Supplemental Information: Tables S1 and S2 and Figures S1 and S2

Cheng J, Goldstein R, Gershenson A, Stec B, Roberts MF (2013)
"The cation- π box is a specific phosphatidylcholine membrane targeting motif,"

The following table compares all proteins in the PDB as of August 2012 that contain choline, phosphocholine, phosphatidylcholine lipid, or similar choline containing ligands. Included in the chart is the identity of the protein, its organism of origin, all PDBs associated with its identity and origin, as well as its state of membrane interaction and the function of the protein. Ligands are abbreviated as per PDB codes, and a list of the abbreviations follows. The binding mode of the choline ligand is also noted, with attention focused on the nature of the coordination of the quaternary amine of the choline or choline portion of the ligand. Proteins are ordered in a general sense by function, i.e. proteins that bind choline specifically, lipid transfer proteins, esterases, etc. Highlighted in bold are any proteins that bind the quaternary amine of choline with a π-cation box type arrangement: that is, the cationic quaternary amine is held between the π systems several (usually 2-4) aromatic residues.

Table SI. Proteins in the PDB that have been shown to bind choline or choline-containing ligands.

Protein (organism)	PDBID	Membrane Interaction	Function	Ligand	Binding Mode
choline binding protein F (*Streptococcus pneumoniae*)	2V04, 2V05, 2VYU, 2X8M, 2X8O, 2X8P	soluble	choline binding protein	CHO	quaternary amine held by π-cation box
choline binding domain from autolysin (*S. pneumoniae*)	1GVM, 1H8G, 1HCX	soluble	choline binding protein	CHO	quaternary amine held by π-cation box
choline binding domain of Spr1274 (*S. pneumoniae*)	3HIA	soluble	choline binding protein	CHO	quaternary amine held in π-cation box; hydroxyl held by a phosphate via an Asn and a Ser
choline binding protein (*Sinorhizobium meliloti*)	2REG	membrane bound	ABC transporter	CHO	quaternary amine held by π-cation box, Asp and Asn form polar contacts
choline acetyl-transferase (human)	2FY3	soluble	generates acetylcholine	CHO	quaternary amine held by π-cation box
OpuBC (*Bacillus subtilis*)	3R6U, 3PPQ	soluble	choline binding protein	CHO	quaternary amine held by π-cation box
acetylcholine binding protein (*Lymnaea stagnalis*)	1UV6, 2XZ5	membrane bound	acetylcholine binding protein	CCE	quaternary amine held by π-cation box
C-reactive protein (human)	1b09	soluble	phosphocholine binding protein	PC	bound via the phosphate
PDV-109 fibronectin module (bovine)	1H8P	interfacial	choline lipid binding protein	PC	quaternary amine is held by a π-cation box
IGA-Kappa MC/PC603 FAB (mouse)	2MCP	soluble	phosphocholine binding protein	PC	quaternary amine is held by a π-cation box; phosphate is coordinated by an Arg

