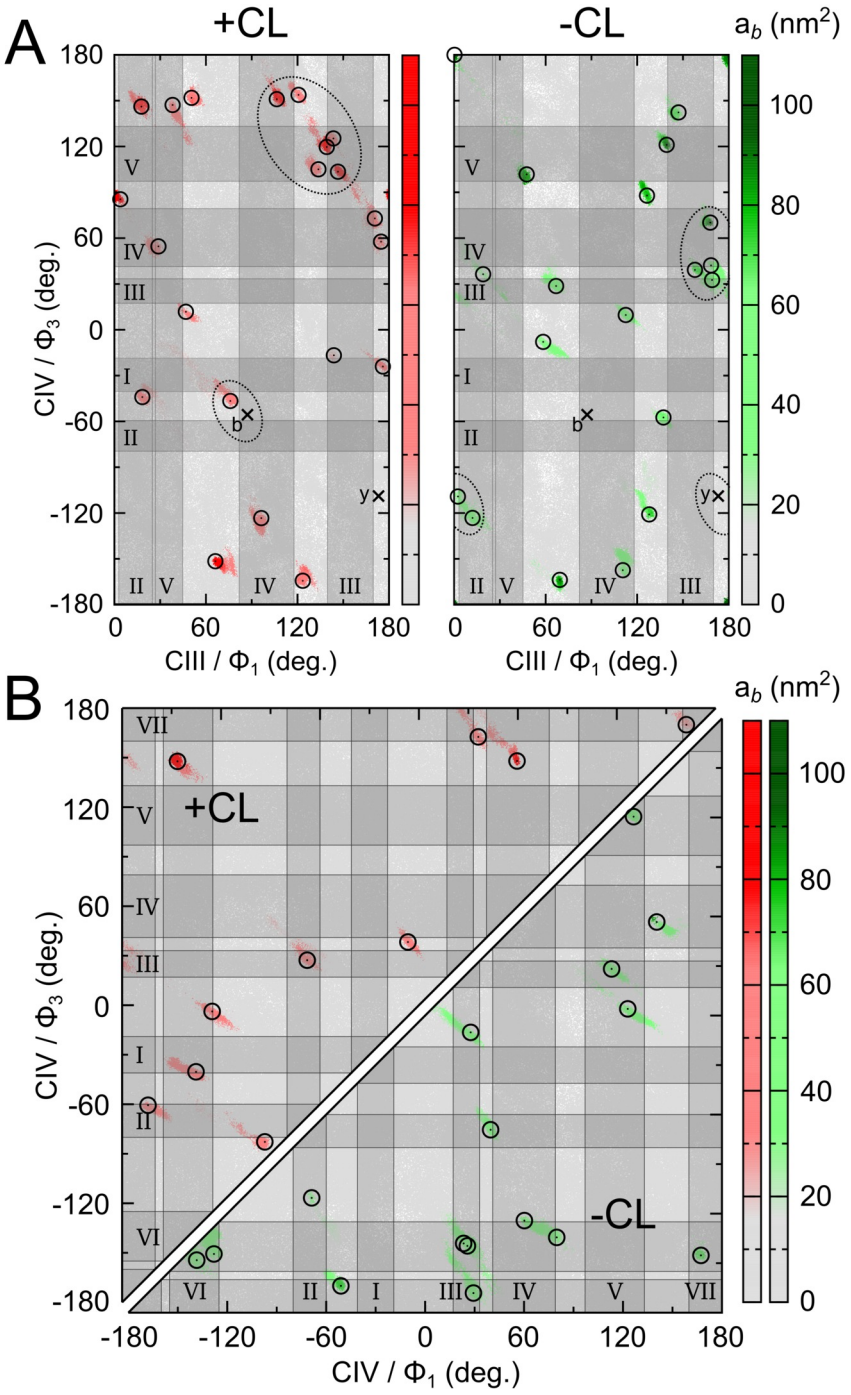



**Figure S8.** Complexes contact map. Projection of the self-assembly trajectories of all associating pairs of CIII and CIV onto the $\Phi_1$ and $\Phi_3$ surface denoting the relative orientation of CIII and CIV, respectively (see Methods). The protein burial associated with the interface, $a_b$, is used as third dimension to emphasize when the complexes are in contact. Both lipid environments are reported for CIII/CIV **(A)** and CIV/CIV **(B)** interactions. The shaded areas correspond to the ranges of angles in which CL binding sites would be facing the interacting complex. Only the sites located on the surface of the proteins and susceptible to interact are thus reported here: *i.e.*, $I_{CIII}$, $VI_{CIII}$, $VII_{CIII}$, and $VI_{CIV}$, $VIIa_{CIV}$, $VIIb_{CIV}$ are not shown. The

black crosses indicate the orientation of the experimental models built from EM data extracted from bovine heart (37, 38) (*b*) and yeast mitochondria (39, 40) (*y*). The experimental orientations are encircled together with interfaces that resemble them in the simulations. Additional circles indicate groups of similar orientations. See text for more details.

