Supporting Information

Xiao et al. 10.1073/pnas.1109290108

SI Methods

Protein Purification. Cell pastes of GST-PTPMT1$^{69–261}$ [protein tyrosine phosphatas (PTP) localized to the Mitochondrion 1) were resuspended in 25 mM Tris (pH 7.5), 150 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5 mM [Tris(2-carboxyethyl)phosphine] (TCEP), 1 mM Pefabloc SC, 1 mM benzamidine-HCl, 1 mM PMSF (phenylmethylsulphonyl fluoride) and disrupted by sonication. GST-PTPMT1$^{69–261}$ was purified by affinity chromatography using GST-Bind Resin (Novagen) and eluted with 50 mM Tris (pH 8.0), 150 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5 mM TCEP, and 15 mM reduced glutathione. PTPMT1$^{69–261}$ mutants were generated using site-directed mutagenesis PCR and purified similarly.

Cell pastes of His_6-SUMO-PTPMT1$^{100–261}$ were resuspended in 50 mM Tris (pH 7.4), 300 mM NaCl, 20 mM imidazole, 0.5 mM TCEP, 1 mM Pefabloc SC, 1 mM benzamidine-HCl, 1 mM PMSF and disrupted by sonication. The fusion protein was purified by Ni-NTA affinity chromatography and digested with the small ubiquitin-like modifier (SUMO) specific protease ULP1. His_6-SUMO and the ULP1 protease were removed by a second Ni-NTA affinity chromatography step. Untagged PTPMT1$^{100–261}$ was adjusted to 20 mM Hepes (pH 7.0), 100 mM NaCl, 2 mM DTT, applied to a Resource-S cation exchange column (GE Healthcare), and eluted with a linear gradient of NaCl. Peak fractions were buffer exchanged into 20 mM Hepes (pH 7.0), 100 mM NaCl, 2 mM DTT, concentrated to 15.5 mg/mL, and flash-frozen with liquid nitrogen. Se-Met substituted PTPMT1 was expressed in the methionine auxotroph *Escherichia coli* strain B834(DE3) (Novagen) using the SelenoMet minimal media plus Nutrient Mix (AthenaES) supplemented with 75 mg/L L(+)-Selenomethionine (Acros). Both Se-Met and C200S mutant of PTPMT1 were purified similarly to the wild-type protein and concentrated to 14.4 mg/mL for crystallization.

Structure Determination. Diffraction data were collected at the Advanced Light Source (ALS) beam line 8.2.1 and processed with HKL2000 (HKL Research) and CCP4 (1). The structure of PTPMT1 was determined by the multiple-wavelength anomalous diffraction (MAD) method using the data collected from a Se-Met crystal. Heavy atom search, phase calculation and refinement, density modification, and initial model building were carried out with the AutoSol (2) program in the Phenix software suite (3). The rest of the molecule was traced manually with Coot (4) and refined using Phenix. 5% randomly selected reflections were used for cross-validation (5), and the refinement strategy included simulated annealing, individual coordinates and *B*-factors refinements, and TLS (translation, libration, and screw-rotation) refinement (6). The structure of the PTPMT1–phosphatidylinositol 5-phosphate [PI(5)P] complex was determined by the molecular replacement method using PHASER (7). PTPMT1 structure obtained above was used as search a model. The model was adjusted using Coot and refined using Phenix. A structure model of PI(5)P is initially derived from PI(3,5)P_2 in the MTMR2-PI(3,5)P_2 structure [Protein Data Bank (PDB) ID code 1ZVR] (8), and the geometry restraint information used for refinement is generated using the eLBOW program (9). The head group of PI(5)P is well-defined by the electron density and has *B*-factors close to the protein molecule; whereas the diacylglycerol moiety is less well defined, has high *B*-factors and possibly partial occupancies. As a result, only a portion of the diacylglycerol group is built into the final structural model.

Thin Layer Chromatography. Silica gel 60 plates (Whatman) were activated by baking for 1 h at 180 °F under vacuum. The dried products of the phosphatidylglycerol phosphate (PGP) phosphatase assays were resuspended in 40 μL of Chloroform and spotted onto the TLC plate. The plate was dried for 10 min in a fume hood and developed in Chloroform/Methanol/glacial acetic acid (65 : 25 : 8). The plate was again dried in a fume hood, exposed to a storage phosphor screen, and analyzed using a Typhoon 9410 (GE Healthcare).

1. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50:760–763.
2. McCoy AJ, Storoni LC, Read RJ (2004) Simple algorithm for a maximum-likelihood SAD function. *Acta Crystallogr D Biol Crystallogr* 60:1220–1228.
3. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213–21.
4. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132.
5. Brunger AT (1992) Free R value: A novel statistical quantity for assessing the accuracy of crystal structures. *Nature* 355:472–475.
6. Afonine PV, Grosse-Kunstleve RW, Adams PD (2005) A robust bulk-solvent correction and anisotropic scaling procedure. *Acta Crystallogr D Biol Crystallogr* 61:850–855.
7. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40:658–674.
8. Begley MJ, et al. (2006) Molecular basis for substrate recognition by MTMR2, a myotubularin family phosphoinositide phosphatase. *Proc Natl Acad Sci USA* 103: 927–932.
9. Moriarty NW, Grosse-Kunstleve RW, Adams PD (2009) Electronic ligand builder and optimization workbench (eLBOW): A tool for ligand coordinate and restraint generation. *Acta Crystallogr D Biol Crystallogr* 65:1074–1080.

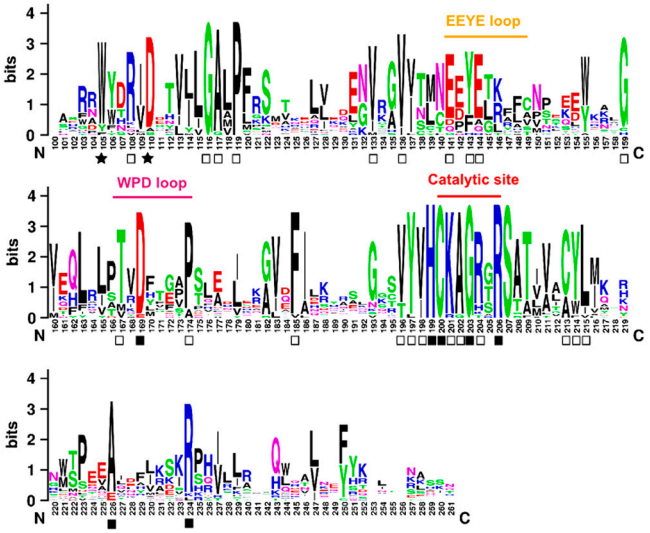

Fig. S1. A conservation logo of the phosphatase domain of PTPMT1. The overall height of the letters indicates the sequence conservation at each position, and the height of each letter reflects the relative frequency of amino acids at that position. The stars denote the residues only conserved in PTPMT1, the filled black rectangles denote residues conserved throughout dual-specificity phosphatases (DSPs), and the empty rectangles denote the residues conserved not only in PTPMT1 but also in some other DSP subfamilies. A total of 28 sequences are used to generate the sequence logo.

