Supplemental Information

Structures of a complete human V-ATPase reveal mechanisms of its assembly

Longfei Wang, Di Wu, Carol V. Robinson, Hao Wu, Tian-Min Fu

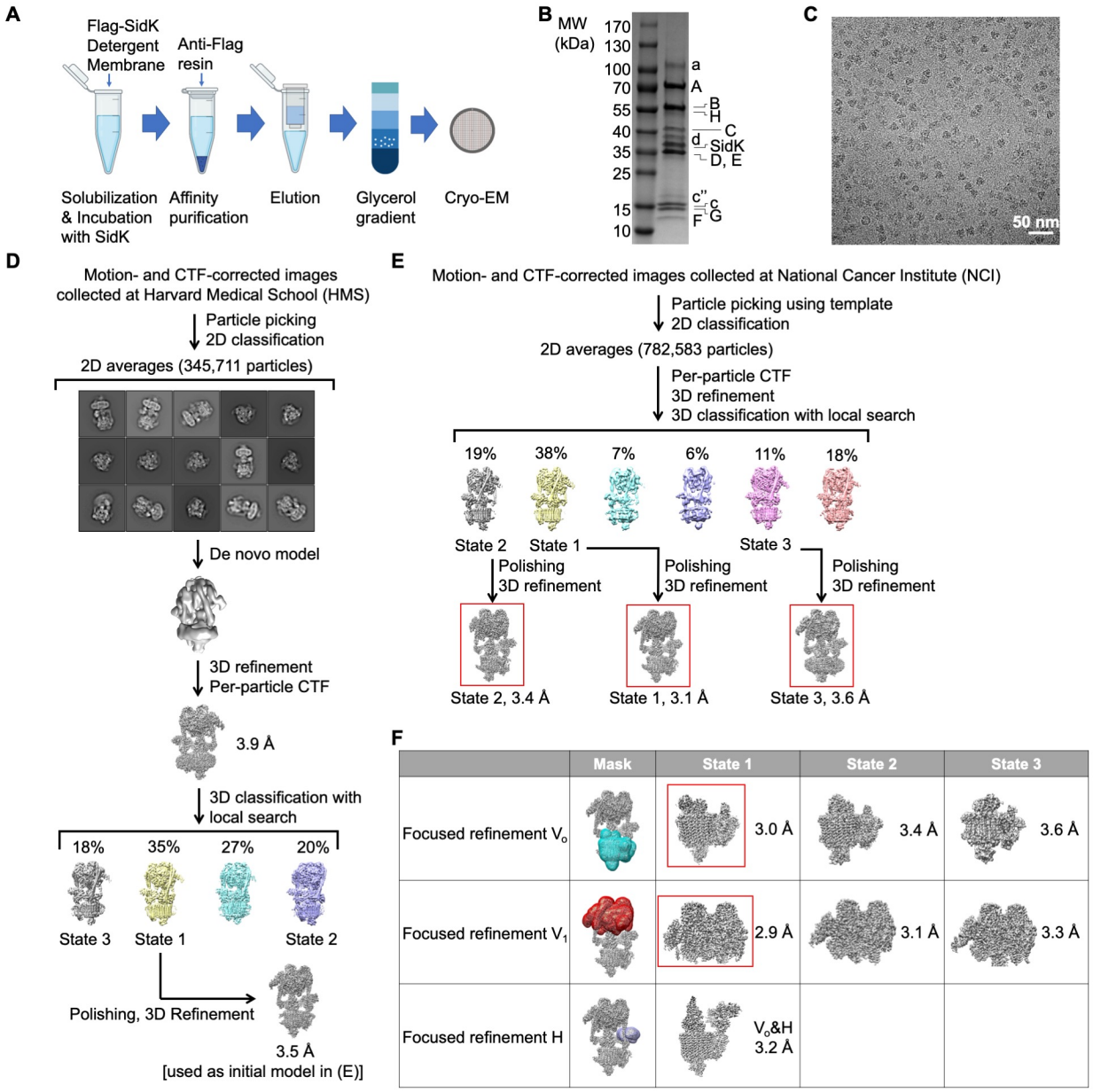

Figure S1. Purification and 3D Reconstruction of Human V-ATPase, Related to Figure 1

(A) Diagram of the purification procedure.