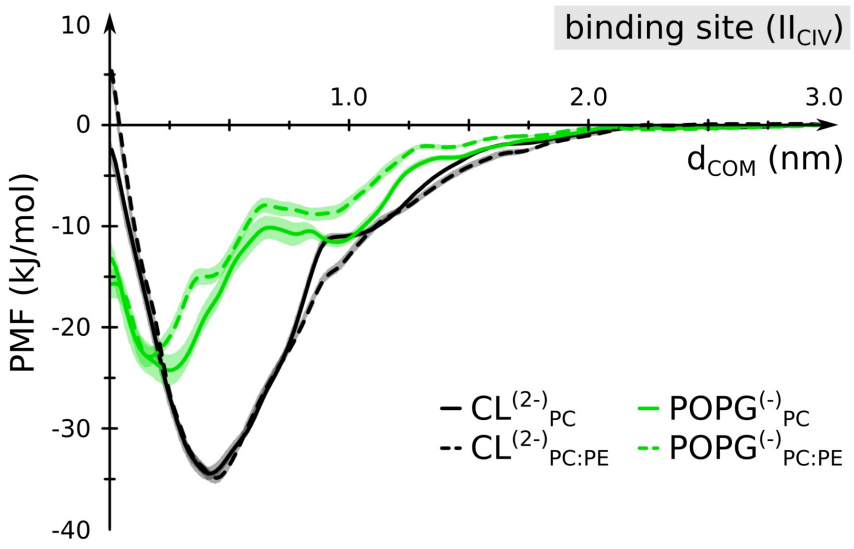



**Figure S9:** Lipid binding affinities. The binding free energy profile (PMF) of CL (black) and PG (green) in a pure POPC (plain lines) and a POPE:POPC mixture at a 1:1 ratio (dashed lines) bilayers are reported. The PMFs were performed on the CL binding site II on CIV. See the Method section for technical details. The data show only little effect of the membrane composition tested on the binding strength of both lipid species.