Protein	PDBID	Membrane Interactions	Function	Ligand	Binding Mode
immune response protein (human)	1DL7	soluble	phosphocholine binding protein	NCH	quaternary amine is held by a π-cation box; polar contacts to phosphate
phosphatidylcholine transfer protein (human)	1LN1, 1LN2, 1LN3	interfacial	PC lipid transfer protein	DLP	quaternary amine is held by a π-cation box; also held by acyl chains, ligand is very tightly bound
phosphatidylinositol transfer protein (human, rat)	1UW5, 2A1L, 1T27	interfacial	PI/PC transfer protein	PC, PCW	acyl chains and phosphate bound; choline group is apparently uncoordinated
Sec14 homolog (yeast)	3B7Q, 3B7Z	soluble	PI/PC transfer protein	6PL	quaternary amine held by π-cation box; lipid chains also bound
autolysisn (*S. pneumoniae*)	2WW5, 2WWC, 2WWD	soluble	peptidoglycanase	CHO	quaternary amine held by π-cation box
Endolysin (phage CP-1)	1OBA	soluble	peptidoglycanase	CHO	quaternary amine held by π-cation box
butyryl cholinesterase (human)	1P0M	soluble	esterase	CHO	quaternary amine held by π-cation box
phosphocholine esterase domain from CPBE (*S. pneumoniae)*	1WRA, 2BIB	interfacial	esterase	PC	mostly bound through phosphate; single π-cation interaction with quaternary amine
Acetylcholinesterase (mouse)	2HA2, 2HA3,2HA6, 2HA5, 2HA7	soluble	esterase	SCU, CHO, SCK, BCH	quaternary amine held by π-cation box, additional polar contacts bind molecule
butyrylcholine esterase (human)	1P0P	soluble	esterase	BCH	quaternary amine held by π-cation box, as well as various polar contacts on the rest of the molecule
Acetylcholinesterase (*Torpedo californica)*	2C58, 2C4H	soluble	esterase	BCH	quaternary amine held by π-cation box
choline kinase (*Plasmodium knowlesi*)	3C5I	soluble	kinase	CHO	quaternary amine held in π-cation box, hydroxide coordinated by an Asp and Glu
choline kinase α2 (human)	2CKQ	soluble	kinase	PC	quaternary amine held in π-cation box; phosphate coordinated by Ser, Asp, Asn
potassium channel (*Burkholderia psedomallei*)	2WLL	membrane bound	ion channel	PLC	bound via the phosphate only
potassium channel *(Magnetospirillum magnetotacticum*)	2x6a, 2x6b, 2x6c	membrane bound	ion channel	PC	binding to phosphate dominant; a single π-cation interaction

Protein	PDBID	Membrane Interactions	Function	Ligand	Binding Mode
Na^+ coupled betaine symporter (*Corynebacterium glutamicum*)	3P03	membrane bound	transporter	CHO	quaternary amine held by π-cation box, additional polar contacts bind molecule
LeuT (*Aquifex aeolicus*)	3USG, 3USL, 3USM,	membrane bound	sodium symporter	PC	bound by the lipid chains; no interaction with phosphocholine
ligand gated ion channel (*Aquifex aeolicus*)	3p4w, 3p50, 3uu5, 3uu8,3uub, 3EAM	membrane bound	ion channel	PLC, GPC	lipid chains are bound; polar contacts to the phosphate and glycerol; no interactions with choline
Xanthorhodopsin (*Salinibacter ruber*)	3DDL	membrane bound	ion channel	PX4	bound by the lipid chains; phosphate is visible but not bound; choline head group is not ordered
NavAB (*Arcobacter butzleri*)	3RVY, 3RVZ, 3RW0	membrane bound	voltage gated ion channel	PX4	bound by the lipid chains; phosphate is visible but not bound; choline head group is not ordered
SR calcium pump (*Oryctolagus cuniculus)*	2ZBD, 3AR2	membrane bound	ion pump	GPC	bound by lipid chains and phosphate; choline is not bound
sodium potassium pump (porcine)	3B8E	membrane bound	ion pump	GPC	bound by phosphate only; lipid chains are not ordered
voltage dependent ion anion channel (mouse)	3EMN	membrane bound	ion channel	MC3	bound by acyl chain and polar contact to glycerol; choline is not ordered
Lipase (*Thermomyces lanuginose*)	1EIN	interfacial	lipase	PLC	bound by the lipid chains only; choline is not bound
lipase-procolipase (human)	1LPA	interfacial	lipase	PLC	bound by the lipid chains only; choline is not bound
phospholipase A_2 (porcine)	1L8S	interfacial	lipase	LPE	bound by the lipid chains only; choline is not bound
Phospholipase (*Bacillus cereus*)	1P6D, 1P6E	interfacial	phospholipase	3PC, PC5	extensive coordination to sulfur (PC5); Glu coordination to the choline
cytochrome BC1 (chicken)	3H1I, 3H1J	membrane bound	mitochondrial electron transport	PLC	ligand is held by acyl chains, with 1 Asp near choline
photosynthetic reaction center (*Rhodobacter sphaeroides* R26)	2HG9,2HH1,2HG3, 2J8C, 1M3X	membrane bound	photosynthetic reaction center	PCK, GPC	held by lipid chains only
mitochondiral ADP/ATP carrier (bovine)	1OKC	membrane bound	ATP transport	GPC	coordinated by phosphate only
cytochrome BC1 complex (yeast)	1KB9	membrane bound	electron transport chain	PCF	held by lipid chain, phosphate makes polar contacts

