The Plant Cell, Vol. 27: 1113–1127, April 2015, www.plantcell.org

Structural Mechanism Underlying the Specific Recognition between the Arabidopsis State-Transition Phosphatase TAP38/PPH1 and Phosphorylated Light-Harvesting Complex Protein Lhcb1[OPEN]

Xuepeng Wei,[a,b] Jiangtao Guo,[a] Mei Li,[a] and Zhenfeng Liu[a,1]

[a] National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China
[b] School of Life Sciences, University of Chinese Academy of Sciences, Beijing100101, China

ORCID ID: 0000-0001-5502-9474 (Z.L.)

During state transitions, plants regulate energy distribution between photosystems I and II through reversible phosphorylation and lateral migration of the major light-harvesting complex LHCII. Dephosphorylation of LHCII and the transition from state 2 to state 1 requires a thylakoid membrane-associated phosphatase named TAP38 or PPH1. TAP38/PPH1 specifically targets LHCII but not the core subunits of photosystem II, whereas the underlying molecular mechanism of their mutual recognition is currently unclear. Here, we present the structures of *Arabidopsis thaliana* TAP38/PPH1 in the substrate-free and substrate-bound states. The protein contains a type 2C serine/threonine protein phosphatase (PP2C) core domain, a Mn^{2+} (or Mg^{2+}) binuclear center and two additional motifs contributing to substrate recognition. A 15-mer phosphorylated N-terminal peptide of Lhcb1 binds to TAP38/PPH1 on two surface clefts enclosed by the additional motifs. The first segment of the phosphopeptide is clamped by a pair of tooth-like arginine residues at Cleft 1 site. The binding adopts the lock-and-key mechanism with slight rearrangement of the substrate binding residues on TAP38/PPH1. Meanwhile, a more evident substrate-induced fitting occurs on Cleft 2 harboring the extended part of the phosphopeptide. The results unravel the bases for the specific recognition between TAP38/PPH1 and phosphorylated Lhcb1, a crucial step in state transitions.

INTRODUCTION

In algae and land plants, imbalanced excitation of the two photosystems (photosystem I [PSI] and photosystem II [PSII]) will occur when the light intensity and quality are constantly fluctuating in the natural environment. To avoid this, plants are capable of responding to the changes in light conditions and counteracting the potential photodamage and photoinhibition effects caused by the imbalance. The state transition is a fundamental photosynthetic acclimation mechanism that operates efficiently to balance the distribution of excitation energy between PSI and PSII under varying light conditions (Bonaventura and Myers, 1969; Murata, 1969a, 1969b). The state transition process in plants is dependent on reversible phosphorylation and dephosphorylation of the major light-harvesting complex LHCII (Allen, 1992). It is also accompanied by reorganization of the thylakoid membranes and lateral migrations of the mobile LHCII pool to distribute excitation energy between the two photosystems. Under low white light or red light (with the wavelengths specific for exciting PSII), overexcitation of PSII leads to reduction of the plastoquinol pool and activation of LHCII kinase by the cytochrome b_6f complex (Vener et al., 1997; Zito et al., 1999). Subsequently, a portion of LHCII, mainly in the form of a $(Lhcb1)_2Lhcb2$ heterotrimer, is phosphorylated, dissociates from the periphery of PSII, and migrates to PSI to rebalance the excitation energy distribution (Pietrzykowska et al., 2014). Meanwhile, phosphorylation of the bulk LHCII, with Lhcb1 as its major protein component, leads to reorganization of the grana structure, which may facilitate phosphorylation and mobilization of the peripheral LHCII. On the other hand, when the light composition is dominated by far-red light (with the wavelengths specifically exciting PSI), or under completely dark conditions, the oxidized plastoquinol pool will induce inactivation of the LHCII kinase (Rochaix, 2013). Alternatively, the LHCII kinase may be inhibited by the ferredoxin-thioredoxin system under high white light conditions (Rintamäki et al., 2000). Through the catalytic action of an LHCII phosphatase, the phospho-LHCII is dephosphorylated; subsequently, the mobile fraction of LHCII returns to the periphery of PSII to reequilibrate the excitation energy.

State transitions require a chloroplast kinase and phosphatase pair for the phosphorylation and dephosphorylation of LHCII, respectively (Rochaix et al., 2012). The identities of the kinase and phosphatase responsible for LHCII phosphorylation and dephosphorylation during state transitions have remained elusive for over 30 years. Stt7 from *Chlamydomonas reinhardtii* and STN7 from *Arabidopsis thaliana* were identified as the kinases crucial for the phosphorylation of LHCII and the transition from state 1 to state 2 (Depège et al., 2003; Bellafiore et al., 2005). In the *stt7* or *stn7* knockout plants, phosphorylation of LHCII stays at a low level regardless of the changing light conditions. The Stt7 and STN7 proteins have a small N-terminal region at the luminal side, a single transmembrane helix, and a large catalytic domain at the stromal side (Rochaix, 2007). The catalytic domain adopts a canonical fold of the Ser/Thr kinase family. It contains an ATP binding site

[1] Address correspondence to liuzf@sun5.ibp.ac.cn.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Zhenfeng Liu (liuzf@sun5.ibp.ac.cn).
[OPEN]Articles can be viewed online without a subscription.
www.plantcell.org/cgi/doi/10.1105/tpc.15.00102

sandwiched between two lobes and an additional Stt7/STN7-characteristic motif inserted near the kinase activation loop region (Guo et al., 2013). Despite extensive efforts, the mechanisms underlying substrate recognition and activity regulation of Stt7/STN7 are still enigmatic.

Investigations on the LHCII phosphatase date back to 1980 when a thylakoid membrane-associated phosphatase activity was reported to be capable of dephosphorylating LHCII (Bennett, 1980). This phosphatase activity was stimulated by $MgCl_2$ and inhibited by NaF. Further work reported that another phosphatase activity, peripherally associated on the thylakoid membrane, can also dephosphorylate thylakoid proteins (Sun et al., 1989). Later on, a 29-kD stromal phosphatase was purified and found to be capable of dephosphorylating LHCII, too (Hammer et al., 1995, 1997). The activities of thylakoid protein phosphatases are independent of the redox states, and the kinetics of dephosphorylation reactions are heterogeneous among different substrates, including LHCII and PSII core subunits D1, D2, and CP43 (Silverstein et al., 1993). More recently, two groups independently identified, through genetic and biochemical approaches, a chloroplast phosphatase specifically involved in LHCII phosphorylation (Pribil et al., 2010; Shapiguzov et al., 2010). This phosphatase was named differently by the two groups, either as the Thylakoid-Associated Phosphatase of 38 kD (TAP38) (Pribil et al., 2010) or as Protein Phosphatase 1 (PPH1) (Shapiguzov et al., 2010). State transitions in *tap38*/*pph1* knockout plants were blocked in state 2 and the dephosphorylation of LHCII was strongly impaired in these mutants. Moreover, in vitro assays indicated that recombinant TAP38 is able to dephosphorylate LHCII directly (Pribil et al., 2010). Therefore, it is believed that TAP38/PPH1 represents the long-sought LHCII phosphatase (Pesaresi et al., 2011; Puthiyaveetil et al., 2012). TAP38/PPH1 is a type 2C serine/threonine protein phosphatase (PP2C) belonging to the metal-dependent protein phosphatases family. TAP38/PPH1 exhibits low sequence identity with the other PP2C members in Arabidopsis and forms a distinct clade with its orthologs from other plant species. The majority of TAP38/PPH1 protein resides in the thylakoid membrane with a predicted membrane-spanning helix at its C-terminal region, while some minor soluble isoforms are also present in the stroma (Shapiguzov et al., 2010).

In addition to TAP38/PPH1, another chloroplast PP2C-type phosphatase named PSII CORE PHOSPHATASE (PBCP) was identified as the one specifically involved in dephosphorylation of PSII core subunits, such as D1, D2, and CP43 (Samol et al., 2012). The function of PBCP is to counteract the action of STN8 kinase (required for the phosphorylation of PSII core subunits) (Bonardi et al., 2005), and it is possibly involved in unfolding of thylakoid membranes during the repair process of PSII after photodamage. Therefore, PBCP and TAP38/PPH1 represent two pivotal chloroplast phosphatases with distinct and specialized activities in the acclimation of the photosynthetic apparatus in response to changes in light conditions. Although the in vivo functions of TAP38/PPH1 and PBCP have been demonstrated, the molecular mechanism underlying the specific recognition between TAP38/PPH1 (or PBCP) and the substrate remains elusive and awaits to be addressed through structural and biochemical approaches. Previously, the structures of other PP2C members suggested that they all share a structurally conserved catalytic core domain bearing a binuclear metal center (Shi, 2009; Barford, 2010). The catalytic reaction mediated by the PP2C members was proposed to occur through the bimolecular nucleophilic substitution (S_N2) mechanism based on the earlier works on human PP2Cα (Das et al., 1996; Jackson et al., 2003). Nevertheless, the details about substrate recognition and catalytic mechanism remain to be revealed through structural studies of a PP2C phosphatase in complex with its natural phosphopeptide/phosphoprotein substrate (Barford, 2010; Dupeux et al., 2011).

To understand the mechanistic details of one key step in state transitions, namely, dephosphorylation of LHCII, we solved the high-resolution crystal structures of the stromal domain of TAP38/PPH1 from Arabidopsis (At-TAP38/PPH1), in both substrate-free and substrate-bound states. Direct and specific interactions between TAP38/PPH1 and the N-terminal phosphorylated polypeptide of Lhcb1 have been observed and are described in detail. Insights into the mechanisms of TAP38/PPH1-pLhcb1 mutual recognition and the catalytic process of TAP38/PPH1 have been obtained.

RESULTS

Substrate Specificity of TAP38/PPH1

To analyze the phosphatase activity and substrate specificity of TAP38/PPH1, we used the native thylakoid phosphoproteins from *pph1-2* plants (lacking endogenous TAP38/PPH1) and various synthetic phosphorylated peptides as the substrates for the purified recombinant stromal domain of Arabidopsis TAP38/PPH1 (At-PPH1-SD, Arg59-Thr351). The results show that the recombinant PPH1-SD protein is highly active toward phosphorylated LHCII (pLHCII) and displayed the fastest kinetics toward pLHCII among the four phosphoproteins detected in protein gel blots (Figures 1A to 1E). Although PPH1-SD is barely active toward pD1, it can also dephosphorylate pD2 and pCP43 in addition to pLHCII, but at much lower efficiencies than toward pLHCII (Figures 1A and 1E). Evidently, TAP38/PPH1 has a strong preference for the pLHCII substrate over the PSII core subunits.