(B) SDS-PAGE showing purified human V-ATPase sample.

(C) A representative cryo-EM image of human V-ATPase.

(D, E) Workflow of 3D reconstruction of dataset I (D) and dataset II (E). The three maps boxed in red rectangles are the final maps of the three states.

(F) Focused refinement and improvement of resolution. Final V_1 and V_o maps are boxed in red rectangles.

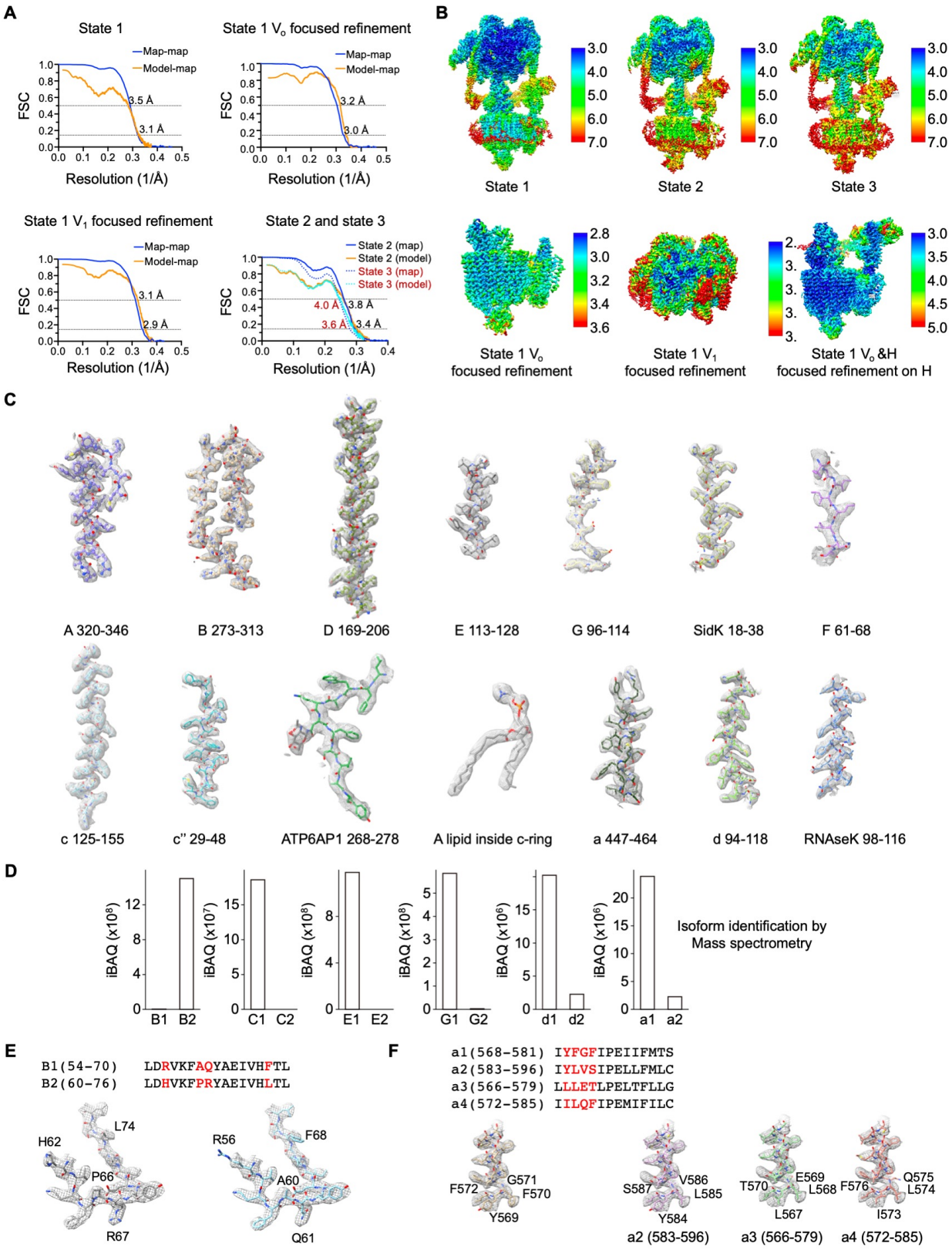

Figure S2. Quality of Reconstructed Human V-ATPase Maps, Related to Figure 1

(A) Fourier shell correlation (FSC) curves of 3D reconstructed human V-ATPase maps.

(B) Local resolutions of the reconstructions correlating with final maps in Figure S1E and S1F. Resolutions are color-coded by scale bars.
(C) Representative segments of the state 1 cryo-EM density map in each subunit fitted with the final atomic model.
(D) The isoforms of V-ATPase subunits identified by mass spectrometry.
(E) Cryo-EM density map of a segment of subunit B that is fitted well with isoform B1 but poorly with isoform B2.
(F) Cryo-EM density map of a segment of subunit a that is fitted well with isoform a1, but less well with isoforms a2, a3, and a4.