Protein	PDBID	Membrane Interactions	Function	Ligand	Binding Mode
cytochrome BC1 complex (yeast)	1P84	membrane bound	electron transport chain	GPC	held by lipid chains only
cytochrome C oxidase (bovine)	1V54, 2DYR, 3AG2, 3AG3, 1V55, 2EIJ, 3ABM, 2ABK, 3AG4, 2EIK, 2EIL, 3ABL, 2DYS, 3AG1, 2ZXW, 2EIM, 2EIN	membrane bound	electron transport chain	PSC	held by choline head group with 1 Tyr, and 1 Asp, extensive polar contacts with phosphate, complete hydrophobic interaction with the acyl chain
cytochrome BC1 complex (yeast)	1P84	membrane bound	electron transport chain	GPC	held by lipid chains only
cytochrome C oxidase (*Paracoccus denitrificans*)	1QLE	membrane bound	electron transport chain	GPC	held by lipid chains and polar contact to phosphate
cytochrome B6F (*Nostoc* sp. PCC7120*)*	2ZT9	membrane bound	electron transport chain	OPC	held by lipid chain; phosphate makes polar contacts
ATP synthase rotor ring (*Bacillus pseudofirmus*)	2X2V	membrane bound	ATP synthesis	DPV	bound by lipid chain and phosphate polar contacts; Glu coordinates quaternary amine
cytochrome B6F (*Mastigocladus laminosus*)	2E74, 2E75, 2E76, 1Vf5, 2D2C	membrane bound	electron transport chain	OPC	bound by lipid chains only; PC head group is far above the protein surface
nuclear liver receptor homolog (human)	4DOS	soluble	binds phospholipids	PLC	lipid chains and phosphate are bound; choline group sticks out above protein surface
Lipovitellin (*Ichtyomyzon unicuspis*)	1LSH	interfacial	lipid and metal storage	PLC	bound by lipid chains only; PC head group is far above the protein surface
bacteriacidal/perme-ability increasing protein (human)	1EWF, 1BO1	interfacial	Lipopolysaccha-ride binding	GPC	lipid chains and phosphate are bound; choline sticks out above the surface of the protein
non-specific lipid transfer protein (wheat)	1BWO, 1MID	membrane bound	lipid transfer protein	LPC	bound by lipid chain; phosphate and glycerol make polar contacts; choline is above the surface of the protein
GM2-activating protein (human)	2AG2	interfacial	lysosomal lipid transfer protein	LP3	bound by lipid chains only
SF-1 (mouse)	3F7D	interfacial	lipid binding nuclear receptor	P42	lipid chains and phosphate are bound; choline group sticks out above protein surface
Wnt inhibitory factor 1 (human)	2YGN, 2YGO, 2YGP, 2YGQ	soluble	growth factor inhibitor	PCF	single Phe within π-cation distance; but mostly held by lipid chains