What is the specific feature of pLHCII making it a more favorable substrate for PPH1 than the others? The phosphothreonine (pThr or t) residues of pLhcb1 and pLhcb2 are located at the third position from the N terminus with two basic residues preceding it, such as Arg-Lys-pThr-Val-Ala (pLhcb1) or Arg-Arg-pThr-Val-Lys (pLhcb2) (Vener et al., 2001). In contrast, those of pD1 (pThr-Ala-Ile-Leu-Glu), pD2 (pThr-Ile-Ala-Leu-Gly), and pCP43 (pThr-Leu-Phe-Asn-Gly) are located at the first position of their N termini without the two preceding basic residues. To investigate whether the first two basic residues of pLhcb1/2 play a role in the specific substrate recognition by PPH1, we have performed a phosphatase activity assay on PPH1-SD using synthetic phosphorylated peptide substrates (Figure 1F). The results show again that PPH1-SD exhibits higher activity toward the N-terminal phospho-pentapeptide of pLhcb1/2 than toward those of pD1, pD2 or pCP43. When the first two basic residues of pLhcb1 (RK) are fused to the N-terminal ends of pD1, pD2, and pCP43 phosphopeptides, the enzyme shows enhanced activity toward these chimeric substrates. On the other hand, when the Arg-Lys or Arg-Arg residues are deleted from the pLhcb1/2 phosphopeptides, the activity of

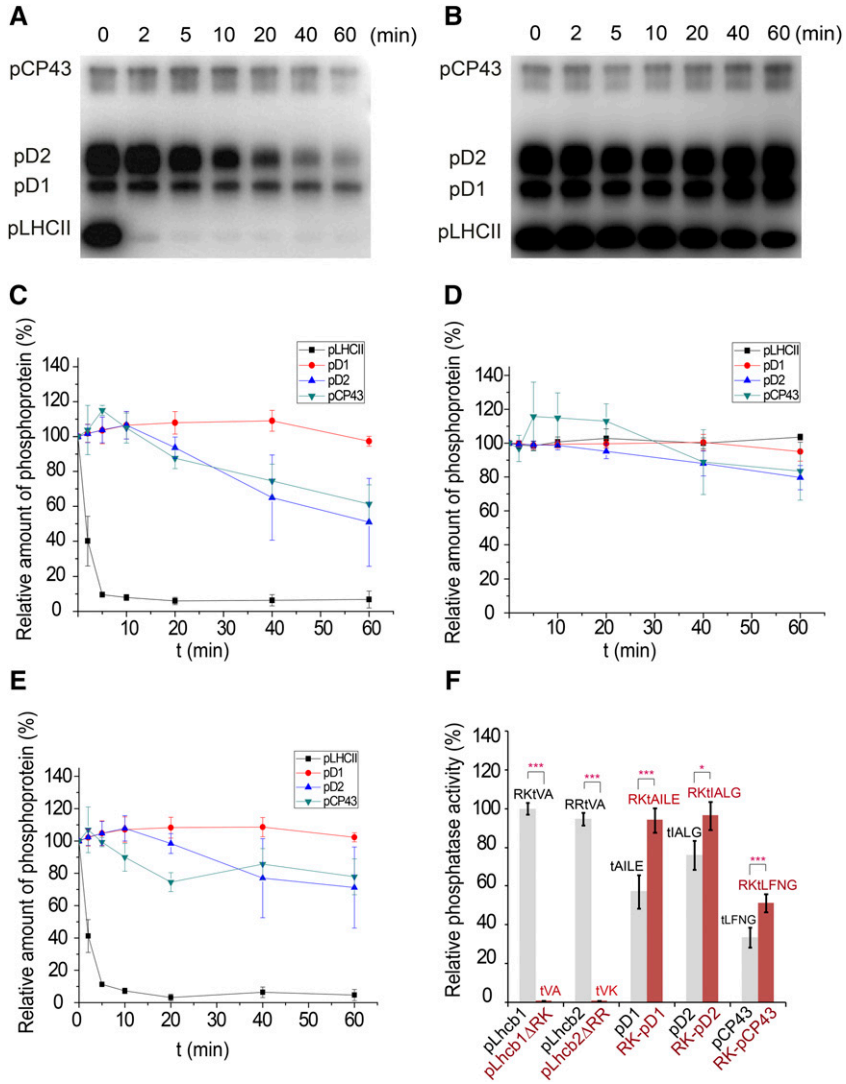

Figure 1. The Phosphatase Activity of Recombinant At-PPH1-SD Protein.

(A) Dephosphorylation of the thylakoid membrane proteins catalyzed by the purified PPH1-SD protein. The phosphoproteins were prepared from the *pph1-2* plant and probed through the protein gel blot using the antiphosphothreonine antibody.

(B) A parallel control experiments were performed without adding the PPH1-SD protein in the reactions. The data show the background activities of endogenous thylakoid phosphatases other than TAP38/PPH1.

(C) The dephosphorylation kinetic curves of pLHCII, pD1, pD2, and pCP43 corresponding to the protein gel blot shown in **(A)**. The intensities of the protein bands at different time points were scaled to those at the beginning of reactions. The relative amount of phosphoproteins = 100% when t = 0 min.

(D) The dephosphorylation kinetic curves corresponding to the control protein gel blot shown in **(B)**.

(E) Processed kinetic curves representing the dephosphorylation activities catalyzed by exogenous PPH1-SD alone. The curves were plotted with the band intensity data subtracting the effects of endogenous phosphatases. The relative amount of phosphoproteins = 100% $-$ ΔI_{PPH1}. ΔI_{PPH1} represents the percentage of band intensity decrease caused by the activity of PPH1-SD. $\Delta I_{PPH1} = [I_0(A) - I_t(A)]/I_0(A) - [I_0(B) - I_t(B)]/I_0(B)$. $I_0(A)$ and $I_0(B)$ represents the 0-min band intensities extracted from A and B panels, respectively. $I_t(A)$ and $I_t(B)$ represents the t-min band intensities extracted from **(A)** and **(B)**, respectively. The error bars in **(C)** to **(E)** indicate the standard errors of the mean values of three independent repeats (n = 3).

(F) The relative phosphatase activities of PPH1-SD toward different synthetic phosphorylated peptides. The activity of PPH1-SD was measured through the phosphatase assay system monitoring the release of free phosphate during the dephosphorylation reaction. Its activity toward pLhcb1 phosphopeptide (RKtVA) was normalized as 100% and the rest data were proportionally scaled. The sequences of the phosphopeptides (pLhcb1, pLhcb1ΔRK, pLhcb2, pLhcb2ΔRR, pD1, RK-pD1, pD2, RK-pD2, pCP43, and RK-pCP43) used for the assay are labeled on top of the columns. The error bars denote the standard errors of the mean values (n = 9 except for pLhcb1ΔRK and pLhcb2ΔRR, whose n = 6; *P < 0.05 and ***P < 0.001 by t test).

PPH1-SD toward the resulting pThr-Val-Ala or pThr-Val-Lys peptides drops to a minimal level (1% of the activity toward Arg-Lys-pThr-Val-Ala) close to the background (Figure 1F). Therefore, the first two basic residues of pLhcb1/2 are crucial for the specificity of TAP38/PPH1 toward pLHCII.

Structure of PPH1 in Substrate-Free State

To understand the structural basis for specific substrate recognition of TAP38/PPH1, we have successfully crystallized the Arabidopsis PPH1-SD and solved its structure at 1.6-Å resolution through x-ray crystallography. The overall structure of PPH1-SD contains a conserved PP2C-type catalytic core domain and two additional motifs (Motifs A and B) (Figures 2A to 2C). The core domain superimposes well with the other members of PP2C-type phosphatases, such as the human PP2Cα (Das et al., 1996) (Supplemental Figure 1A) and the Arabidopsis ABSCISIC ACID INSENSITIVE1 (ABI1) phosphatase involved in the abscisic acid signaling pathway (Supplemental Figure 1B) (Miyazono et al., 2009; Yin et al., 2009), despite the fact that TAP38/PPH1 only shares ~21% sequence identity with either of them. In brief, the core domain of PPH1-SD harbors a central β-sandwich structure containing two sheets of five antiparallel β-strands (Figure 2A). The β-sandwich is flanked by two α-helices on one side and three more α-helices on the other side, leading to the formation of a tightly packed 2α-5β-5β-3α core domain. Motifs A (residues 196 to 237) and B (residues 238 to 277) are consecutively inserted in the loop region between β7 and β8 (Figure 2A). Motif A contains an α-helix followed by a long loop, while motif B folds into a hairpin-like structure with a pair of antiparallel α-helices. Motif B covers the segment of sequence (residues 242 to 274), which is a characteristic feature of the PPH1 family identified through bioinformatics analysis (Shapiguzov et al., 2010). Motifs A and B are not only covalently linked together, but also form extensive noncovalent interactions with each other (Supplemental Figures 2A and 2B). These interactions include the van der Waals interactions (Trp220-Phe262 and Trp220-Ile266) at the buried hydrophobic interface and hydrogen bonds (such as Cys227-Arg257 and Ala217-Lys269) as well as ionic interactions (such as Asp229-Arg240) at the solvent-exposed interface. Furthermore, the two stably associated motifs are attached to the core domain through the loop region between Val-232 and Ile-239 (Supplemental Figure 2C). This loop interacts with the core domain through hydrogen bonds, hydrophobic interactions, and ionic interactions.

At the top region of PPH1-SD, two narrow clefts form between the two β-sheets of the core domain (Cleft 1) and between the two additional motifs (Cleft 2) (Figure 2B). Cleft 1 harbors a Mn^{2+}-Mn^{2+} binuclear center and its bottom surface is highly electronegative. The negative charges carried by the acidic residues (Asp-93, Asp-296, and Asp-339) in this area are neutralized by the two Mn^{2+} ions (Mn1 and Mn2). The PPH1-SD protein used for crystallization was purified in the presence of 2 mM Mn^{2+}, and the x-ray anomalous difference map computed with data collected near the K edge of Mn confirmed that the two metal binding sites in PPH1-SD do contain Mn^{2+} (Supplemental Figure 1C). In the activity assay on the effect of divalent cations (Figure 2D), PPH1 exhibits optimal activity with Mg^{2+} and maintains 40% activity with Mn^{2+}, while the others, like Ca^{2+} or Zn^{2+}, do not support its activity. The binding sites for Mg^{2+} should overlap with those of Mn^{2+} considering that their coordination geometries are both octahedral and either of them supports the catalytic activity of PPH1/TAP38.

The binuclear center structure in PPH1/TAP38 closely resembles the Mn^{2+}-Mn^{2+} center found in the human PP2Cα structure (Das et al., 1996) (Figures 2E and 2F). The two metal ions in PPH1/TAP38 are spaced 3.9 Å apart, almost identical to that of PP2Cα (3.8 Å). Each Mn^{2+} ion is coordinated with the surrounding amino acid residues and water molecules in an octahedral geometry. The metal-oxygen bond lengths range from 2.1 to 2.4 Å. The two metal-O bonds linking the interstitial water molecule to Mn1 and Mn2 are slightly longer than the other 10 bonds (2.3/2.4 Å versus 2.1/2.2 Å). Above the binuclear center, a sulfate ion is located 5.1 and 4.7 Å from Mn1 and Mn2, respectively (Figure 2A). It forms strong ionic interaction with a nearby Arg residue, namely, Arg-69. As the sulfate ion is a close mimic of free phosphate ion (the product of dephosphorylation reaction), it is predictable that the phosphopeptide substrate should also bind around this region.