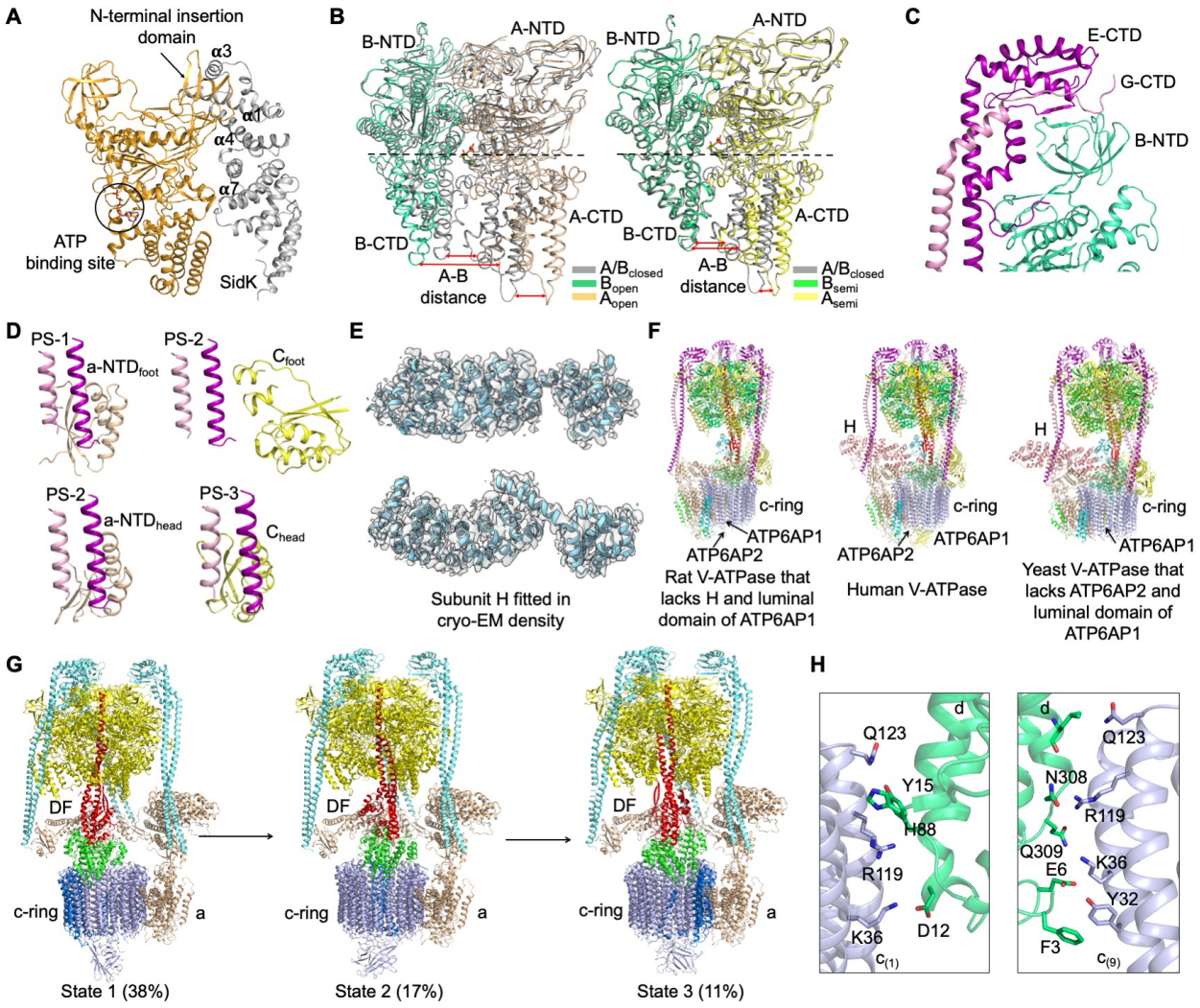

Figure S3. Subunits in the V_1 and V_o Complex, Related to Figure 2

(A) Close-up view of the interface between subunit A (orange) and SidK (gray).

(B) Structure comparison of AB heterodimers in open (wheat/green), closed (gray), and semi-open (yellow/lime) conformations.

(C) Interaction interfaces between the subunit E (purple)-G (pink) complex and subunit B (green).

(D) N-terminal domains of PS-1, PS-2, and PS-3 interact with different regions of subunit a (wheat) and subunit c (yellow), as illustrated.

(E) Ribbon diagram of subunit H fitted into the cryo-EM map at a contour level of 2.0 σ.

(F) Comparison between human, rat, and yeast V-ATPase structures with highlighted differences.

(G) Three rotational states of human V-ATPase structures and their abundance among all the cryo-EM particles.

(H) Detailed quasi-equivalent interactions of subunit d with $c_{(1)}$ and $c_{(9)}$, highlighting the importance of R119 and K36 of c for both interactions.