Protein	PDBID	Membrane Interactions	Function	Ligand	Binding Mode
Wnt inhibitory factor 1 (human)	2YGN, 2YGO, 2YGP, 2YGQ	soluble	growth factor inhibitor	PCF	single Phe within π-cation distance; but mostly held by lipid chains
calmodulin (bovine)	3IF7	soluble	Ca^{2+} binding protein	SPU	lipid chains bound; choline is coordinated by two Glu
Rhodopsin (squid)	2Z73	membrane bound	GPCR	GPC	bound by lipid chain; phosphate and glycerol make polar contacts; choline not ordered
rhomboid protease GLPG (*E. coli*)	2XTV	membrane bound	protease	MC3	lipid chains bound; polar contact with phosphates; cholines are above the surface of the protein
SNARE ykt6 longin domain (rat)	3KYQ	interfacial	membrane fusion protein	DPV	bound by lipid chains, polar contacts, and a single π-cation interaction with quaternary amine
pore-forming cytolysin sticholysin II (*Stichodactyla helianthus*)	1O72	interfacial	pore forming protein	PC	quaternary amine held by π-cation box
leukociden F (*Staphylococcus aureus*)	3LKF	interfacial	self assembling channel forming protein	PC	quaternary amine held by a π-cation; phosphate bound by Arg
lens specific aquaporin (*Ovis aries*)	2B6O	membrane bound	water pore forming protein	MC3	bound by lipid chain; phosphate makes polar contacts; choline is above the surface of the protein
bacterial dynamin-like protein (*Nostoc punctiforme*)	2W6D	membrane bound	membrane fusion protein	CPL	held by phosphates or lipid chains, not choline
GM2 activator protein (human)	1TJJ	soluble	lipid transfer protein	PFS	lipid chains are bound; polar contacts are formed with the phosphates
phosphoethanolamine methyltranserase (*Plasmodium falciparum*)	3UJ9, 3UJC, 3UJD	soluble	methyl transferase	PC	mostly coordinated by the phosphate; one π-cation interaction with quaternary amine
CD1D antigen (mouse)	2FIK, 2GAZ, 2H26, 1ZHN	soluble	lipid binding antigen protein	6PL, PC6	acyl chains only; PC head group is far above the protein surface
MD-1 lymphocyte antigen (mouse)	3M7O	soluble	antigen	L9R	ligand bound via lipid chains and polar contacts with phosphate; one π-cation interaction with the amine
CD1b3 (bovine)	3L9R	soluble	antigen?	L9R	bound via acyl chains; phosphocholine is above the surface of the protein

Protein	PDBID	Membrane Interactions	Function	Ligand	Binding Mode
L-ficolin (human)	2J0H	soluble	lectin-like immune defense protein	ACH	anchored via acetyl group
BmrR (*B. subtilis*)	3Q5S	soluble	multidrug resistance gene	ACH	anchored via acetyl group
PaCTD (*Pseudomonas arvilla*)	2AZQ, 1DMH, 1DLT, 1DLM, 1DLQ	interfacial	1,2 dioxygenase	PCF	bound via lipid chains and glycerol group; phosphocholine sticks up above the protein
hydoxyquinol 1,2 dioxygenase (*Nocardioides simplex*)	1TMX, 3N9T, 1S9A	soluble	dioxygenase	HGX	bound via acyl chains; phosphocholine is not ordered
Rho 1,2-CTD (*Rhodococcus opacus*)	3HHY,3HJS, 3I51, 3HKP, 3I4Y, 3HGI, 3HHX,3HJQ, 3I4V, 3HJ8	soluble	dioxygenase	6PL	bound by lipid chain and polar contacts to the glycerol

Ligand abbreviations in Table SI.