The Cleft 2 site is located in the region between Motifs A and B and appears shallower than the Cleft 1 site (Figure 2B). The surface of Cleft 2 is nearly neutral or slightly electropositive. These two clefts are connected to each other and form a C-shaped groove on the top surface of PPH1. Such a surface groove appears to be ideal for binding the elongated N-terminal phosphopeptide of Lhcb1 or Lhcb2. Alternatively, it may also provide the binding sites for potential unidentified regulatory factors (either small molecules or proteins).

A Mutant of TAP38/PPH1 in Complex with a Phosphopeptide Substrate

To obtain a stable complex between TAP38/PPH1 and the phosphopeptide substrate, the major obstacle lies in the fact that the active enzyme only binds the substrate transiently and will release the products immediately after catalysis. To tackle this problem, there are two ways of blocking the catalytic process in theory to our knowledge. One way is by designing mutations on the enzyme to abolish or attenuate the dephosphorylation reaction without affecting its substrate binding affinity and the other is using nonhydrolyzable substrate mimics to form a complex with the wild-type enzyme. The attempts to cocrystallize wild-type PPH1-SD with the phosphopeptide mimics, such as Arg-Lys-Asp-Val-Ala and Arg-Lys-Glu-Val-Ala peptides, were unsuccessful, presumably due to their low-affinity binding. On the other hand, we have also chosen the mutagenesis strategy and screened for low-activity mutants of PPH1-SD for cocrystallization with the phosphorylated peptide substrate. A number of single-site mutations on the accessory acidic residues surrounding the binuclear center have been designed (Figure 3A). The purified mutant proteins were tested for their activities toward the Arg-Lys-pThr-Val-Ala phosphopeptide (the N-terminal pentapeptide of pLhcb1). As shown by the results of activity assay (Figure 3B), the D180N and D180E mutants exhibited the most dramatic loss of activity among the six mutants tested. The Asp-180 is located 5.7 Å from Mn1, and the space between them is filled by two water molecules (Figure 3A). The proper size and negative charge of the amino acid side chain on the 180 position of TAP38/PPH1 are both crucial for its optimal catalytic activity.

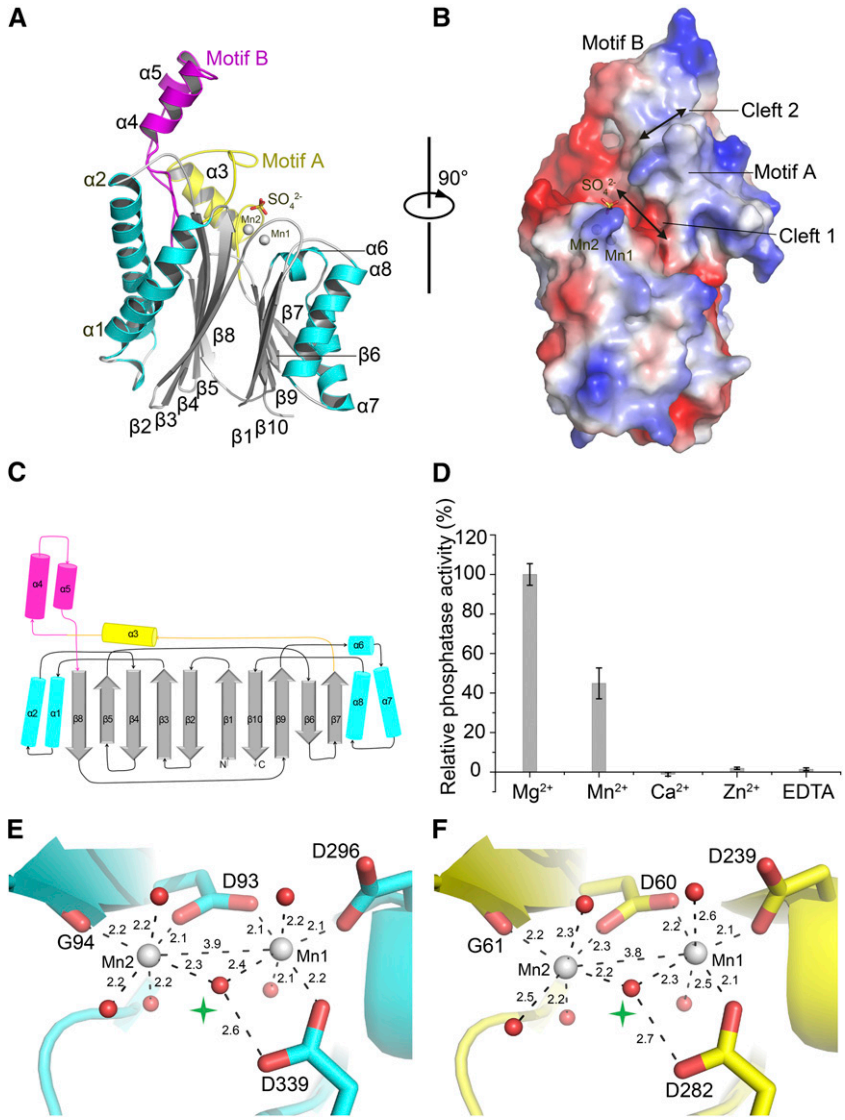

Figure 2. Structure of At-PPH1-SD in the Substrate-Free State.

(A) Side view of the PPH1-SD structure in cartoon presentation. The catalytic core domain is colored in cyan and silver, while two additional motifs (Motif A and Motif B) are colored in yellow and magenta, respectively. The Mn^{2+} ions (Mn1 and Mn2) are shown as silver spheres, and a sulfate ion above the binuclear center is presented as a stick model.

(B) A rotated view of PPH1-SD in the electrostatic-potential surface presentation. The red and blue colors on the surfaces indicate electronegative and electropositive regions, respectively.

(C) Schematic drawing of the topological structure of PPH1-SD. The α-helices are shown as cylinders and the β-strands are presented as block arrows. The color codes are consistent with those in **(A)**.

(D) The relative phosphatase activity of PPH1-SD in the presence of different divalent cations. RKtVA peptide was used as the substrate for the assay. The error bars denote the standard errors of the mean values (n = 3).

(E) and **(F)** Local structures of the binuclear centers of PPH1 **(E)** and human PP2Cα **(F)**. The water molecules are presented as red spheres and the interstitial ones linking Mn1 and Mn2 are indicated by green stars.

To solve the structure of PPH1-SD in complex with the phosphopeptide substrate, the D180E mutant was cocrystallized with a synthetic 15-mer phosphopeptide whose sequence (RKtVAKPKGPSGSPW) is identical to the N-terminal region of pLhcb1. The structure of PPH1-phosphopeptide complex was solved at 2.0-Å resolution and the high-quality electron density map reveals well-defined features for the first 10 amino acid residues including the phosphothreonine residue (RKtVAKPKGP) (Figure 3C). The electron density for the last five C-terminal residues (Ser-Gly-Ser-Pro-Trp) of the phosphopeptide is too weak to be interpretable, suggesting this part is disordered due to its high flexibility.

Previously, the structure of a synthetic 15-mer phosphopeptide of the N-terminal region of pea (*Pisum sativum*) Lhcb1 (sequence: RKSAtTKKVASSGSP) in solution was characterized through various spectroscopic techniques (Nilsson et al., 1997). The backbone of the pea Lhcb1 N-terminal phosphopeptide is mostly extended, while 12% α-helical region and 45% β-structure (β-sheet and β-turn) have also been detected. Unlike the solution-state pea Lhcb1 N-terminal phosphopeptide, the Arabidopsis Lhcb1 N-terminal phosphopeptide bound to TAP38/PPH1 adopts an all-coil secondary structure. Its backbone is also mostly extended, while the region around pThr3 bears a notable bending (Figures 3D to 3F). The difference may arise from the sequence variation between the two phosphopeptides and presumably also from the conformational changes of the phosphopeptide upon binding to TAP38/PPH1.

Interactions between PPH1 and the N-Terminal Phosphopeptide of Lhcb1

As shown in Figure 3C, the N-terminal phosphopeptide of Lhcb1 is accommodated in the elongated C-shaped surface groove of the TAP38/PPH1. The first five-residue segment (Arg-Lys-pThr-Val-Ala) fits well into Cleft 1, while the next five-residue segment (Lys-Pro-Lys-Gly-Pro) binds to Cleft 2. The extended phosphopeptide backbone resembles the C character. It matches nicely with the shape of the surface groove on TAP38/PPH1. There is a 96° bending around $pThr3_{Lhcb1}$ that directs the phosphothreonine toward Mn2 of the binuclear center (Figure 3D). Unexpectedly, the phosphate group of $pThr3_{Lhcb1}$ forms a direct metal-O coordination bond (2.1-Å bond length) with Mn2 by replacing a water molecule (Figures 3E and 5A). This water molecule is coordinated with Mn2 in the substrate-free state. To secure the position of pThr3, $Arg69_{PPH1}$ from the β1-β2 loop interacts with the phosphate group through a strong ionic bond (2.6-Å bond length) (Figure 3E). Moreover, several water molecules surrounding the phosphate group further connect it to the nearby residues through hydrogen bonds.

Besides the interactions involving $pThr3_{Lhcb1}$, additional hydrogen bonds, hydrophobic interactions, and van der Waals interactions have formed between the other regions of the phosphopeptide and the enzyme (Figures 3E and 3F). For instance, the terminal NH_2 group of the phosphopeptide is hydrogen bonded to two water molecules that are attached to $Asp296_{PPH1}$ and $Gln338_{PPH1}$, respectively. The first two basic residues (Arg-1 and Lys-2) of the phosphopeptide are important for enhancing the selectivity, as shown by the activity assay using substrates with or without these two residues (Figure 1F). These two residues not only interact with the enzyme through hydrogen bonds, but they also carry positive charges that will help to attract the peptide toward the negatively charged surface of the binding site (Figure 4D). The side chain of $Arg1_{Lhcb1}$ is recognized by the backbone carbonyl group of $Asn223_{PPH1}$ (Figure 3E). Moreover, this Arg residue also pairs up with the side chain of $Arg225_{PPH1}$, leading to an electrostatic-defying but stable Arg-Arg short-range interaction (Magalhaes et al., 1994; Pednekar et al., 2009). The carbonyl of $Lys2_{Lhcb1}$ is recognized by the side chain of $Arg225_{PPH1}$, and its side chain is linked to $Glu71_{PPH1}$ and $Glu73_{PPH1}$ through the water molecules between them. These specific interactions involving the first two residues of pLhcb1 explain the preference of TAP38/PPH1 for the substrates with two basic residues at their N termini. In the other regions, three additional pairs of hydrogen bonds mediate the interactions between the peptide backbone and the surface residues of the Cleft 1 site of TAP38/PPH1 (Figure 3E). In detail, the carbonyl of $pThr3_{Lhcb1}$ is hydrogen-bonded to the backbone amide of $Ala231_{PPH1}$; the amide of $Val4_{Lhcb1}$ is attached to the backbone carbonyl of $Asp229_{PPH1}$; the carbonyl of $Ala5_{Lhcb1}$ is associated with the side chain carboxyl group of $Asp229_{PPH1}$. These interactions serve to secure the binding of the phosphopeptide within the cavity of the Cleft 1 site.