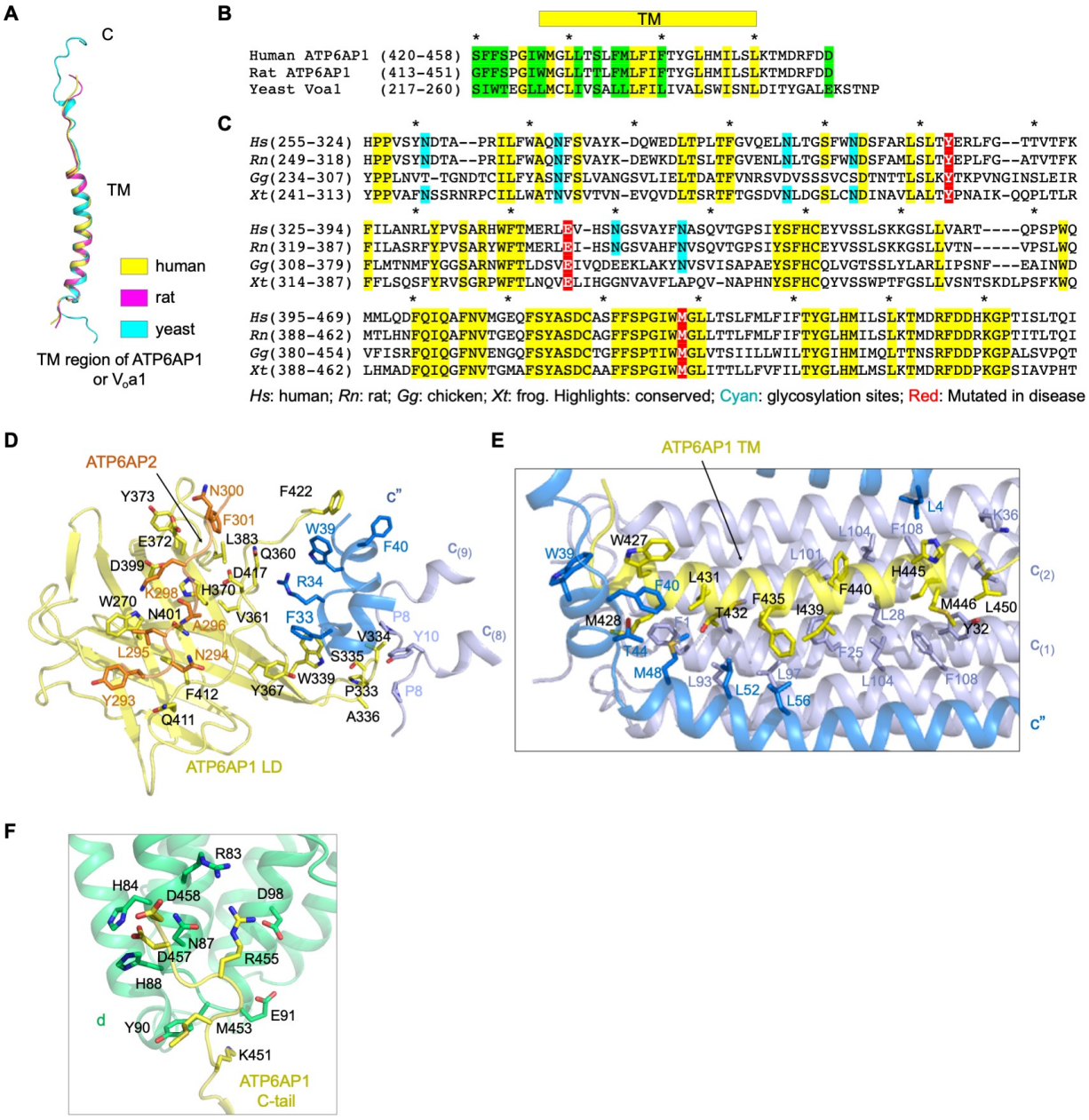

Figure S4. ATP6AP1 and Its Interaction Network, Related to Figure 4

(A) Overlaid structures of the transmembrane region of human ATP6AP1 (yellow), rat ATP6AP1 (magenta), and yeast V_oa1 (cyan).

(B) Sequence alignment of the transmembrane region of human ATP6AP1, rat ATP6AP1, and yeast V_oa1. Identical and similar residues are highlighted in yellow and green, respectively.

(C) Sequence alignment of luminal domain of human ATP6AP1 with LAMP-1. Identical residues are highlighted in yellow, glycosylation sites in cyan, and missense disease mutation sites in red.

(D) Detailed interactions of ATP6AP1 LD (yellow) with ATP6AP2 (orange), c" (marine), $c_{(8)}$ (blue), and $c_{(9)}$ (blue).

(E) Detailed interactions of ATP6AP1 transmembrane region (yellow) with c" (marine), $c_{(1)}$ (blue), and $c_{(2)}$ (blue).

(F) Detailed interactions between ATP6AP1 cytosolic tail (yellow) and subunit d (green).