Abbreviation	compound
3PC	(3S)-3,4-di-N-hexanoyloxybutyl-1-phosphocholine
6PL	1-palmitoyl-2-stearoyl-sn-glycero-3-phosphocholine
ACH	acetylcholine
BCH	butyrylthiocholine
CCE	carbamylcholine
CHO	choline
CPL	1-palmitoyl-2-linoleoyl-phosphatidylcholine
DLP	1,2-dilinoleoyl-sn-glycero-3-phosphocholine
DPV	n-dodecylphosphocholine
GPC	glycerophosphocholine
HGX	1-heptadecanoyl-2-tridecanoyl-3-glycerolphosphonylcholine
L9R	1-stearoyl-2-oleoyl-sn-glycero-3-phosphocholine
LP3	(7R)-4,7-dihydroxy-N,N,N-trimethyl-10-oxo-3,5,9-trioxa-4-phosphaheptacosan-1-aminium 4-oxide
LPC	(1-myristoyl-glycerol-3-yl)phosphonylcholine
LPE	1-O-octadecyl-sn-glycerol-3-phosphocholine
MC3	1,2-dimyristoyl-rac-glycerol-3-phosphocholine
NCH	p-nitrophenylphosphocholine
OPC	dioleoylphosphatidylcholine
P42	1-stearoyl-2-palmitoyl-sn-glycero-3-phosphocholine
PC	phosphocholine
PC5	1,2-di-N-pentanoyl-sn-glycero-3-dithiophosphocholine
PC6	7-[(dodecanoyloxy)methyl]-4-hydroxyl-N,N,N-trimethyl-9-oxo-3,5,8-trioxa- 4-phosphado-triacontan-1-aminium 4-oxide
PCF	1,2-dipalmitoylphosphatidylcholine
PCK	1,2-distearoyl(9,10-dibromo)-sn-glycero-3-phosphocholine
PCW	1,2-dioleoyl-sn-glycero-3-phosphocholine
PFS	1-O-octadecyl-2-acetyl-sn-glycerol-3-phosphocholine
PLC	diundecylphosphatidylcholine
PSC	1-palmitoyl-2-linoleoyl-sn-glycerol-3-phosphocholine
PX4	1,2-dimyristoyl-sn-glycero-3-phosphocholine
SCK	succinyl dicholine
SCU	succinylcholine

Table S1 shows one overwhelming theme: if a protein specifically binds choline, either as the substrate or product of a reaction, or as a moiety to be transferred, it does so via a protein π-choline cation box. Proteins that bind lipids and are non-specific as to the head group, or proteins which are integral membrane proteins, bind ligand mostly via the lipid chains, or possibly via interactions with the phosphate. Out of 12 proteins that have choline as a ligand, all 12 bound choline via a π-cation box. Similarly, of the 10 proteins that contain phosphocholine, 9 bound the quaternary amine in a π-cation box, while one did not. Conversely, of the 41 proteins that bound phospholipids with a choline head group, only a few had π-cation interactions, and of these there was usually only a single defined interaction of the cation with an aromatic group. Overall, of the 67 different proteins that bind choline or choline containing ligands, 43% of them bind the choline via a π-cation interaction.