The second part of the phosphopeptide, namely, the Lys-Pro-Lys-Gly-Pro segment, is bound to the Cleft 2 region through the following specific interactions with TAP38/PPH1 (Figures 3E and 3F). The backbone amide and carbonyl groups of $Lys8_{Lhcb1}$ are hydrogen-bonded to $Cys227_{PPH1}$ and $Trp220_{PPH1}$, respectively (Figure 3E). Moreover, hydrophobic and van der Waals interactions, involving $Pro7_{Lhcb1}$ and $Pro10_{Lhcb1}$ (with $Arg257_{PPH1}$/$Trp258_{PPH1}$ and $Trp220_{PPH1}$/$Phe262_{PPH1}$, respectively), reinforce the binding of this segment in the Cleft 2 region (Figure 3F). The presence of multiple proline or glycine residues in the N-terminal region of Lhcb1 prevents the formation of regular secondary structure (such as α-helix or β-sheet) within this area and endows it the flexibility to adapt into the Cleft 2 site in an extended conformation.

Conformational Changes in PPH1 Induced by Phosphopeptide Binding

By comparing the structures of PPH1 in the substrate-free and substrate-bound states, we have observed significant conformational changes in PPH1 in the regions involved in binding the Lhcb1 phosphopeptide. Three loops around the Cleft 1 site, namely, $loop_{β1-β2}$, $loop_{α3-α4}$, and $loop_{α8-β10}$, are attracted toward the phosphopeptide (Figure 4A). A closer look at these regions reveals that the side chains of Arg-69 (on $loop_{β1-β2}$) and Arg-225 (on $loop_{α3-α4}$) have moved ~1 Å toward the substrate to form an ionic interaction and hydrogen bond with it, respectively (Figure 4B). At the Cleft 2 site, three amino acid residues lining its wall (Arg-257, Trp-258, and Phe-262) have gone through more profound conformational changes induced by substrate binding (Figure 4B). The side chains of Arg-257 and Trp-258 change their conformations through a rotamer-switch mechanism (switching of the amino acid side chain from one conformer to another through rotation around a single bond) to make space for the second part of the phosphopeptide. The aromatic ring of $Phe262_{PPH1}$ is repositioned to avoid steric hindrance with $Pro10_{Lhcb1}$. Moreover, the polypeptide backbone of this region of Motif B (between Lys-250 and Gly-270) shifts outward to open the groove on Cleft 2 for substrate binding (Figure 4A).

The substrate binding sites of TAP38/PPH1 are flexible, allowing them to adopt two different conformations in the substrate-bound and substrate-free states. In the substrate-free state, the Cleft 1 site on TAP38/PPH1 is in a relatively open state (Figure 4C), while the Cleft 2 site appears to be closed (Figure 4E). When the substrate binds, the Arg-69 and Arg-225 residues from the Cleft 1 site close up and lock on the phosphopeptide like a pair of teeth (Figure 4D). In the meantime, the wall-lining residues on the Cleft 2 site move away from the groove and open it to accommodate the remaining part of the phosphopeptide (Figure 4F). The structural evidence presented hereby directly demonstrates that TAP38/PPH1

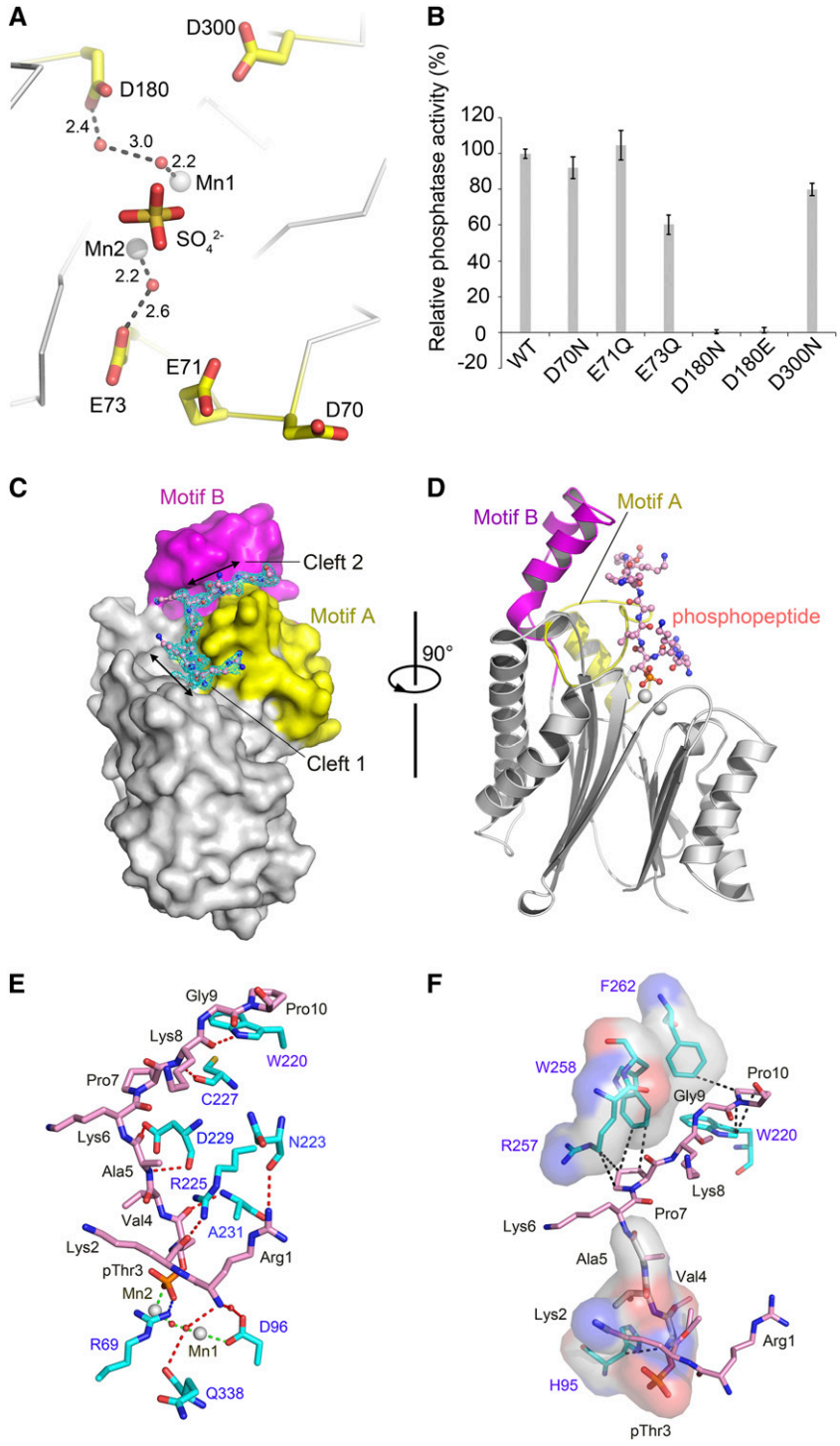

Figure 3. Identification of a Low-Activity Mutant of At-PPH1-SD for Trapping the Phosphopeptide Substrate.

(A) Locations of the five accessory acidic residues around the active site chosen for mutagenesis screening. These residues are not directly involved in coordinating the binuclear center but may participate in the catalytic process. They are positioned nearby the product mimic (sulfate ion). The unit for the numbers labeled nearby the dash lines is Å.

(B) The phosphatase activity of the PPH1-SD mutants relative to the wild-type protein. The error bars denote the standard errors of the mean values (n = 6).

specifically targets pLhcb1 and can respond to the binding of substrate by readjusting the shapes of its binding sites according to the substrate.

DISCUSSION

Structural Mechanism Underlying the Substrate Preference of TAP38/PPH1

Phosphorylation and dephosphorylation of LHCII are fundamental events occurring in the state transition processes in algae and plants (Bennett, 1980; Allen et al., 1981). Despite the long history of research on these phenomena, it is only recently that a chloroplast kinase and phosphatase pair, named Stt7/STN7 and TAP38/PPH1, respectively, have been shown to be essential for the reversible modification of LHCII (Rochaix et al., 2012). By solving and analyzing the structures of TAP38/PPH1 in substrate-free and substrate-bound states, we have provided direct evidence for the specific recognition between TAP38/PPH1 and pLhcb1. The structure of TAP38/PPH1 in complex with the N-terminal phosphopeptide of Lhcb1 demonstrates that the elongated phosphopeptide is recognized by two surface clefts sites on TAP38/PPH1. It curves into a C-shaped molecule fitting well into the grooves of these two connected clefts (Figure 3C). The first part of the phosphopeptide (Arg-Lys-pThr-Val-Ala) docks into the Cleft 1 region in a way resembling the classical lock-and-key mechanism. In the absence of substrate, the Cleft 1 site is already wide open (Figure 4C). Upon binding of Arg-Lys-pThr-Val-Ala peptide in this region, the nearby amino acid residues are attracted toward the peptide and clamp onto the substrate (Figure 4D). On the other hand, the Cleft 2 site in the substrate-free state is occluded by the bulky side chains of Arg-257, Trp-258, and Phe-262 and has a relatively shallow groove on its surface (Figure 4E). Upon binding with the second part (Lys-Pro-Lys-Gly-Pro) of the phosphopeptide, this site deepens its groove to accommodate the peptide (Figure 4F). The change is achieved through repositioning and rotamer-switching of its wall-lining residues (Figure 4B). Therefore, the induced-fit mechanism dominates on the second cleft site of TAP38/PPH1 for the binding and recognition of pLhcb1.

The specificity of TAP38/PPH1 toward pLhcb1 (and presumably also pLhcb2) is achieved through two distinct strategies. First, the two basic residues on the N-terminal region of Lhcb1 (Arg-Lys) enhance the enzyme's activity toward pLhcb1 against pD1, pD2, and pCP43 (Figure 1F). The structure of the TAP38/PPH1-phosphopeptide complex clearly demonstrates that the enzyme's Cleft 1 site matches well with the Arg-Lys residues both in shape and surface charge (Figure 4D). Second, TAP38/PPH1 employs a second binding site, namely, Cleft 2, to further increase its specificity toward pLhcb1 or pLhcb2. This site recognizes and binds the extended part (Lys-Pro-Lys-Gly-Pro) of the Lhcb1 phosphopeptide following the first five residues (Figure 4F). In Lhcb2, the sequence of the corresponding part is Ser-Thr-Pro-Gln-Ser, which may also fit in the Cleft 2 region as it mainly recognizes the main chain of the peptide substrate. As for the unfavorable substrates of TAP38/PPH1, namely, pD1, pD2, and pCP43 of the PSII core, their first 10 residues at the N-terminal regions are tAILERRESE (pD1), tIALGKFTKD (pD2), and tLFNGTLALA (pCP43), respectively. The first five-residue segments and the extended parts (underscored regions) of these PSII core subunits are largely different from those of pLhcb1 (RKtVAKPKGP) and pLhcb2 (RRtVKSTPQS), making them incompatible binding partners in either Cleft 1 or Cleft 2 sites of TAP38/PPH1.