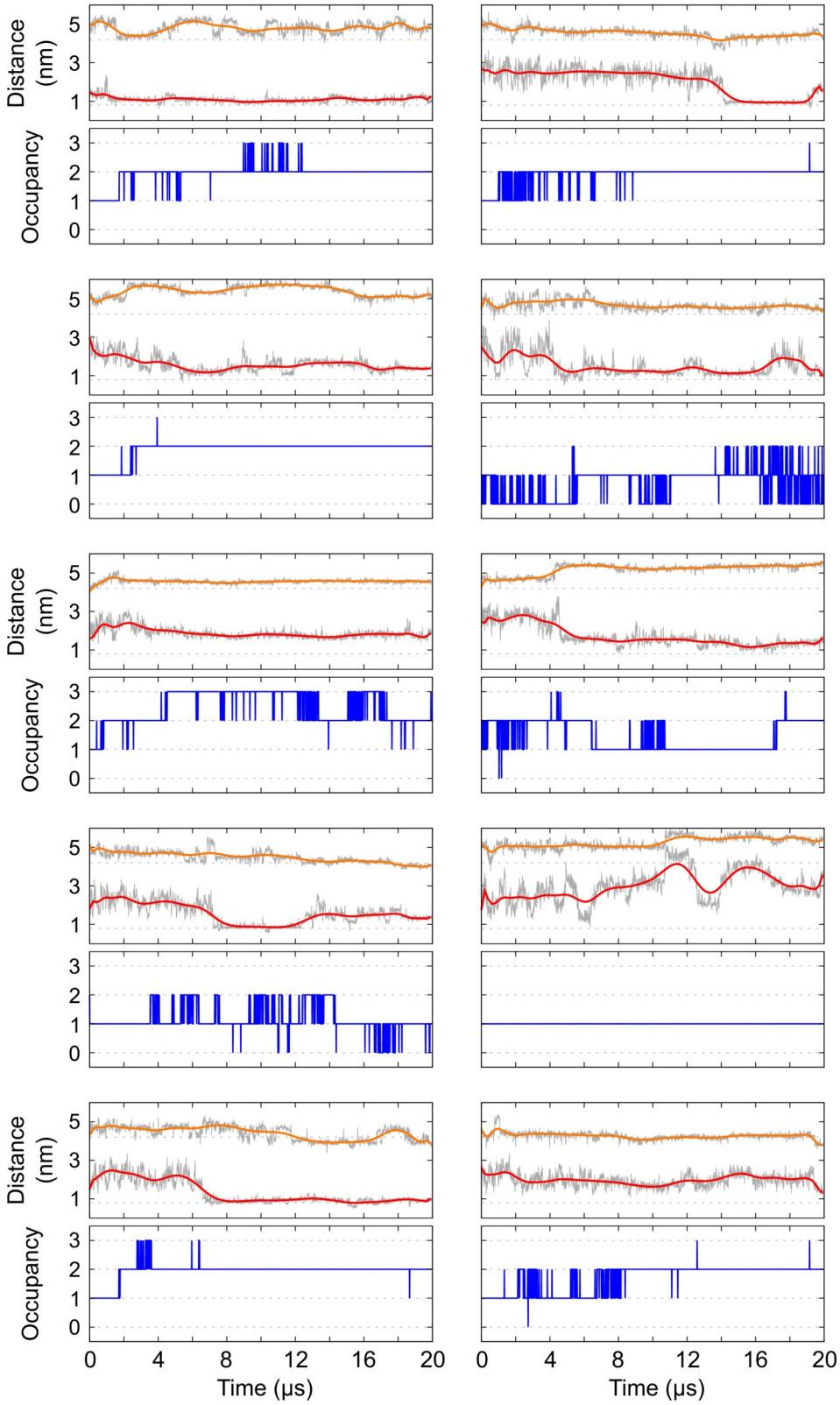



**Figure S10** (first part; caption next page)

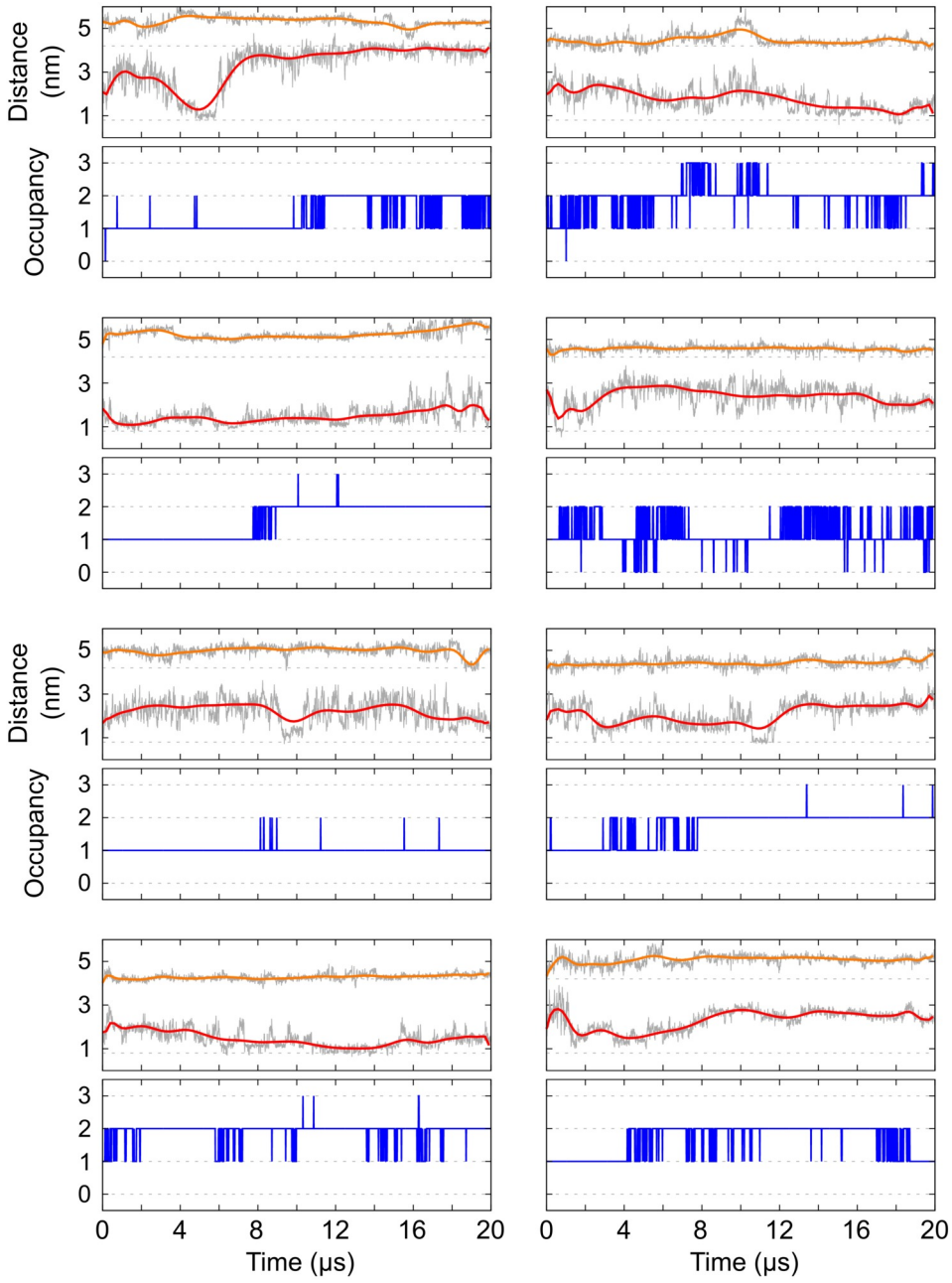



**Figure S10** (second part)**:** Dynamic of the subunit K for each CIII (each monomer) and evolution of the occupancy of the cavity closed by this subunit. For each set of graphs, the top panel depicts a pair of distances representative of the motion of the two subsections of the subunit K (see Fig. 3A in (1)); the bottom panel reports the occupancy of the inner cavity.