HEPES binding to N254Y/H258Y

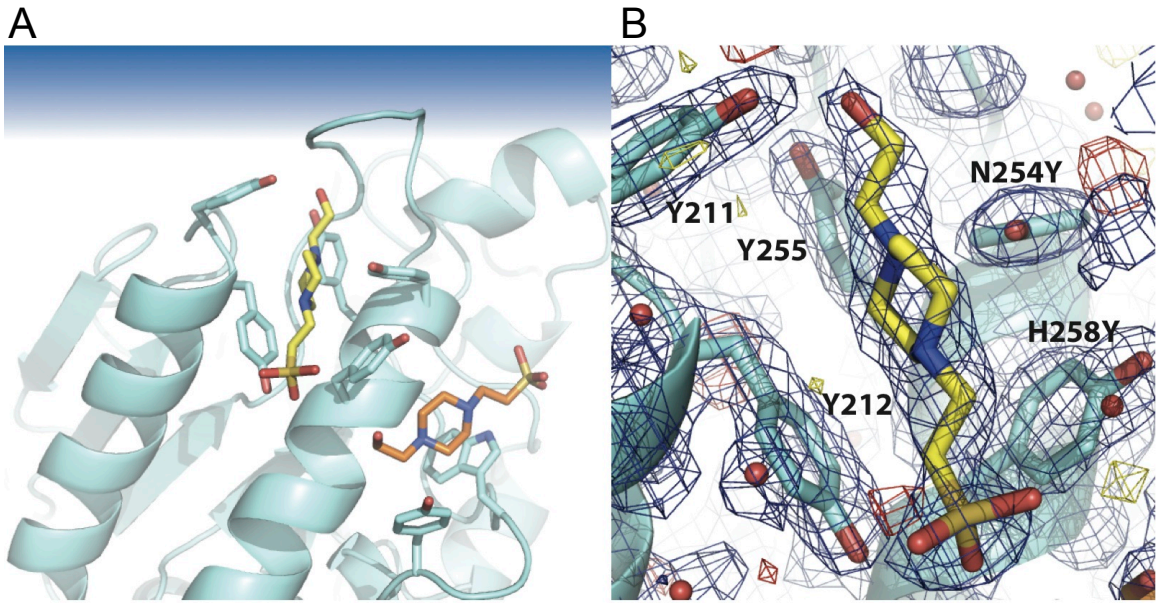

Figure S1. (A) N254Y/H254Y structure with HEPES in the choline binding site (PDB 4F2Y). Representative electron density for HEPES ligands (B). Electron density shown in dark blue and contoured at 1σ, is shown with the model superimposed.

Initially, we generated crystals of N254Y/H258Y from a solution containing 10% isopropanol, 22% PEG 4000, and 100 mM HEPES, pH 7.5. Prior to crystallization, the protein was incubated with 30 mM glycerophosphocholine and 1 mM diC_7PC for 2 h, though neither of these molecules was evident in the structure. Instead two molecules of HEPES were observed in the helix F/G region (Figure S2). In contrast, when wild type *S. aureus* PI-PLC was crystallized under similar conditions, no small molecules were bound in this region. In this N254Y/H258Y structure (PDB entry 4I9M) one HEPES molecule lies in choline binding site 1, directly between helices F and G, with a cationic nitrogen held between the π systems of Tyr212 and Tyr258. Likewise the cationic nitrogen of the second molecule of HEPES is held in choline binding site 2, between the π systems of Trp287, Tyr290 and Tyr258. Tyr254 provides edge face interactions with Tyr258. In order to accommodate the HEPES molecule in the binding pockets, the side chain of Tyr212 must rotate downward from its position in the unliganded double mutant structure by 94°, while the side chain of Tyr258 must rotate upward 99°. The rotation of Tyr212 forms the left sidewall of binding site one, while Tyr258 forms the right side of this binding site, as well as the left side of binding site two.

Previously, we obtained a structure of H258Y crystallized in acidic conditions (Goldstein et al (2012) *Biochemistry* 51: 2579-2587), but more recently obtained crystals of this mutant protein from a solution containing 10% isopropanol, 22% PEG 4000, and 100 mM HEPES, pH 7.5. The basic H258Y structure (PDB entry 4I9T) has no clear indication of HEPES occupying the sites seen in N254Y/H258Y, although there was some electron density consistent with a low occupancy (<30%) ligand bound in site 2. However, the refinement is best with a small PEG molecule at that site. The reason for the decreased affinity for HEPES in the H258Y single mutant becomes clear when examining the side chains of the binding pockets. Without the mutation N254Y, the edge face interaction between Tyr254 and Tyr258 is lost, this allows the side chain of Tyr258 to exist in either the two distinct positions – one without a ligand (57%) or one where a ligand is present (43%). More importantly, without the mutation of Asn254 to the significantly larger Tyr residue, the side chain of Tyr212 is not completely rotated down into the mutant conformation, and instead appears primarily in the conformation of the WT structure (70%) rather than that of the HEPES bound mutant structure (30%). This implies that both N254Y and H258Y are needed for discrete PC binding to the *S. aureus* PI-PLC.