Comparing TAP38/PPH1 with PBCP

The sequence alignment between TAP38/PPH1 and PBCP shows two large insertions present in TAP38/PPH1 but absent in PBCP (Supplemental Figure 2A). These two segments (residues 196 to 230 and 257 to 279) cover most of the Motif A (196 to 237) and part of its Motif B (238 to 277) in the structure of TAP38/PPH1. Motif A is not only involved in binding the Arg-Lys-pThr-Val-Ala part of the phosphopeptide around the Cleft 1 site, but also contributes to the formation of the inner walls of the Cleft 2 site on one side (Figure 3C). The second insertion (257 to 279) belongs to Motif B and overlaps with the previously reported PPH1-characteristic domain identified through phylogenetic comparisons of various PP2C phosphatases (Shapiguzov et al., 2010). This insertion contains the key residues (Arg-257, Trp-258, and Phe-262) lining the inner walls of the Cleft 2 site on the other side across Motif A (Figure 3C). Built with these two additional motifs, TAP38/PPH1 is thereby able to specifically recognize the N-terminal phosphopeptide of pLhcb1 and presumably also that of pLhcb2 considering its high sequence similarity with pLhcb1.

As for PBCP, its Cleft 1 site should be flat and shallow in shape since the two tooth-like Arg residues in PPH1 (Arg-69 and Arg-225 in Figure 4D) are missing in PBCP (Supplemental Figure 2A).

Figure 3. (continued).

(C) Side-view surface presentation of the D180E mutant of PPH1-SD in complex with the 15mer phosphopeptide of pLhcb1. The core domain, Motif A, and Motif B of PPH1-SD are colored in silver, yellow, and magenta, respectively. The $2F_o$-F_c electron density of the phosphopeptide is shown as cyan meshes with $1.0\times\sigma$ contour level.

(D) A rotated view of the complex structure. For clarity, the phosphopeptide is presented as stick model, while the PPH1-SD is shown as cartoon presentation with the same color codes as in **(C)**.

(E) The polar interactions between the phosphopeptide and PPH1-SD. The red, blue, and green dotted lines indicate the hydrogen bonds, ionic bond, and coordinate bond, respectively. The carbon atoms of the amino acid residues from PPH1-SD and the phosphopeptide are colored in cyan and pink, respectively. The water molecules are shown as red spheres.

(F) The hydrophobic and van der Waals interactions between the phosphopeptide and PPH1-SD. The regions on PPH1-SD involved in binding the phosphopeptide through hydrophobic interactions are shown as transparent surface models superposed with stick models. The van der Waals interactions are indicated by the dark dotted lines.

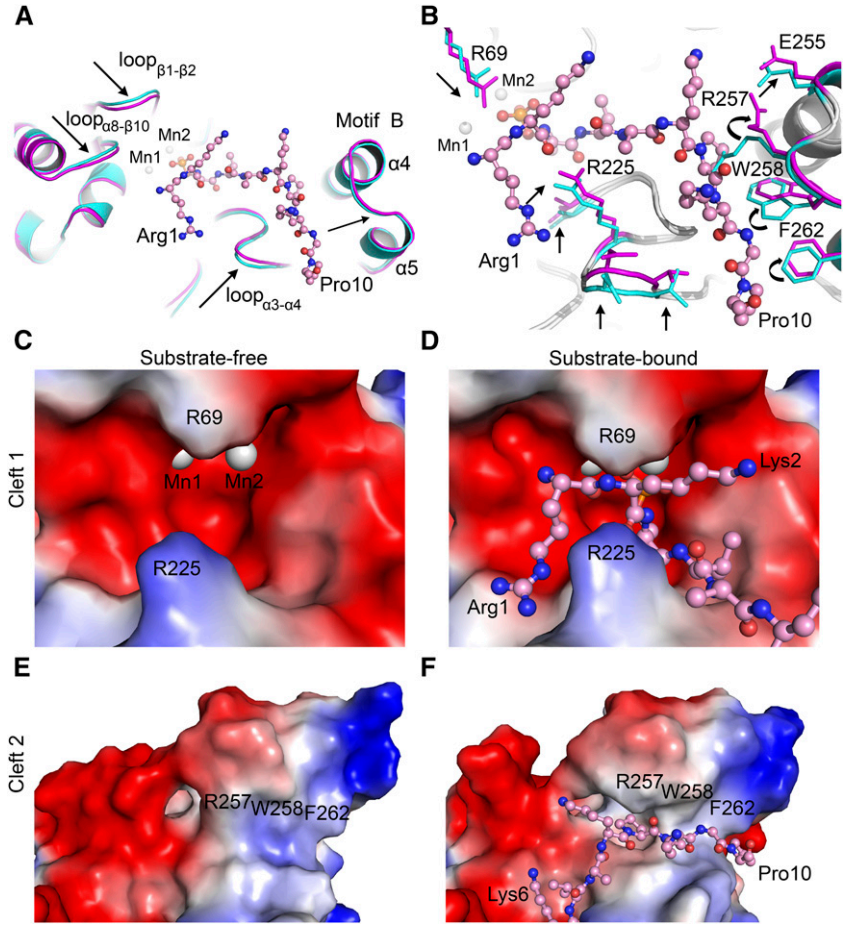

Figure 4. Conformational Changes in At-PPH1-SD Induced by the Binding of Phosphopeptide.

(A) Movement of the loop regions on PPH1-SD induced by phosphopeptide binding. The phosphopeptide is shown in ball-and-stick model, while the PPH1-SD is presented as cartoon models. The substrate-free and substrate-bound states of PPH1-SD are colored in cyan and magenta, respectively. The arrows indicate the movements of polypeptide backbone of PPH1-SD induced by substrate binding.

(B) Rearrangement of the side chains of amino acid residues on PPH1-SD upon substrate binding. The residues with evident conformational changes are indicated by the arrows.

(C) and **(D)** Comparing the shapes of Cleft 1 site on PPH1-SD under substrate-free **(C)** and substrate-bound **(D)** states.

(E) and **(F)** Cleft 2 site under substrate-free **(E)** and substrate-bound **(F)** states. The electrostatic potential surface models of PPH1-SD (zoom-in views on the Cleft 1 or Cleft 2 regions) are presented. Red, electronegative; blue, electropositive.

This feature may be adapted to the N-terminal regions of its preferred substrates, namely, pD1, pD2, and pCP43 (Samol et al., 2012). These PSII core subunits all have phosphothreonine as their first residues at the N termini followed by two or three nonpolar residues. In the absence of Motif A and half of Motif B in PBCP, its Cleft 2 site is most likely also missing. For future study, it remains to be discovered if PBCP contains any additional binding sites for the extended part of the N-terminal peptides of the pD1, pD2, or pCP43 proteins. Moreover, the recombinant PBCP prefers Mn^{2+} and only exhibits minimal activity with Mg^{2+} (Samol et al., 2012), while TAP38/PPH1 can accept both Mg^{2+} and Mn^{2+} for its catalytic activity. Therefore, the active-site structure of PBCP might be somehow different from that of TAP38/PPH1, making it more selective for Mn^{2+} than Mg^{2+}.

Although the in vitro activity assays indicate that recombinant TAP38/PPH1 can also dephosphorylate pD2 and pCP43 (with relatively lower efficiency) in addition to pLHCII (Figures 1A, 1E, and 1F), it is actually highly specific toward pLHCII under in vivo conditions (Pribil et al., 2010; Shapiguzov et al., 2010). In the *pph1* or *tap38* knockout plants, the phosphorylation levels of CP43, D1, and D2 are unaffected (Pribil et al., 2010; Shapiguzov et al., 2010). We suggest that the accessibility of the potential substrates to TAP38/PPH1 under in vivo conditions should be taken into consideration to explain its high specificity toward pLHCII. As TAP38/PPH1 is mainly localized in the membranes of stromal lamellae of thylakoids (Shapiguzov et al., 2010), it can readily access the pLHCIIs associated at the peripheral regions of PSI-LHCI supercomplexes (also localized in the stromal lamellae; Supplemental Figure 2D). On the other hand, the PSII complexes are mainly

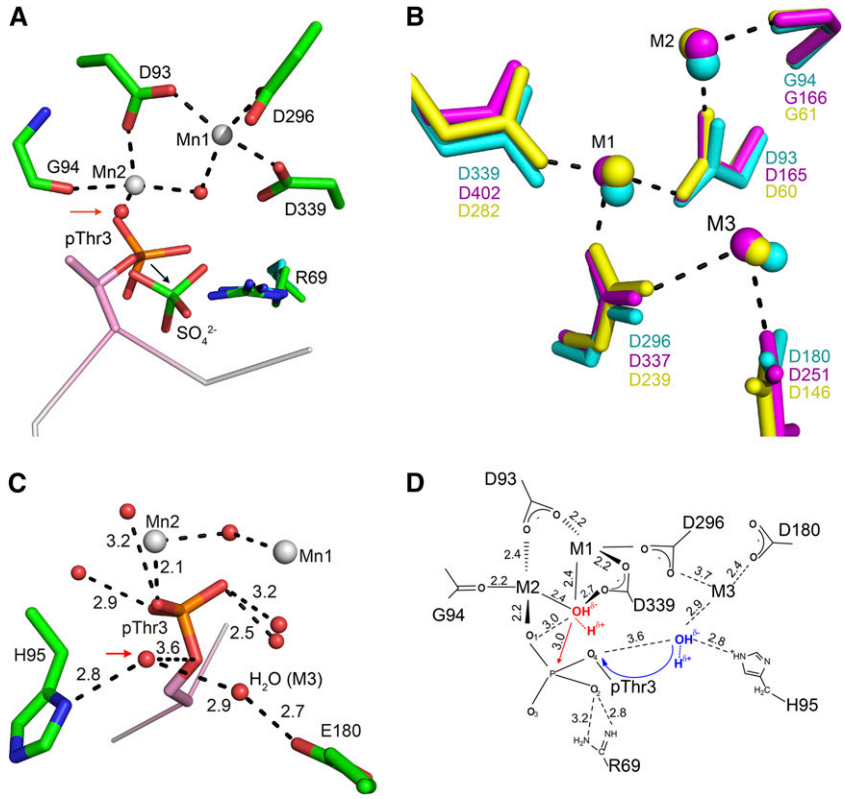

Figure 5. Structural Bases for the Catalytic Mechanism of TAP38/PPH1.

(A) The relative position of the pThr3 residue on the substrate with respect to the product mimic (sulfate ion). The Mn2-coordinating water molecule being replaced by the phosphate group of pThr is indicated by the red arrow. The local structure around the binuclear center of PPH1-SD is shown. The dark dash lines indicate the coordinate bonds around Mn1 and Mn2.

(B) Superposition of the local structures of the active sites of PPH1 (cyan), ABI2 (magenta), and human PP2Cα (yellow). The potential third metal site is labeled as "M3."

(C) The water molecules and amino acid residues potentially involved in protonation of the leaving group on pThr3 residue. The putative proton donor for the leaving group is indicated by the arrow.