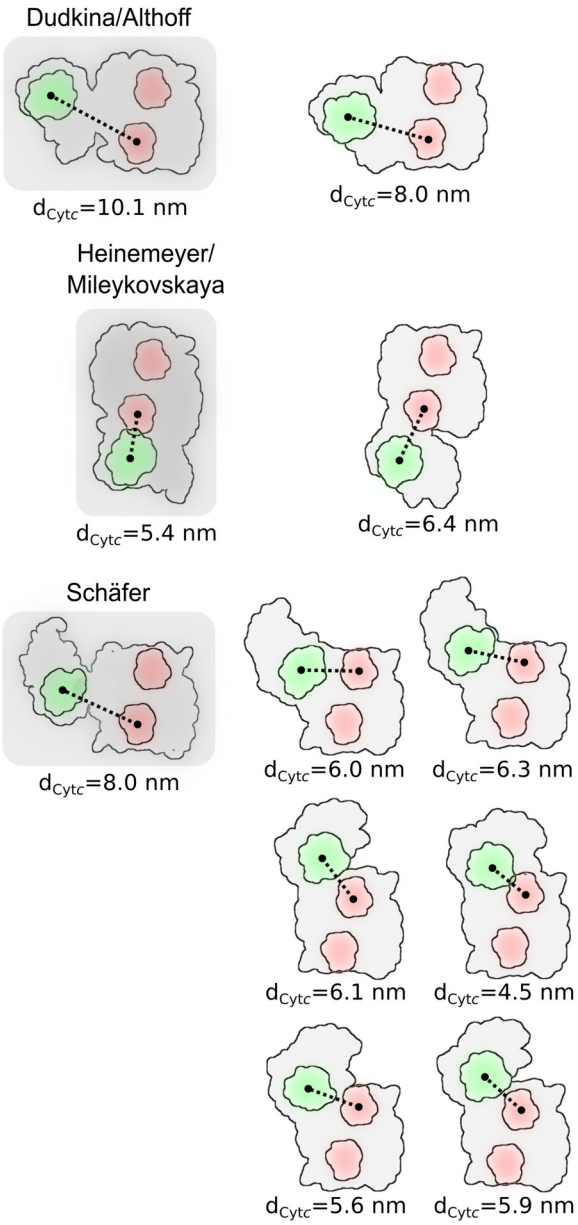



**Figure S11:** Distance between cytochrome *c* binding sites on CIII and CIV. The distances are taken between the cytochrome *c* bound to the CIII (PDB structure 3xc5 (41); red) and the subunit II from the CIV (green) as reported by Mileykovskaya *et al.* (40).

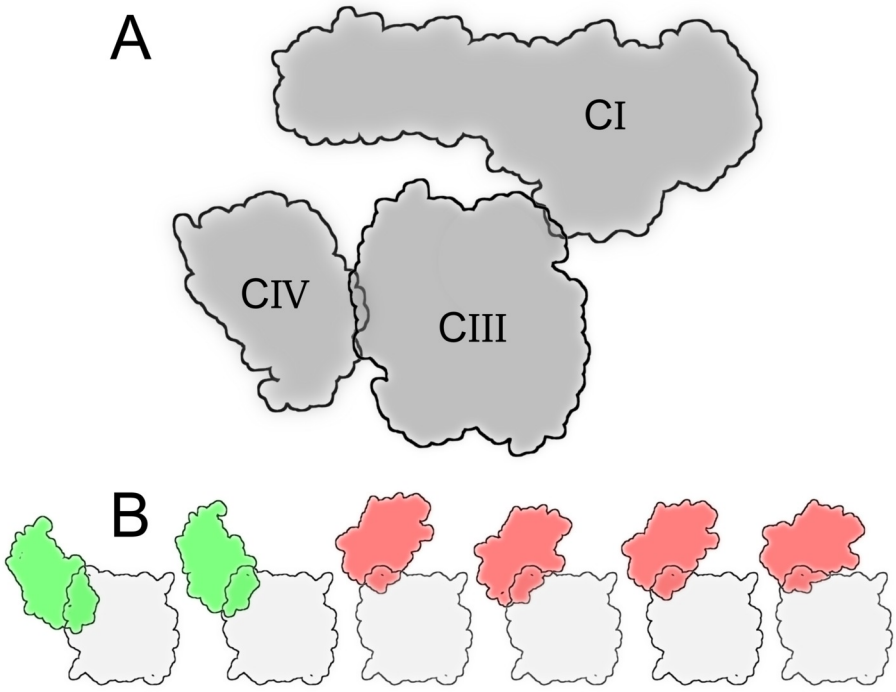



**Figure S12.** Fit of the favored interface found in our simulation (Fig. 2H,K) onto the $CI_1CIII_2CIV_1$ model of bovine heart mitochondria by Althoff *et al.* (37). **(A)** Althoff's model. **(B)** The six conformations contributing to the most populated orientation found in the simulations (Fig. 2H,K). In green are cases where the model would fit into the $CI_1CIII_2CIV_1$. In red are cases where CIV would clash with CI.