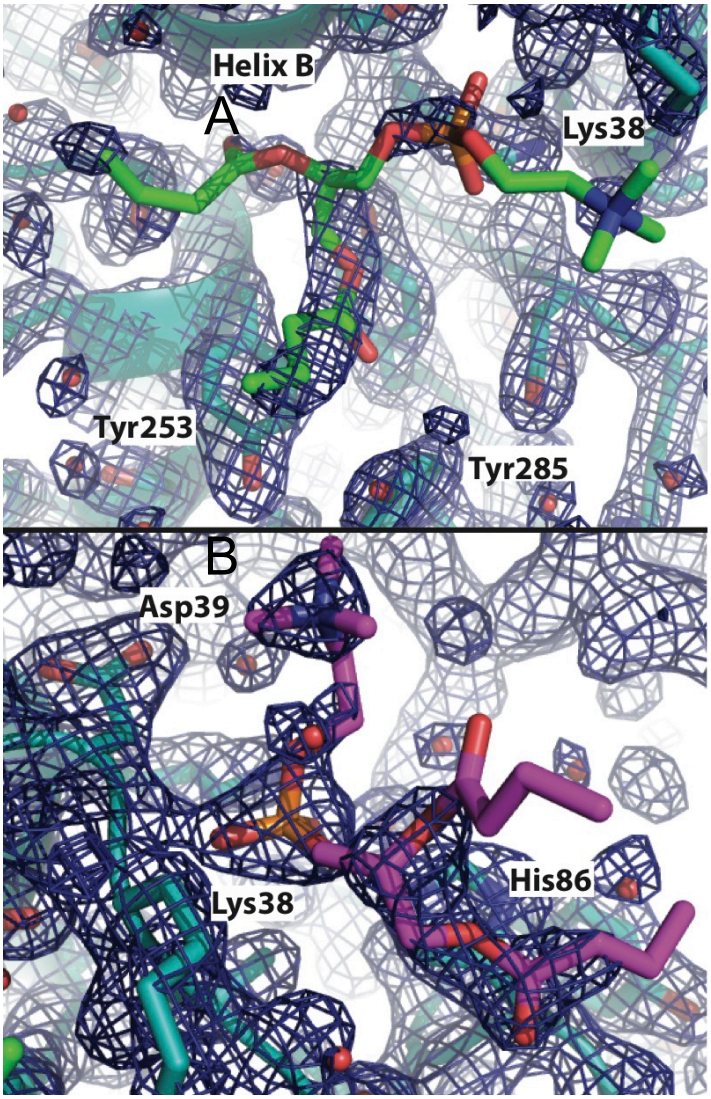

Figure S2. View of the two auxiliary diC_4PC molecules found in the N254Y/H258Y crystal.

(A) One phospholipid, under helix B, is bound most likely by hydrophobic interactions of one acyl chain with Tyr253 and Tyr585, although the overall the density is mediocre.

(B) The secondary auxiliary phospholipid is bound in the anion binding site described in the WT protein (Cheng et al. (2012) *J. Biol. Chem.* 287, 40317-40327). The phosphate is coordinated by Lys38 and His86; densities for the phosphate, quaternary amine, and glycerol backbone are excellent, while the lipid chains are more mobile. Electron density, shown in dark blue in both, is for 2Fo-Fc at 1 σ.

Table S2. Crystallographic data for selected *S. aureus* PI-PLC crystals.

Crystal	N254Y/H258Y + HEPES	H258Y: Basic
PDB ID	4I9M	4I9T
Resolution Range (Å)	2.19-30.85	1.99-37.63
No. of reflections	15898	20933
Reflections in free set	795	1072
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell		
a (Å)	86.22	85.59
b (Å)	56.12	57.83
c (Å)	61.71	60.78
Completeness	99.5%	99.3%
R_{merge}	11.7	7.6
Protein molecules in A.U.	1	1
R_{cryst} [a]	0.1580	0.1786
R_{free} [b]	0.2358	0.2324
No. residues	304	305
No. non-hydrogen protein atoms	2437	2498
No. H_2O molecules	342	195
No. SO_4^{-2} ions	1	1
r.m.s.d. bonds (Å)	0.02	0.20
r.m.s.d. angles (^{o})	1.70	1.98
Ramachandran plot (%)		
Most favored	90.1	96.49
Additionally allowed	9.9	3.19
Generously allowed	0	0.32
Disallowed	0	0
Average B-factor (Å^2)	33.45	25.64

[a] $R_{cryst} = \{\Sigma\ (||F_o| - |F_c||) / |F_o|\}$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure factor amplitudes, respectively.

[b] Brunger (1992).