(D) A schematic presentation of the proposed reaction mechanism catalyzed by PPH1/TAP38. The water molecule proposed to launch nucleophilic attack on P-O bond is colored in red. The other water molecule donating a proton for the leaving group of pThr3 is colored in blue. The unit for the numbers labeled near the noncovalent bonds is Å.

localized in the tightly stacked grana thylakoid membranes and surrounded by the minor and major light-harvesting complexes (LHCII) at their peripheral regions (Dekker and Boekema, 2005; Daum and Kühlbrandt, 2011). Thus, it would be fairly unlikely for the membrane-bound PPH1 to migrate laterally into the grana region and breach multiple layers of LHCIIs to access the PSII core in order to dephosphorylate the pD1, pD2, and pCP43 subunits. To dephosphorylate the PSII core subunits, plants use a specialized phosphatase, namely, PBCP, which is a soluble protein enriched in the stroma that can readily diffuse to the grana region to gain access to the PSII core subunits (Samol et al., 2012).

Implication for the Features of Substrate Binding Site in Stt7/STN7

A pair of tooth-like Arg residues located on the wall of the Cleft 1 site of TAP38/PPH1, namely, Arg-225 and Arg-69, are directly involved in binding the first three Arg-Lys-pThr residues of pLhcb1 (Supplemental Figure 3A). Curiously, it appears that the *Micromonas* Stt7 kinase homolog (Guo et al., 2013) also contains a pair of basic residues (Lys-320 and Arg-364) near its active site (Supplemental Figure 3B). They correspond to Lys-304 and Arg-344 in Chlamydomonas Stt7 (Cr-Stt7) or Lys-281 and Arg-321 in Arabidopsis STN7 (At-STN7), respectively. Moreover, the Stt7/STN7 kinases harbor an additional hairpin motif that contributes to the formation of the putative Cleft 2 region of the kinases (Supplemental Figure 3B). Although more evidence is needed to demonstrate that Stt7 or STN7 phosphorylates LHCII directly, these observations suggest that the potential peptide binding sites of the state transition kinases may have features resembling those of TAP38/PPH1.

Insights into the Catalytic Mechanism of PP2C Phosphatases

Besides the two high-affinity metal binding sites (M1 and M2) at the binuclear center, various other PP2C members have been reported to contain a third metal binding site (M3) nearby the

Table 1. Data Collection and Structure Refinement Statistics

Data Collection	At-PPH1-SD	D180E-Phosphopeptide
Wavelength (Å)	1.0000	1.0000
Space group	$P2_1$	$P2_12_12_1$
Cell dimensions		
a, b, c (Å)	55.40, 73.88, 82.00	60.29, 64.08, 100.08
α, β, γ (°)	90.00, 109.14, 90.00	90.00, 90.00, 90.00
Resolution (Å)	50–1.60 (1.66–1.60)[a]	50–2.0 (2.07–2.0)*
Completeness (%)	98.4 (97.2)[a]	99.8 (100)[a]
Redundancy	3.8 (3.7)[a]	5.4 (5.5)[a]
I/σ(I)	22.0 (2.6)[a]	14.3 (4.6)[a]
R_{merge}	0.053 (0.493)[a]	0.123 (0.482)[a]
Refinement		
Resolution (Å)	30–1.60	40–2.0
R_{work}/R_{free}	0.180/0.203	0.193/0.217
No. of Atoms		
All	5177	2603
Protein	4589	2371
Ligand/ion	14	2
Water	574	230
B-factor (Å^2)		
All	32.4	24.8
Protein	31.4	24.3
Ligand/ion	38.8	11.0
Water	40.2	30.0
Root mean square deviations		
Bond lengths (Å)	0.017	0.012
Bond angles (°)	1.690	1.835
Ramachandran plot (%)[b]	97.8, 2.2, 0	97.0, 3.0, 0

[a]Numbers in parentheses represent the values for the highest resolution shell.

[b]Percentage of amino acid residues in most favored, allowed, and outlier regions in the Ramachandran plot (showing the relationship between the two backbone dihedral angles of each amino acid residue in the protein structure).

binuclear center (Pullen et al., 2004; Rantanen et al., 2007; Wehenkel et al., 2007; Soon et al., 2012) (Figure 5B). This additional low-affinity Mg^{2+}/Mn^{2+} binding site is required for their catalytic activity and involves two conserved Asp residues as well as an accessory Asp residue (Su et al., 2011; Tanoue et al., 2013). In the crystal structure of wild-type TAP38/PPH1 or D180E mutant in complex with phosphopeptide, a water molecule instead of a Mn^{2+} or Mg^{2+} ion occupies the potential M3 site (Supplemental Figure 4). This water molecule is hydrogen bonded to the Asp-180 (or Glu-180 in the D180E mutant) of TAP38/PPH1 and nearly overlaps with the third metal ion found in ABI2 phosphatase (Figure 5B). The lack of a third metal ion at this site is mainly because its binding site is affected either by a lysine residue (Lys-261) from the adjacent molecule in the crystal lattice (Supplemental Figure 4A) or by the D180E mutation. Lys-261 forms stable ionic interactions with the carboxyl group of Asp-180 and may prevent the third metal ion from binding to this area (Supplemental Figure 4A). In solution, this site should be open to accept the third Mg^{2+}/Mn^{2+} ion and is likely a dynamic metal binding site that is in constant exchange with solvent molecules.

The potential M3 binding site in PPH1 is contributed by D180 and D296, which are highly conserved among various PP2C members (Figure 5B). Mutation of Asp-180 in PPH1 (corresponding to D146 in human PP2Cα) to Glu or Asn residues causes a dramatic loss of activity toward phosphorylated peptide substrate (Figure 3B). In the D180E mutant, the third metal ion might not be positioned accurately near the binuclear center or the substrate pThr group. This may lead to the attenuation of enzyme activity by blocking a key step of the catalytic reaction, namely, protonation of the leaving group on the Thr residue of the substrate. Previous mutagenesis and kinetics study on human PP2Cα strongly suggested that His-62 (His-95 in PPH1) is involved in protonation of the leaving group (Thr residue) oxygen atom during catalytic reaction (Jackson et al., 2003). The H62Q mutation causes a significant loss of activity on human PP2Cα with 20-fold lower K_{cat} value compared with the wild-type enzyme. In the crystal structure of PPH1, the potential M3 site and His-95 are connected by a water molecule (Figure 5C), suggesting that His-95, D180, and the third metal may cooperate to protonate the leaving group of Thr residue on the peptide substrate.

The hydrolysis of a phosphoester bond catalyzed by the PP2C phosphatases involves the S_N2 reaction mechanism (Das et al., 1996). Based on the latest knowledge of the catalytic mechanism of human PP2Cα (Jackson et al., 2003; Tanoue et al., 2013) and the crystal structures of TAP38/PPH1 in substrate-free and substrate-bound states, we hereby propose an updated model on the catalytic mechanism of PP2C (Figure 5D). (1) First, Mn2 binds directly to the phosphate group of phosphothreonine through a coordination bond (2.1 Å). An oxygen atom (O1) of the phosphate

group of phosphothreonine replaces the water molecule that was coordinated to Mn2 in the substrate-free state (Figure 5A). The strong and direct interaction between Mn2 and the phosphate group draws electrons toward the metal ion and polarizes the phosphorus atom of the phosphate group, making it more electrophilic. In order for Mn2 to coordinate the phosphate group of pThr3 directly, the backbone of the phosphopeptide bends around the pThr residue to direct the phosphate group toward Mn2. The accurate positioning of phosphate above Mn2 is further reinforced by $Arg69_{PPH1}$ through ionic interaction. (2) The water molecule bridging metals 1 and 2 was proposed to be a nucleophile in the reaction. It is activated by a nearby conserved Asp residue (D339 in At-PPH1 or D282 in human PP2Cα) (Das et al., 1996). When activated by Asp-339 in At-TAP38/PPH1, this interstitial water molecule between Mn1 and Mn2 will deprotonate into hydroxide ion (OH^-) and launch nucleophilic attack on the phosphorus atom of the phosphate group. (3) Hydrolysis of the P-O monoester bond will require protonation of the leaving group, which may be facilitated by nearby amino acid residues, such as H95 in At-PPH1 (or H62 in human PP2Cα) (Jackson et al., 2003). In the high-resolution structure of TAP38/PPH1 in complex with Lhcb1 phosphopeptide, His-95 is located 4.9 Å from the P-O bond without forming direct interaction with it. Instead, His-95 interacts indirectly with the P-O bond through a bridging water molecule (Figure 5C). Therefore, the proton is most likely relayed by this water molecule to the leaving group oxygen atom. The role of the third metal ion in this region may facilitate the process by coordinating the local residues and/or assist the deprotonation of proton donors, such as His-95 or Asp-180. (4) After hydrolysis, the free phosphate ion will be pulled away from the Mn2 binding site by $Arg69_{PPH1}$ to a nearby position, as indicated by its mimic, namely, the sulfate ion shown in Figure 5A. The previous structure of human PP2Cα contains a free phosphate ion within its active site (Das et al., 1996), resembling our substrate-free TAP38/PPH1 structure with a sulfate ion bound in the corresponding position of free phosphate. This state likely represents the posthydrolysis state with dephosphorylated protein/peptide released but the other product (phosphate) still bound.

METHODS

Protein Expression and Purification

The full-length *Arabidopsis thaliana* PPH1 (at4g27800.1, the first isoform) contains a large soluble domain at the stromal side followed by a single transmembrane helix at its C-terminal region (residues 367 to 385). The recombinant full-length PPH1 protein carrying an N-terminal hexahistidine tag is found mainly in the inclusion bodies when it is expressed in the BL21 (DE3) *Escherichia coli* strain using the pET15b vector. The gene is inserted between the *Nde*I and *Bam*HI sites of the vector. The expressed protein can be purified in the denatured form and then refolded according to the protocol reported by Pribil et al. (2010). The refolded full-length PPH1 appears to be fairly pure and soluble in non-ionic detergent (octyl-glucopyranoside), but only exhibits weak phosphatase activity toward the Arg-Lys-pThr-Val-Ala phosphopeptide substrate (Supplemental Figure 5).