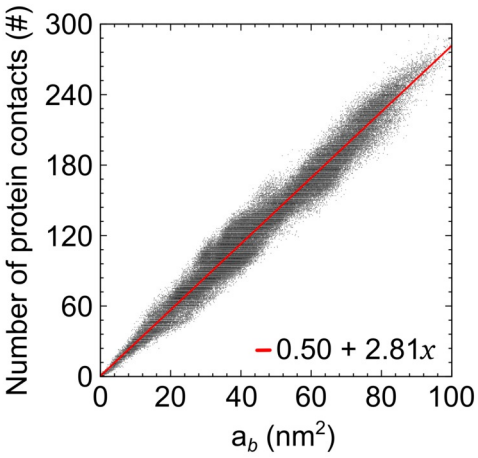



**Figure S13.** Correlation plot between protein buried surface, $a_b$, and number of protein contacts. A conversion factor of 2.81 was derived using all available values, *e.g.* all values of all CIII/CIV and CIV/CIV interfaces formed in the 20 μs simulations in both systems (+CL and –CL).

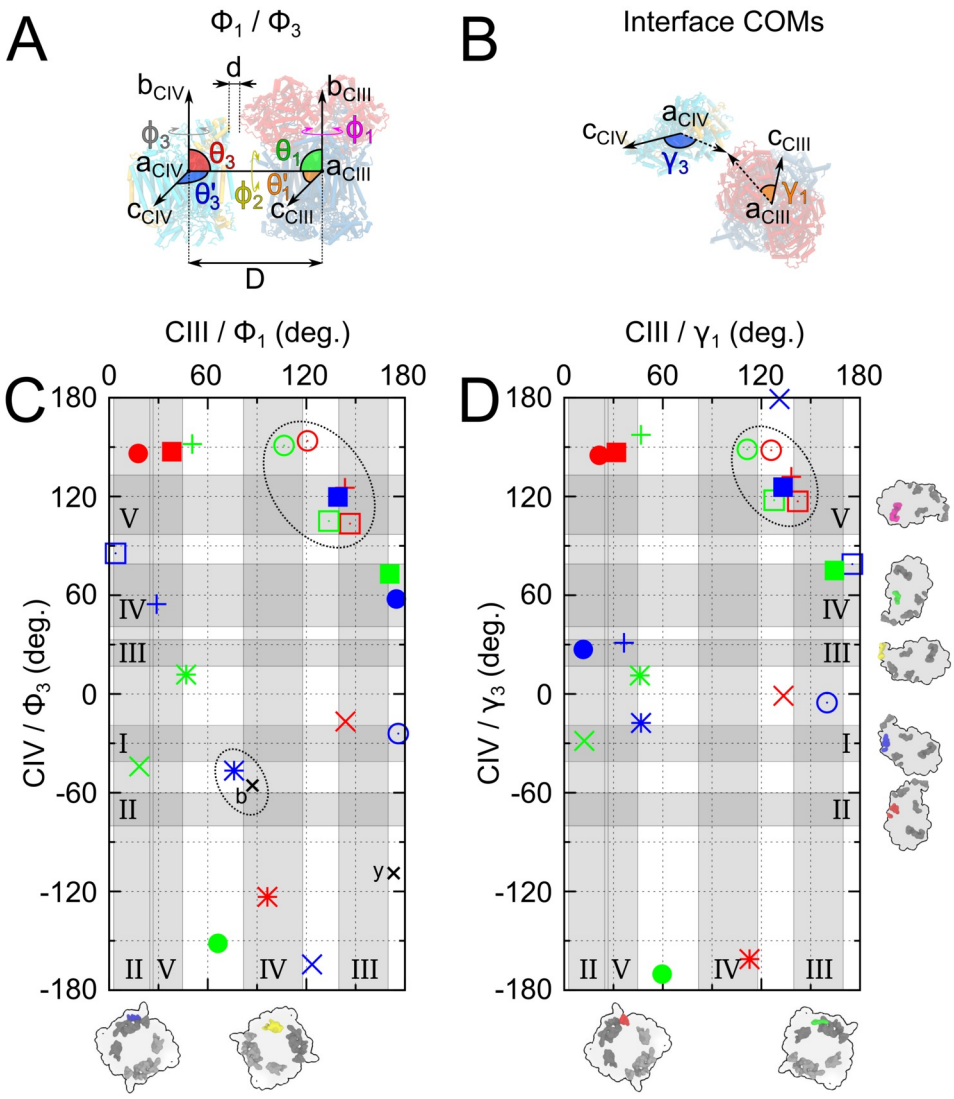



**Figure S14.** Definition and comparison of the angles reporting the relative orientation of the complexes. **(A)** Angles defined by the virtual bond analysis (19), and **(B)** by the center of masses (COMs) of the interfaces formed. **(C)** Projection of all associating pairs of CIII and CIV in the self-assembly trajectory onto the $\Phi_1$ and $\Phi_3$ surface denoting the relative orientation of CIII and CIV, respectively. **(D)** Projection of all associating pairs of CIII and CIV in the self-assembly trajectory onto the angle the $\gamma_1$ and $\gamma_3$ surface. The values are reported only for the system containing CL, and were averaged in both cases over the last microsecond of simulation. Note the similarity of the two definitions. At the bottom and on the right size of the figure are reported the orientation of the complexes, representing the angle ranges for which each site is facing the other complex. These ranges are indicated on the maps by shaded areas. Circled are the dots representing the new interface found in our simulations. Color and shapes are varied for each interface to ease the comparison between the two definitions. On the plot of Panel C is reported the relative orientation ($\Phi_1/\Phi_3$) for the