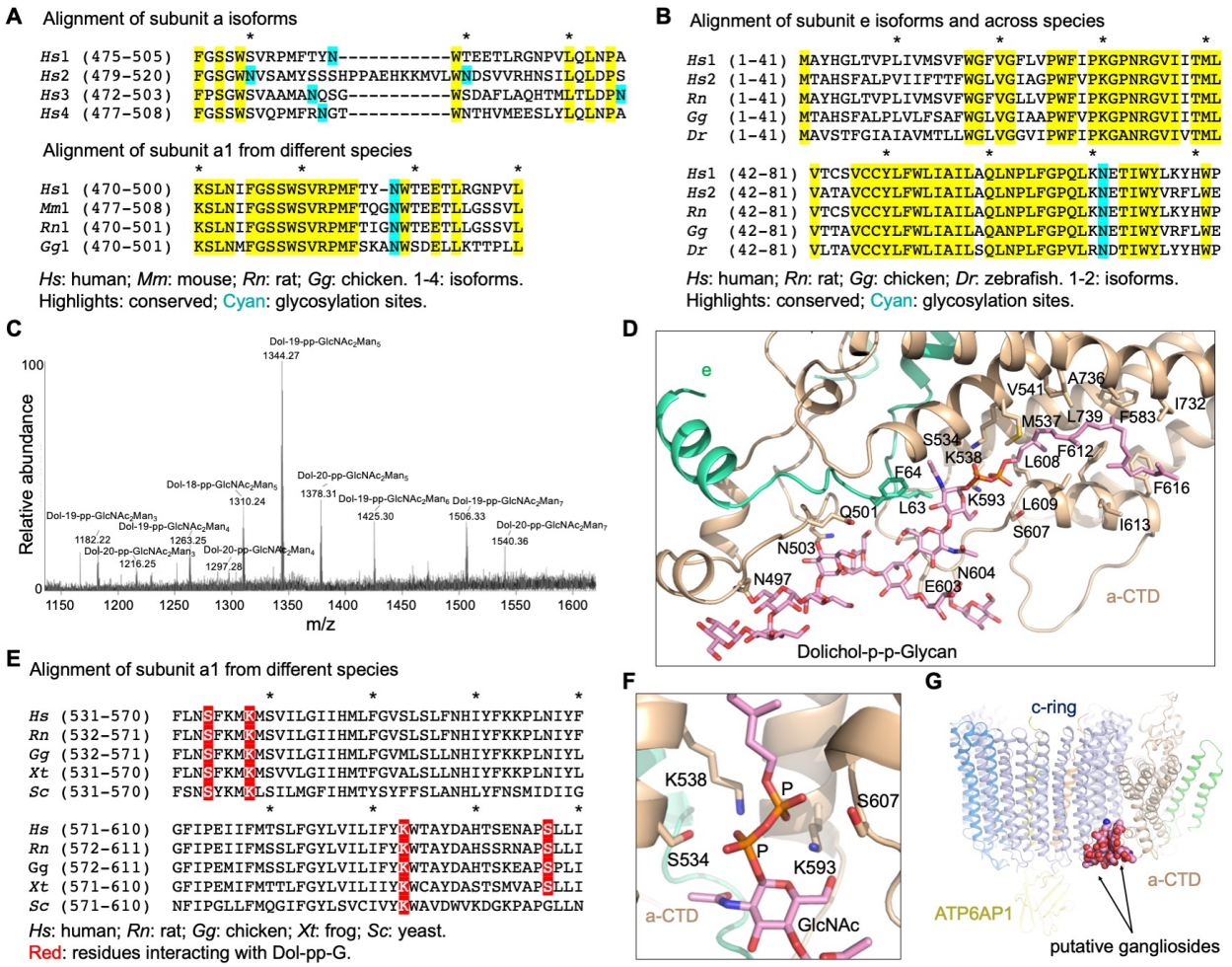

Figure S5. N-linked Glycosylation and Glycolipids in the V_o Complex, Related to Figure 5

(A) Sequence alignment of the N-linked glycosylation regions of human (Hs) a1, a2, a3, and a4 (top), and human, mouse (Mm), rat (Rn), and chicken (Gg) a1. Identical residues and N-glycosylated residues are highlighted in yellow and cyan, respectively.