After several rounds of construct optimization through serial truncations from the N and/or C termini, we have found that the region covering most of the stromal domain of PPH1 (Arg59-Thr351, PPH1-SD) expresses as a soluble form (instead of inclusion bodies) in the cytosol of *E. coli*. The cDNA encoding the PPH1-SD protein was ligated into pET24a vector between *Nde*I and *Xho*I sites. Mutants were generated through the QuikChange site-directed mutagenesis protocol (Agilent Technologies). The recombinant PPH1-SD protein expressed in the BL21(DE3) strain was fused with a hexahistidine tag at its C terminus. The *E. coli* cells were cultured in LB-kanamycin medium supplemented with 2 mM $MnCl_2$. When the cell density reached OD_{600nm} = 0.8 to 1.0, protein expression was induced with 0.1 mM isopropyl β-D-1-thiogalactopyranoside for 16 to 20 h at 16°C. After expression, the cells were harvested by centrifugation at 4700*g* for 10 min, and the pellets were frozen and stored at −20°C. Subsequently, the cell pellets were resuspended in a lysis buffer containing 20 mM Tris-HCl, pH 8.0, 500 mM NaCl, and 5% glycerol, and then sonicated on a Misonix S-4000 sonicator. The cell lysate was centrifuged at 48,900*g* for 30 min and the supernatant was collected. The protein was purified through a Ni-NTA metal affinity column (Qiagen). The buffers used for Ni-NTA column purification were as follows. Equilibration buffer contained 20 mM Tris-HCl, pH 8.0, 500 mM NaCl, and 5% glycerol. Washing buffer contained 20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 5% glycerol, and 50 mM imidazole. Elution buffer contained 20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 5% glycerol, and 200 mM imidazole. After being eluted from the Ni-NTA column and concentrated to ~10 mg/mL, the protein sample was further purified through gel filtration chromatography using a Superdex 200 10/300 GL column (GE Healthcare) in a buffer containing 20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 5% glycerol, 2 mM $MnCl_2$, and 0.2% β-mercaptoethanol. The highly purified PPH1-SD protein shows much higher (20 times) phosphatase activity than the refolded full-length protein (Supplemental Figure 5). The protein was concentrated to 10 to 11 mg/mL with A_{260}/A_{280} = 0.5 to 0.7 before crystallization. The 15mer phosphopeptide of Lhcb1 (sequence: RKtVAKPKGPSGSPW, "t" indicates the phosphothreonine residue) and all other peptides were obtained through chemical synthesis by a local vendor (ChinaPeptides).

Protein Crystallization

All crystals were grown at 16°C using the sitting-drop vapor diffusion method (McPherson, 2004). The substrate-free PPH1-SD crystals were grown by mixing 0.5 μL of 10 mg/mL protein with 0.5 μL well solution containing 0.1 M Bis-Tris, pH 6.5, 0.4 M Li_2SO_4, and 30% polyethylene glycol (PEG) 3350. For cocrystallization of PPH1-SD (D180E) mutant with the 15mer Lhcb1 phosphopeptide, the protein was firstly mixed with the synthetic 15mer phosphopeptide (dissolved in double-distilled water) at a molar ratio of 1:10 (protein:phosphopeptide). Subsequently, 0.5 μL of the well solution with 0.2 M sodium malonate, pH 6.0, and 20% PEG 3350 was added to 0.5 μL of the protein-phosphopeptide mixture in sitting drops. The drops were then equilibrated against 100 μL well solution in the sealed chambers of 48-well plates (XtalQuest). Crystals with the morphology of a thin plate usually appeared 2 d after the drops were set up and matured to 0.2 mm × 0.1 mm × 0.02 mm (length × width × thickness) for 1 week. For cryoprotection, the substrate-free PPH1-SD crystals were quickly soaked in a solution of 0.1 M Bis-Tris, pH 6.5, 0.4 M Li_2SO_4, 32% PEG 3350, and 15% glycerol for 1 min before being frozen in liquid nitrogen. The crystals of PPH1-SD (D180E)-phosphopeptide complex were cryoprotected in a solution of 0.2 M sodium malonate, pH 6.0, 22% PEG 3350, 1.5 mM phosphopeptide, and 15% glycerol.

Data Collection, Phasing, and Structure Refinement

The x-ray diffraction data of PPH1-SD and PPH1-SD (D180E)-15mer phosphopeptide complex were collected at Photon Factory (beamline NW12A) and the Shanghai Synchrotron Radiation Facility (beamline BL17U), respectively. The NW12A at Photon Factory is equipped with an ADSC Quantum 210 detector, while the BL17U at the Shanghai Synchrotron Radiation Facility uses an ADSC Q315 CCD detector. The data were both

processed with the HKL2000 program. To solve the phases of the wild-type PPH1-SD structure, automatic molecular replacement (AutoMR) was run in the PHENIX suite (Adams et al., 2010) using a search model of human protein phosphatase 1A (Protein Data Bank ID: 3FXJ). The search yielded a clear solution with a Z-score of 12.6 and a log likelihood gain of 205 and an output model with two molecules per asymmetric unit. Subsequently, the phenix.autobuild program was used to rebuild the structural model according to the input sequence of PPH1. The initial model (R_{work} = 36.9%/R_{free} = 38.5%) was manually corrected using the Coot program (Emsley and Cowtan, 2004) and subjected to further refinement using the phenix.refine program. The final model of both molecules in the asymmetric unit covers all residues of PPH1-SD (Arg-59 to Thr-351 of PPH1). Two manganese ions, one sulfate ion, and numerous water molecules have been located in each PPH1-SD monomer.

The structure of the PPH1-SD(D180E)-phosphopeptide complex was also solved through molecular replacement using the wild-type PPH1-SD structure as the input model. The top solution has a Z-score of 48.8 and a log likelihood gain of 2764. The model was rebuilt by the phenix.autobuild program, manually corrected using Coot, and further refined using the phenix.refine program. Each asymmetric unit contains one copy of the PPH1-SD(D180E) molecule and one molecule of Lhcb1 phosphopeptide continuously traced from Arg-1 to Pro-10. The statistics of data collection and structure refinement are listed in Table 1. The sequence alignment shown in Supplemental Figure 2A was generated using the ClustalW program (McWilliam et al., 2013) and readjusted manually.

Phosphatase Activity Assay

The Promega Ser/Thr phosphatase assay system was used to analyze the activity of PPH1-SD in dephosphorylating the phosphopeptide substrates. Each reaction was performed in a 50-μL sample containing 50 mM imidazole, pH 7.2, 5 mM $MgCl_2$, 0.2 mM EGTA, 0.1 mg/mL BSA, and 500 μM synthetic phosphorylated peptide. To start the reaction, 1 μg wild-type or mutant PPH1-SD was added to the system. After incubation at 30°C for 20 min, the reaction was stopped by adding 50 μL molybdate dye solution provided by the assay kit. The absorbance at 600 nm (A_{600}) was measured on a 96-well plate reader (Thermo Varioskan Flash 3100). For the metal ion-dependent activity assay, the PPH1-SD protein was stripped off the endogenous metal ion with EDTA, purified through a gel-filtration column, and then used for the phosphatase activity assay supplemented with different divalent cations (Mg^{2+}, Mn^{2+}, Ca^{2+}, or Zn^{2+}).

Furthermore, seeds of the Arabidopsis *pph1-2* (SAIL_514-C03) mutant were obtained from the European Arabidopsis Stock Center (NASC) and used for preparing native thylakoid phosphoproteins in order to characterize the activity of recombinant PPH1-SD toward native substrates. The plants were grown in a growth chamber with 12 h light and 12 h dark at 22°C for 3 to 4 weeks. The thylakoids were prepared according to a published protocol (Casazza et al., 2001) and frozen at −80°C until use. Before the assay, the thylakoid sample was solubilized in 1% β-dodecyl-maltoside (DDM) to a final concentration of chlorophyll at 0.5 mg/mL. Each reaction was performed in a 500-μL system containing 50 mM imidazole, pH 7.2, 5 mM $MgCl_2$, 0.2 mM EGTA, 0.1 mg/mL BSA, and dodecyl-maltoside-solubilized thylakoid proteins with 0.1 μg/μL chlorophyll. To initiate the PPH1-catalyzed dephosphorylation reaction, 2.5 μL purified recombinant PPH1-PD at a concentration of 1 μg/μL was added to the system and then incubated at 30°C. Samples were taken after the reaction continued for 0, 2, 5, 10, 20, 40, and 60 min. To stop the reaction, the samples were mixed immediately with SDS loading buffer. The proteins in the samples were resolved by 15% SDS-PAGE with 6 M urea, and each lane was loaded with 0.5 μg chlorophyll. The SDS-PAGE was run at 100 V overnight to separate the bands of LHCII, D1, D2, and CP43. After electrophoresis, the proteins on the gel were transferred to polyvinylidene fluoride membrane (presoaked with 100% methanol) under 250-mA constant current for 150 min. The membrane was blocked in 8% nonfat milk overnight at 4°C. The phosphothreonine antibody from Cell Signaling was used in 1:20,000 dilution to probe the phosphorylated thylakoid proteins on the membrane. After being incubated with the secondary antibody conjugated with horseradish peroxidase, the membrane was developed by adding the western lightning Ultra ECL horseradish peroxidase substrate (Perkin-Elmer), and the images were taken on a chemiluminescence CCD imaging system (ChemiScope 3500 mini imager; Clinx Science Instruments). The band intensity was measured using ImageJ software.

Accession Numbers

The protein sequence data from this article can be found in the Uniprot databank (www.uniprot.org) under the following accession numbers: P49599 (Arabidopsis TAP38/PPH1), O64730 (Arabidopsis PBCP), P0CJ48 (Arabidopsis Lhcb1), and Q9SHR7 (Arabidopsis Lhcb2). The Protein Data Bank accession codes for the structures of Arabidopsis ABI1 and ABI2, human PP2Cα, *Micromonas* Stt7, and the cAMP-dependent protein kinase in complex with substrates are 3JRQ, 3UJK, 1A6Q, 4IX4, and 1JBP, respectively. Coordinates and structure factors for At-PPH1-SD (PDB: 4YZG) and At-PPH1-SD(D180E)-pLhcb1 phosphopeptide complex (PDB: 4YZH) have been deposited in the Protein Data Bank.

Supplemental Data

Supplemental Figure 1. Characterization of the core domain structure of At-PPH1-SD.

Supplemental Figure 2. The factors related to the distinct substrate specificities of PPH1 and PBCP.

Supplemental Figure 3. Comparing the peptide binding sites of TAP38/PPH1 and a Stt7 kinase homolog.

Supplemental Figure 4. The potential third metal ion binding sites in the wild type and D180E mutant of At-PPH1-SD.

Supplemental Figure 5. Comparing the refolded full-length At-PPH1 protein with the At-PPH1-SD protein.

ACKNOWLEDGMENTS

We thank J.D. Rochaix, W.R. Chang, and M.R. Fan for discussion; X.Y. Liu and X.B. Liang for technical support in biochemistry and for handling protein crystal samples for data collection; the staff at Shanghai Synchrotron Radiation Facility and Photon Factory for their technical support during x-ray diffraction data collection; and the European Arabidopsis Stock Center (NASC) for providing the Arabidopsis mutant seeds. The project is financially supported by National 973 Project Grants 2011CBA00903 and the Strategic Priority Research Program of CAS (XDB08020302). Z.F.L. is supported by the National Thousand Young Talents Program from the Office of Global Experts Recruitment in China.

AUTHOR CONTRIBUTIONS

X.P.W. carried out the molecular biology, biochemistry, data collection, and structure determination and analysis. J.T.G. constructed the initial expression vectors of PPH1 and examined protein levels. M.L. contributed to plant growth, biochemistry, and structure analysis. Z.F.L. designed and coordinated the project. X.P.W. and Z.F.L. wrote the article.

Received February 3, 2015; revised March 26, 2015; accepted April 5, 2015; published April 17, 2015.

REFERENCES

Adams, P.D., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. **66:** 213–221.

Allen, J.F. (1992). Protein phosphorylation in regulation of photosynthesis. Biochim. Biophys. Acta **1098:** 275–335.