bovine (*b*) and yeast (*y*) experimental models (37, 39). The difference between the location of interfaces in $\Phi_1/\Phi_3$ when compared to $\gamma_1/\gamma_3$ maps reflects the fact that protein contacts mostly dot not form in between the centers of masses of the two proteins. Thus $\gamma_1/\gamma_3$ are better reporting on the locations of the interfaces on the complexes but do not report on the relative orientations of the complexes, where $\Phi_1/\Phi_3$ are more precise.

## Supplementary Tables

**Table S1.** Lipid molar ratios (CL:POPC and POPC:CL) at different levels of resolution of the protein surfaces and their interfaces. The values were extracted from the analysis in Figure 3 and S4.

**Annular lipids**

| | | CIII | | | | CIV | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | +CL | | –CL | | +CL | | –CL | |
| time (μs) | | 0 | 20 | 0 | 20 | 0 | 20 | 0 | 20 |
| | CL | 28 | 32 | – | – | 15 | 15 | – | – |
| | POPC | 120 | 90 | 160 | 130 | 75 | 63 | 100 | 85 |
| ratio | CL:POPC | 0.23 | 0.36 | | | 0.20 | 0.24 | | |
| | POPC:CL | 4.3 | 2.8 | | | 5.0 | 4.2 | | |

**Interfacial lipids**

| | | CIII/CIV | | | | CIV/CIV | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | +CL | | –CL | | +CL | | –CL | |
| interface status | | before | after | before | after | before | after | before | after |
| | CL | 10 | 9 | – | – | 8 | 6 | – | – |
| | POPC | 35 | 20 | 50 | 30 | 35 | 20 | 47 | 27 |
| ratio | CL:POPC | 0.29 | 0.45 | | | 0.23 | 0.30 | | |
| | POPC:CL | 3.5 | 2.2 | | | 4.4 | 3.3 | | |

**Shared lipids**

| | | CIII/CIV | | | | CIV/CIV | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | +CL | | –CL | | +CL | | –CL | |
| interface status | | before | after | before | after | before | after | before | after |
| | CL | – | 4 | – | – | – | 2.5 | – | – |
| | POPC | – | 6 | – | 10 | – | 6 | – | 7 |
| ratio | CL:POPC | | 0.67 | | | | 0.42 | | |
| | POPC:CL | | 1.5 | | | | 2.4 | | |

**Table S2.** Cardiolipin binding sites' occupation. CL's occupation of the binding sites defined previously (1, 2) for the complexes CIII and CIV are reported while CIII and CIV are in isolation (*isolated*), *i.e.* as reference, averaged over the 9 and 27 copies of the complexes for the last μs of the self-assembly simulation (*in SCs*) and averaged over the bindings sites on either CIII and CIV when involved in a contact with another complex (*shared*). The numbers in parenthesis indicate the number of cases used in the averaging of shared binding sites. Occupancy values in isolation have been corrected to account for the increased CL:POPC ratio from 1:17 and 1:20 for CIII and CIV respectively to 1:15 in the SC self-assembly simulation.

| **CIII** | *isolated* | *in SCs* | *shared* with **CIV** | |
|---|---|---|---|---|
| | | | before | after |
| Ia | 1.08[a] | 0.87 | – | – |
| Ib | 1.50[a,b] | 1.12[b] | – | – |
| II | 1.07[a] | 0.86 | 0.92 | 0.81 (6) |
| III | 0.74 | 0.60 | 0.54 | 0.69 (8) |
| IVa | 0.85 | 0.53 | 0.73 | 0.65 (7) |
| IVb | 0.78 | 0.48 | 0.58 | 0.39 (9) |
| V | 0.23 | 0.16 | 0.18 | 0.10 (6) |
| average[c] | 0.73 | 0.47 | 0.59 | 0.53 |

| **CIV** | *isolated* | *in SCs* | *shared* with **CIII** | | *shared* with **CIV** | |
|---|---|---|---|---|---|---|
| | | | before | after | before | after |
| I | 1.29[a] | 0.78 | 0.80 | 0.69 (3) | 0.80 | 0.81 (2) |
| II | 0.80 | 0.29 | 0.01 | 0.40 (2) | 0.13 | 0.13 (3) |
| III | 0.73 | 0.36 | 0.43 | 0.53 (4) | 0.55 | 0.28 (3) |
| IV | 0.48 | 0.31 | 0.25 | 0.39 (4) | 0.32 | 0.25 (2) |
| Va | 0.97 | 0.38 | 0.57 | 0.33 (10) | 0.34 | 0.31 (2) |
| Vb | 0.87 | 0.37 | 0.25 | 0.34 (6) | 0.74 | 0.31 (2) |
| VI | 0.63 | 0.21 | – | – | 0.15 | 0.37 (3) |
| VIIa | 1.16[a] | 0.66 | – | – | 0.53 | 0.18 (3) |
| VIIb | 0.67 | 0.24 | – | – | 0.33 | 0.26 (3) |
| average[d] | 0.86 | 0.42 | 0.39 | 0.45 | 0.48 | 0.35 |

[a] indicates an occupancy exceeding 1 due to renormalization. [b] indicates an occupancy exceeding 1 due to motion of CL within site $I_{CIII}$. [c] sites Ia and Ib were not included. [d] sites VI, VIIa,b were not included.