(B) Sequence alignment of human e1 (Hs1), human e2 (Hs2), rat e1 (Rn), chicken e1 (Gg), and zebrafish e1 (Dr). Identical residues and N-glycosylated residues are highlighted in yellow and cyan, respectively.

(C) Identification of Dolichol-p-p-glycan (Dol-pp-G) by mass spectrometry.

(D) Detailed interactions of subunits a (wheat) and e (green) with the Dolichol-p-p-glycan (pink).

(E) Sequence alignment of Dol-pp-G coordinating segments of human (Hs), rat (Rn), chicken (Gg), frog (Xt), and yeast (Sc) subunit a. Residues that are critical for coordinating the glycolipid are highlighted in red.

(F) Detailed interactions between a1 and the di-phosphate of Dol-pp-G.

(G) Two putative ganglioside molecules (pink) at the interface between c-ring (blue) and subunit a (wheat).

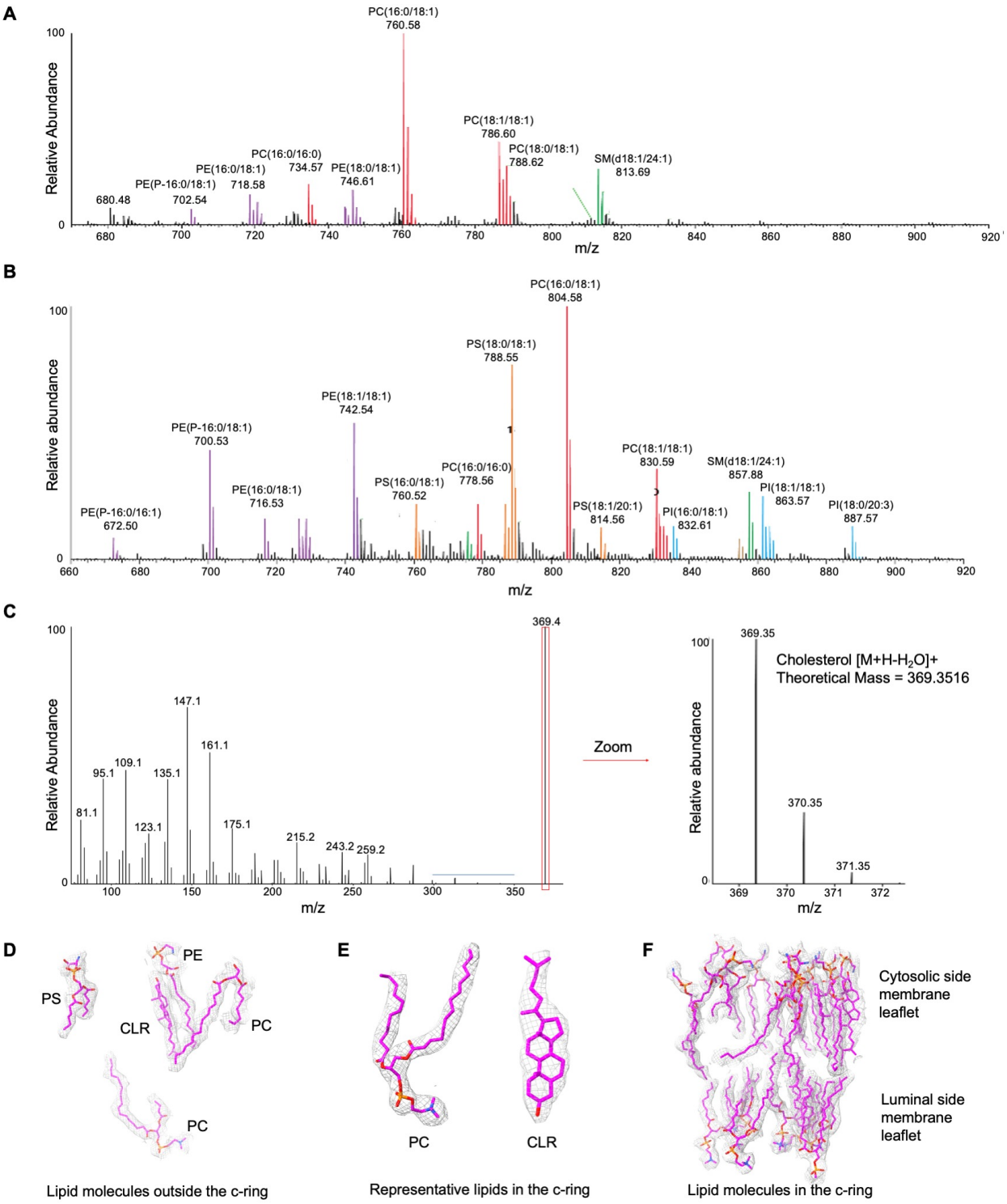

Figure S6. Lipids in the V_o Complex, Related to Figure 6

(A) Identification of phospholipids by mass spectrometry at positive mode.

(B) Identification of phospholipids by mass spectrometry at negative mode.

(C) Identification of cholesterol (CLR) by mass spectrometry.

(D) Lipids outside c-ring superimposed with the cryo-EM density map (2.0 σ). Five lipid molecules were identified in the luminal and cytosolic leaflets.

(E) Examples of modeled PC and CLR in cryo-EM density maps (2.0 $\boldsymbol{\sigma}$).
(F) Lipids inside c-ring superimposed with the cryo-EM density map (2.0 $\boldsymbol{\sigma}$). 24 lipid molecules were identified in the luminal and cytosolic leaflets.