Allen, J.F., Bennett, J., Steinback, K.E., and Arntzen, C.J. (1981). Chloroplast protein phosphorylation couples plastoquinone redox state to distribution of excitation energy between photosystems. Nature **291:** 25–29.

Barford, D. (2010). The structure and topology of protein serine/threonine phosphatases. In Handbook of Cell Signaling, R.A. Bradshaw and E.A. Dennis, eds (San Diego, CA: Academic Press), pp. 677–681.

Bellafiore, S., Barneche, F., Peltier, G., and Rochaix, J.-D. (2005). State transitions and light adaptation require chloroplast thylakoid protein kinase STN7. Nature **433:** 892–895.

Bennett, J. (1980). Chloroplast phosphoproteins. Evidence for a thylakoid-bound phosphoprotein phosphatase. Eur. J. Biochem. **104:** 85–89.

Bonardi, V., Pesaresi, P., Becker, T., Schleiff, E., Wagner, R., Pfannschmidt, T., Jahns, P., and Leister, D. (2005). Photosystem II core phosphorylation and photosynthetic acclimation require two different protein kinases. Nature **437:** 1179–1182.

Bonaventura, C., and Myers, J. (1969). Fluorescence and oxygen evolution from *Chlorella pyrenoidosa.* Biochim. Biophys. Acta **189:** 366–383.

Casazza, A.P., Tarantino, D., and Soave, C. (2001). Preparation and functional characterization of thylakoids from *Arabidopsis thaliana.* Photosynth. Res. **68:** 175–180.

Das, A.K., Helps, N.R., Cohen, P.T., and Barford, D. (1996). Crystal structure of the protein serine/threonine phosphatase 2C at 2.0 A resolution. EMBO J. **15:** 6798–6809.

Daum, B., and Kühlbrandt, W. (2011). Electron tomography of plant thylakoid membranes. J. Exp. Bot. **62:** 2393–2402.

Dekker, J.P., and Boekema, E.J. (2005). Supramolecular organization of thylakoid membrane proteins in green plants. Biochim. Biophys. Acta **1706:** 12–39.

Depège, N., Bellafiore, S., and Rochaix, J.D. (2003). Role of chloroplast protein kinase Stt7 in LHCII phosphorylation and state transition in *Chlamydomonas.* Science **299:** 1572–1575.

Dupeux, F., Antoni, R., Betz, K., Santiago, J., Gonzalez-Guzman, M., Rodriguez, L., Rubio, S., Park, S.-Y., Cutler, S.R., Rodriguez, P.L., and Márquez, J.A. (2011). Modulation of abscisic acid signaling *in vivo* by an engineered receptor-insensitive protein phosphatase type 2C allele. Plant Physiol. **156:** 106–116.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. **60:** 2126–2132.

Guo, J., Wei, X., Li, M., Pan, X., Chang, W., and Liu, Z. (2013). Structure of the catalytic domain of a state transition kinase homolog from *Micromonas* algae. Protein Cell **4:** 607–619.

Hammer, M.F., Sarath, G., and Markwell, J. (1995). Dephosphorylation of the thylakoid membrane light-harvesting complex-II by a stromal protein phosphatase. Photosynth. Res. **45:** 195–201.

Hammer, M.F., Markwell, J., and Sarath, G. (1997). Purification of a protein phosphatase from chloroplast stroma capable of dephosphorylating the light-harvesting complex-II. Plant Physiol. **113:** 227–233.

Jackson, M.D., Fjeld, C.C., and Denu, J.M. (2003). Probing the function of conserved residues in the serine/threonine phosphatase PP2Calpha. Biochemistry **42:** 8513–8521.

Magalhaes, A., Maigret, B., Hoflack, J., Gomes, J.N., and Scheraga, H.A. (1994). Contribution of unusual arginine-arginine short-range interactions to stabilization and recognition in proteins. J. Protein Chem. **13:** 195–215.

McPherson, A. (2004). Introduction to protein crystallization. Methods **34:** 254–265.

McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y.M., Buso, N., Cowley, A.P., and Lopez, R. (2013). Analysis tool web services from the EMBL-EBI. Nucleic Acids Res. **41:** W597–W600.

Miyazono, K., et al. (2009). Structural basis of abscisic acid signalling. Nature **462:** 609–614.

Murata, N. (1969a). Control of excitation transfer in photosynthesis. I. Light-induced change of chlorophyll a fluorescence in *Porphyridium cruentum.* Biochim. Biophys. Acta **172:** 242–251.

Murata, N. (1969b). Control of excitation transfer in photosynthesis. II. Magnesium ion-dependent distribution of excitation energy between two pigment systems in spinach chloroplasts. Biochim. Biophys. Acta **189:** 171–181.

Nilsson, A., Stys, D., Drakenberg, T., Spangfort, M.D., Forsén, S., and Allen, J.F. (1997). Phosphorylation controls the three-dimensional structure of plant light harvesting complex II. J. Biol. Chem. **272:** 18350–18357.

Pednekar, D., Tendulkar, A., and Durani, S. (2009). Electrostatics-defying interaction between arginine termini as a thermodynamic driving force in protein-protein interaction. Proteins **74:** 155–163.

Pesaresi, P., Pribil, M., Wunder, T., and Leister, D. (2011). Dynamics of reversible protein phosphorylation in thylakoids of flowering plants: the roles of STN7, STN8 and TAP38. Biochim. Biophys. Acta **1807:** 887–896.

Pietrzykowska, M., Suorsa, M., Semchonok, D.A., Tikkanen, M., Boekema, E.J., Aro, E.-M., and Jansson, S. (2014). The light-harvesting chlorophyll *a/b* binding proteins Lhcb1 and Lhcb2 play complementary roles during state transitions in *Arabidopsis.* Plant Cell **26:** 3646–3660.

Pribil, M., Pesaresi, P., Hertle, A., Barbato, R., and Leister, D. (2010). Role of plastid protein phosphatase TAP38 in LHCII dephosphorylation and thylakoid electron flow. PLoS Biol. **8:** e1000288.

Pullen, K.E., Ng, H.L., Sung, P.Y., Good, M.C., Smith, S.M., and Alber, T. (2004). An alternate conformation and a third metal in PstP/Ppp, the *M. tuberculosis* PP2C-Family Ser/Thr protein phosphatase. Structure **12:** 1947–1954.

Puthiyaveetil, S., Ibrahim, I.M., and Allen, J.F. (2012). Oxidation-reduction signalling components in regulatory pathways of state transitions and photosystem stoichiometry adjustment in chloroplasts. Plant Cell Environ. **35:** 347–359.

Rantanen, M.K., Lehtiö, L., Rajagopal, L., Rubens, C.E., and Goldman, A. (2007). Structure of *Streptococcus agalactiae* serine/threonine phosphatase. The subdomain conformation is coupled to the binding of a third metal ion. FEBS J. **274:** 3128–3137.

Rintamäki, E., Martinsuo, P., Pursiheimo, S., and Aro, E.-M. (2000). Cooperative regulation of light-harvesting complex II phosphorylation via the plastoquinol and ferredoxin-thioredoxin system in chloroplasts. Proc. Natl. Acad. Sci. USA **97:** 11644–11649.

Rochaix, J.-D. (2007). Role of thylakoid protein kinases in photosynthetic acclimation. FEBS Lett. **581:** 2768–2775.

Rochaix, J.-D. (2013). Redox regulation of thylakoid protein kinases and photosynthetic gene expression. Antioxid. Redox Signal. **18:** 2184–2201.

Rochaix, J.-D., Lemeille, S., Shapiguzov, A., Samol, I., Fucile, G., Willig, A., and Goldschmidt-Clermont, M. (2012). Protein kinases and phosphatases involved in the acclimation of the photosynthetic apparatus to a changing light environment. Philos. Trans. R. Soc. Lond. B Biol. Sci. **367:** 3466–3474.

Samol, I., Shapiguzov, A., Ingelsson, B., Fucile, G., Crèvecoeur, M., Vener, A.V., Rochaix, J.D., and Goldschmidt-Clermont, M. (2012). Identification of a photosystem II phosphatase involved in light acclimation in *Arabidopsis.* Plant Cell **24:** 2596–2609.

Shapiguzov, A., Ingelsson, B., Samol, I., Andres, C., Kessler, F., Rochaix, J.D., Vener, A.V., and Goldschmidt-Clermont, M. (2010). The PPH1 phosphatase is specifically involved in LHCII dephosphorylation and state transitions in *Arabidopsis*. Proc. Natl. Acad. Sci. USA **107:** 4782–4787.

Shi, Y. (2009). Serine/threonine phosphatases: mechanism through structure. Cell **139:** 468–484.

Silverstein, T., Cheng, L., and Allen, J.F. (1993). Chloroplast thylakoid protein phosphatase reactions are redox-independent and kinetically heterogeneous. FEBS Lett. **334:** 101–105.

Soon, F.F., et al. (2012). Molecular mimicry regulates ABA signaling by SnRK2 kinases and PP2C phosphatases. Science **335:** 85–88.

Su, J., Schlicker, C., and Forchhammer, K. (2011). A third metal is required for catalytic activity of the signal-transducing protein phosphatase M tPphA. J. Biol. Chem. **286:** 13481–13488.

Sun, G., Bailey, D., Jones, M.W., and Markwell, J. (1989). Chloroplast thylakoid protein phosphatase is a membrane surface-associated activity. Plant Physiol. **89:** 238–243.

Tanoue, K., Miller Jenkins, L.M., Durell, S.R., Debnath, S., Sakai, H., Tagad, H.D., Ishida, K., Appella, E., and Mazur, S.J. (2013). Binding of a third metal ion by the human phosphatases PP2Cα and Wip1 is required for phosphatase activity. Biochemistry **52:** 5830–5843.

Vener, A.V., Harms, A., Sussman, M.R., and Vierstra, R.D. (2001). Mass spectrometric resolution of reversible protein phosphorylation in photosynthetic membranes of *Arabidopsis thaliana*. J. Biol. Chem. **276:** 6959–6966.

Vener, A.V., van Kan, P.J., Rich, P.R., Ohad, I., and Andersson, B. (1997). Plastoquinol at the quinol oxidation site of reduced cytochrome *bf* mediates signal transduction between light and protein phosphorylation: thylakoid protein kinase deactivation by a single-turnover flash. Proc. Natl. Acad. Sci. USA **94:** 1585–1590.

Wehenkel, A., Bellinzoni, M., Schaeffer, F., Villarino, A., and Alzari, P.M. (2007). Structural and binding studies of the three-metal center in two mycobacterial PPM Ser/Thr protein phosphatases. J. Mol. Biol. **374:** 890–898.

Yin, P., Fan, H., Hao, Q., Yuan, X., Wu, D., Pang, Y., Yan, C., Li, W., Wang, J., and Yan, N. (2009). Structural insights into the mechanism of abscisic acid signaling by PYL proteins. Nat. Struct. Mol. Biol. **16:** 1230–1236.

Zito, F., Finazzi, G., Delosme, R., Nitschke, W., Picot, D., and Wollman, F.A. (1999). The Qo site of cytochrome b_6f complexes controls the activation of the LHCII kinase. EMBO J. **18:** 2961–2969.