**Table S3.** Effect of modifying head group charges and acyl chain on the binding strength of CL and POPG. Energies are reported in $kJ.mol^{-1}$. $CL^{2-}$ is considered the physiological species (although one could argue on its exact protonation state in physiological conditions) with two tails per glycerol group, head group, (2t|2t), is used as reference to determine the relative position of the other lipid species. All energies reported in here are thus differences relative to this reference; a null value meaning an equivalent binding strength to the binding site, a positive value meaning a less favorable interaction. The relative binding strength is assumed to be reflected by the position of minimum of the PMF shown in Fig. 4 in the main manuscript. Dashes indicate undetermined values (conditions not simulated). mlCL and dlCL denote monolyso (one lipid tail missing) and dilyso (two lipid tails are missing) CL, respectively. In the case of dlCL we tested the removal of the tails on the same glycerol (2t|0t) or one on each (1t|1t). The second case turned out to be really hard to sample, the bulky head of CL not being properly supported by the lacking tails when single- and double-charged (when attraction to the binding site is the highest). Note that it is relatively difficult to associate accurate errors to the values reported.

| | **CL** (4 tails) | **mlCL** | **dlCL** | | **POPG** |
|---|---|---|---|---|---|
| | 2t\|2t | 2t\|1t | 2t\|0t | 1t\|1t | 2t |
| -2 *e* | *reference* | 0 | +4 | +12 | +15 |
| -1 *e* | +8 | – | +10 | +12 | +10 |
| 0 *e* | +8 | – | +22 | +22 | – |

## Supplementary References


1. 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 $bc_1$. *J Am Chem Soc* 135:3112–3120.

2. Arnarez C, Marrink S-J, Periole X (2013) Identification of cardiolipin binding sites on cytochrome *c* oxidase at the entrance of proton channels. *Sci Rep* 3:1263.

3. Gao X, et al. (2002) The crystal structure of mitochondrial cytochrome $bc_1$ in complex with famoxadone: the role of aromatic-aromatic interaction in inhibition. *Biochemistry* 41(39):11692–11702.

4. Esser L, et al. (2004) Crystallographic studies of quinol oxidation site inhibitors: a modified classification of inhibitors for the cytochrome $bc_1$ complex. *J Mol Biol* 341(1):281–302.

5. Huang L-S, Cobessi D, Tung EY, Berry EA (2005) Binding of the respiratory chain inhibitor antimycin to the mitochondrial $bc_1$ complex: a new crystal structure reveals an altered intramolecular hydrogen-bonding pattern. *J Mol Biol* 351(3):573–597.

6. Tsukihara T, et al. (1996) The whole structure of the 13-subunit oxidized cytochrome *c* oxidase at 2.8 A. *Science* 272(5265):1136–1144.

7. Daum G (1985) Lipids of mitochondria. *Biochim Biophys Acta* 822(1):1–42.

8. van Meer G, Voelker DR, Feigenson GW (2008) Membrane lipids: where they are and how they behave. *Nat Rev Mol Cell Biol* 9(2):112–124.

9. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* 4(3):435–447.

10. Marrink S-J, Risselada HJ, Yefimov S, Tieleman DP, de Vries AH (2007) The MARTINI force field: coarse grained model for biomolecular simulations. *J Phys Chem B* 111:7812–7824.

11. Monticelli L, et al. (2008) The MARTINI coarse-grained force field: extension to proteins. *J Chem Theory Comput* 4(5):819–834.

12. Periole X, Cavalli M, Marrink S-J, Ceruso MA (2009) Combining an elastic network with a coarse-grained molecular force field: structure, dynamics, and intermolecular recognition. *J Chem Theory Comput* 5:2531–2543.

13. Marrink S-J, Tieleman DP (2013) Perspective on the Martini model. *Chem Soc Rev* 42(16):6801–6822.

14. Periole X, Huber T, Marrink S-J, Sakmar TP (2007) G protein-coupled receptors self-assemble in dynamics simulations of model bilayers. *J Am Chem Soc* 129:10126–

10132.

15. Periole X, Knepp AM, Sakmar TP, Marrink S-J, Huber T (2012) Structural determinants of the supramolecular organization of G protein-coupled receptors in bilayers. *J Am Chem Soc* 134(26):10959–10965.

16. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. *J Chem Phys* 81(8):3684–3690.

17. Dahlberg M (2007) Polymorphic phase behavior of cardiolipin derivatives studied by coarse-grained molecular dynamics. *J Phys Chem B* 111(25):7194–7200.