Table S1. Cryo-EM data collection, refinement and validation statistic, Related to Figure 1

Datasets	I (HMS)	II (NCI)				
Structures		State 1	State 1 Focused V_o	State 1 Focused V_1	State 2	State 3
Cryo-EM maps						
EMDB ID		21847	21844	21845	21848	21849
Magnification	81,000	81,000				
Voltage (keV)	300	300				
Electron exposure (e/Å^2)	63.5	50.1				
Defocus range (μm)	-0.8 to -2.2	-0.8 to -2.2				
Pixel size (Å)	1.06	1.08				
Symmetry imposed	C1	C1				
Initial particles (no.)	345,711	782,583				
Final particles (no.)	130,533	301,510			150,243	90,369
Map resolution (Å) FSC threshold	3.5 0.143	3.1 0.143	3.0 0.143	2.9 0.143	3.4 0.143	3.6 0.143
Models						
PDB ID		6WM2	6WLW	6WLZ	6WM3	6WM4
Initial model used (PDB code)		3J9T, 6O7U 6VQ6	6WM2	6WM2	6WM2	6WM2
Model resolution (Å) FSC threshold		3.5 0.5	3.2 0.5	3.1 0.5	3.8 0.5	4.0 0.5
Map sharpening *B* factor (Å^2)		108.77	68.99	71.82	118.82	143.92
Model composition Non-hydrogen atoms Protein residues Ligands		72376 9179 17	21361 2747 16	42233 5335 1	72042 9179 1	72042 9179 1
R.m.s. deviations Bond lengths (Å) Bond angles (°)		0.007 0.755	0.006 0.675	0.007 0.732	0.007 0.820	0.005 0.883
Validation MolProbity score Clashscore Poor rotamers (%)		3.12 18.63 15.41	2.44 8.71 9.34	3.16 16.96 10.57	3.06 18.88 14.50	2.92 18.40 10.08
Ramachandran plot Favored (%) Allowed (%) Disallowed (%)		0.25 6.79 92.96	0 3.58 96.42	0.42 6.62 92.96	0.23 5.86 93.91	0.16 5.67 94.17