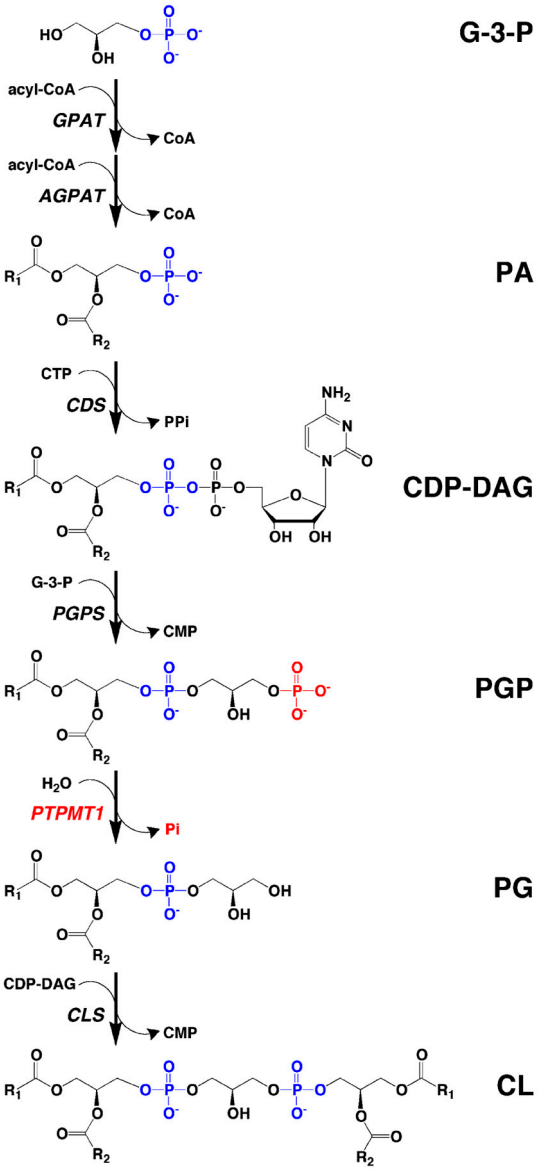

Fig. S2. PTPMT1 functions in the cardiolipin biosynthesis pathway. Glycerol-3-phosphate (G-3-P) is first acylated by the sequential activities of glycerol-3-phosphate acyltransferase (GPAT) and 1-acylglycerol-3-phosphate acyltransferase (AGPAT). This results in the formation of phosphatidic acid (PA), which is then converted to a reactive cytidinediphosphate-diacylglycerol (CDP-DAG) intermediate by the CDP-DAG synthase (CDS). Through the activity of the PGP synthase (PGPS), CDP-DAG further reacts with a second molecule of glycerol-3-phosphate to form PGP. PGP is subsequently dephosphorylated, resulting in the formation of phosphatidylglycerol (PG). Finally, PG condenses with another molecule of CDP-DAG to form cardiolipin (CL), catalyzed by the cardiolipin synthase (CLS). Figure adapted from ref. 1.

1 Voet D, Voet JG (2004) Biochemistry (Wiley, Hoboken, NJ), 3rd Ed, Vol 1

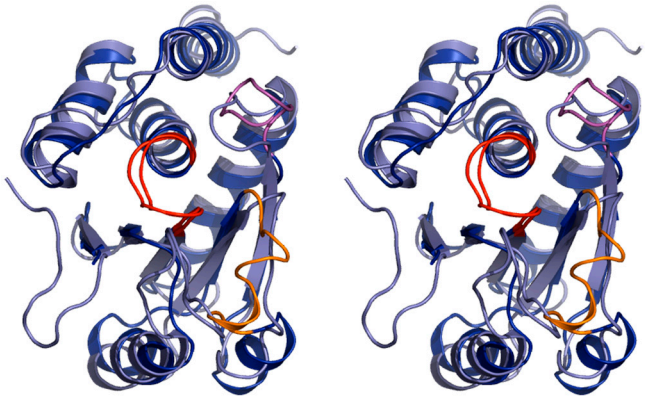

Fig. S3. Superimposition of PTPMT1 and PTEN. PTPMT1 and PTEN are shown using the same color schemes as in Fig. 2.

Table S1. Data collection, phasing and refinement statistics

	Apo	Se-Met			PI(5)P
Data collection					
Space group	$P4_32_12$		$P4_32_12$		C2
Cell dimensions					
a, *b*, *c* (Å)	36.45, 36.45, 230.02		36.83, 36.83, 228.98		90.02, 67.91, 32.09
α, β, γ (°)	90, 90, 90		90, 90, 90		90, 96.6, 90
		Peak	*Inflection*	*Remote*	
Wavelength (Å)	1.0	0.9798	0.9800	0.9573	1.0
Resolution (Å)	1.93	2.0	2.0	2.0	2.05
R_{sym} or R_{merge}	5.1 (70.5)	5.6 (50.4)	4.8 (49.5)	4.8 (53.0)	6.2 (24.0)
$I/\sigma I$	61.9 (3.0)	44.0 (2.7)	44.7 (2.7)	43.2 (2.7)	16.7 (3.7)
Completeness (%)	99.9 (100.0)	99.9 (100.0)	99.9 (100.0)	99.9 (100.0)	97.8 (92.6)
Redundancy	12.7 (13.2)	7.4 (7.0)	7.4 (7.0)	7.4 (7.6)	3.7 (3.4)
Refinement					
Resolution (Å)	1.93				2.05
No. reflections	12762				11902
R_{work}/R_{free}	22.7/24.3				17.3/21.5
No. atoms					
Protein	1257				1248
Ligand/ion	10				31
Water	48				141
B-factors					
Protein	51.7				26.5
Ligand/ion	53.6				46.9
Water	52.3				36.5
R.m.s deviations					
Bond lengths (Å)	0.004				0.003
Bond angles (°)	0.776				0.772

Each dataset was collected from a single crystal. Values in parentheses are for highest-resolution shell.