18. Eisenhaber F, Lijnzaad P, Argos P, Sander C, Scharf M (1995) The double cubic lattice method: Efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *J Comput Chem* 16(3):273–284.

19. Boresch S, Tettinger F, Leitgeb M, Karplus M (2003) Absolute binding free energies: A quantitative approach for their calculation. *J Phys Chem B* 107(35):9535–9551.

20. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J Comput Chem* 13(8):1011–1021.

21. Johnston JM, Wang H, Provasi D, Filizola M (2012) Assessing the relative stability of dimer interfaces in G protein-coupled receptors. *PLOS Comput Biol* 8(8):e1002649.

22. Schmidt MR, Stansfeld PJ, Tucker SJ, Sansom MSP (2013) Simulation-based prediction of phosphatidylinositol 4,5-bisphosphate binding to an ion channel. *Biochemistry* 52(2):279–281.

23. Weingarth M, et al. (2013) Structural determinants of specific lipid binding to potassium channels. *J Am Chem Soc* 135(10):3983–3988.

24. Petrov D, Zagrovic B (2014) Are current atomistic force fields accurate enough to study proteins in crowded environments? *PLOS Comput Biol* 10(5):e1003638.

25. Stark AC, Andrews CT, Elcock AH (2013) Toward optimized potential functions for protein–protein interactions in aqueous solutions: osmotic second virial coefficient calculations using the MARTINI coarse-grained force field. *J Chem Theory Comput* 9:4176–4185.

26. Sengupta D, Marrink S-J (2010) Lipid-mediated interactions tune the association of glycophorin A helix and its disruptive mutants in membranes. *Phys Chem Chem Phys* 12(40):12987–12996.

27. Schäfer LV, et al. (2011) Lipid packing drives the segregation of transmembrane helices into disordered lipid domains in model membranes. *Proc Natl Acad Sci USA* 108(4):1343–1348.

28. Hénin J, Pohorille A, Chipot C (2005) Insights into the recognition and association of transmembrane alpha-helices. The free energy of alpha-helix dimerization in glycophorin A. *J Am Chem Soc* 127(23):8478–8484.

29. Yano Y, Matsuzaki K (2006) Measurement of thermodynamic parameters for hydrophobic mismatch 1: self-association of a transmembrane helix. *Biochemistry* 45(10):3370–3378.

30. Acín-Pérez R, Fernández-Silva P, Peleato ML, Pérez-Martos A, Enriquez JA (2008) Respiratory active mitochondrial supercomplexes. *Mol Cell* 32:529–539.

31. Baradaran R, Berrisford JM, Minhas GS, Sazanov LA (2013) Crystal structure of the entire respiratory complex I. *Nature* 494(7438):443–448.

32. Vukotic M, et al. (2012) Rcf1 mediates cytochrome oxidase assembly and respirasome formation, revealing heterogeneity of the enzyme complex. *Cell Metab* 15(3):336–347.

33. Chen Y-C, et al. (2012) Identification of a protein mediating respiratory supercomplex stability. *Cell Metab* 15(3):348–360.

34. Strogolova V, Furness A, Robb-McGrath M, Garlich J, Stuart RA (2012) Rcf1 and Rcf2, members of the hypoxia-induced gene 1 protein family, are critical components of the mitochondrial cytochrome $bc_1$-cytochrome *c* oxidase supercomplex. *Molecular and Cellular Biology* 32(8):1363–1373.

35. Böttinger L, et al. (2012) Phosphatidylethanolamine and cardiolipin differentially affect the stability of mitochondrial respiratory chain supercomplexes. *J Mol Biol* 423(5):677–686.

36. Kadenbach B, Jarausch J, Hartmann R, Merle P (1983) Separation of mammalian cytochrome c oxidase into 13 polypeptides by a sodium dodecyl sulfate-gel electrophoretic procedure. *Anal Biochem* 129(2):517–521.

37. Althoff T, Mills DJ, Popot J-L, Kühlbrandt W (2011) Arrangement of electron transport chain components in bovine mitochondrial supercomplex $I_1III_2IV_1$. *EMBO J* 30(22):4652–4664.

38. Dudkina NV, Kudryashev M, Stahlberg H, Boekema EJ (2011) Interaction of complexes I, III, and IV within the bovine respirasome by single particle cryoelectron tomography. *Proc Natl Acad Sci USA* 108(37):15196–15200.

39. Heinemeyer J, Braun HP, Boekema EJ, Kouřil R (2007) A structural model of the cytochrome *c* reductase/oxidase supercomplex from yeast mitochondria. *J Biol Chem* 282(16):12240–12248.

40. Mileykovskaya E, et al. (2012) Arrangement of the respiratory chain complexes in saccharomyces cerevisiae supercomplex $III_2IV_2$ revealed by single particle cryo-electron microscopy. *J Biol Chem* 287(27):23095–23103.

41. Solmaz SRN, Hunte C (2008) Structure of complex III with bound cytochrome *c* in reduced state and definition of a minimal core interface for electron transfer. *J Biol Chem* 283(25):17542–17549.