Molecules and Cells

Structural and Biochemical Characterization of the Two *Drosophila* Low Molecular Weight-Protein Tyrosine Phosphatases DARP and Primo-1

Hye Seon Lee[1,2], Yeajin Mo[1,2], Ho-Chul Shin[1], Seung Jun Kim[1,*], and Bonsu Ku[1,*]

[1]Disease Target Structure Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon 34141, Korea, [2]These authors contributed equally to this work.
*Correspondence: ksj@kribb.re.kr (SJK); bku@kribb.re.kr (BK)
https://doi.org/10.14348/molcells.2020.0192
www.molcells.org

The *Drosophila* genome contains four low molecular weight-protein tyrosine phosphatase (LMW-PTP) members: Primo-1, Primo-2, CG14297, and CG31469. The lack of intensive biochemical analysis has limited our understanding of these proteins. Primo-1 and CG31469 were previously classified as pseudophosphatases, but CG31469 was also suggested to be a putative protein arginine phosphatase. Herein, we present the crystal structures of CG31469 and Primo-1, which are the first *Drosophila* LMW-PTP structures. Structural analysis showed that the two proteins adopt the typical LMW-PTP fold and have a canonically arranged P-loop. Intriguingly, while Primo-1 is presumed to be a canonical LMW-PTP, CG31469 is unique as it contains a threonine residue at the fifth position of the P-loop motif instead of highly conserved isoleucine and a characteristically narrow active site pocket, which should facilitate the accommodation of phosphoarginine. Subsequent biochemical analysis revealed that Primo-1 and CG31469 are enzymatically active on phosphotyrosine and phosphoarginine, respectively, refuting their classification as pseudophosphatases. Collectively, we provide structural and biochemical data on two *Drosophila* proteins: Primo-1, the canonical LMW-PTP protein, and CG31469, the first investigated eukaryotic protein arginine phosphatase. We named CG31469 as DARP, which stands for *Drosophila* ARginine Phosphatase.

Keywords: crystal structure, DARP, low molecular weight-protein tyrosine phosphatase, Primo-1, protein arginine phosphatase

INTRODUCTION

Protein tyrosine phosphatases (PTPs) are a family of enzymes that mediate the reversible dephosphorylation of phosphorylated tyrosine residues on substrate proteins, one of the most common and critical post-translational modifications (Tonks, 2006). They are closely associated with the regulation of various intracellular signaling and cellular processes by controlling protein-protein interaction, protein stability, and enzyme activity (Ardito et al., 2017; Hunter, 1995). The *Drosophila* genome contains 44 PTPs, which are less than half of 109 PTPs encoded by the human genome (Hatzihristidis et al., 2015). The 44 *Drosophila* PTP proteins can be divided into four subgroups: 37 class I PTPs that include 16 classical PTPs and 21 dual-specificity phosphatases (DUSPs), four class II low molecular weight-protein tyrosine phosphatases (LMW-PTPs), two class III CDC25 proteins, and one eyes-absent phosphatase (Hatzihristidis et al., 2015). The number of members of each PTP subgroup is higher in humans than in *Drosophila*, except for LMW-PTPs, of which the human genome contains only one (1/109, 0.9%), whereas the *Drosophila* genome

Received 22 September, 2020; revised 18 November, 2020; accepted 19 November, 2020; published online 23 December, 2020

eISSN: 0219-1032

contains four (4/44, 9.1%), including Primo-1, Primo-2, CG14297, and CG31469 (Hatzihristidis et al., 2015).

LMW-PTPs are found in a broad spectrum of genera from bacteria to higher eukaryotes, including *Drosophila* and human. While most PTP proteins consist of multiple domains, LMW-PTPs comprise only a single PTP domain. LMW-PTPs act on phosphotyrosine, but not on phosphoserine and phosphothreonine (Caselli et al., 2016); however, a case of atypical substrate preference of LMW-PTP was also reported. Based on structural and biochemical analyses, Fuhrmann et al. (2013) reported that YwlE, an LMW-PTP protein from *Bacillus subtilis*, is active toward phosphoarginine rather than phosphotyrosine, rendering it a novel type of phosphatase, protein arginine phosphatase. Moreover, the authors conducted a bioinformatics search, identifying CG31469 from *D. melanogaster* as the first putative eukaryotic protein arginine phosphatase (Fuhrmann et al., 2013).

The four *Drosophila* LMW-PTP members, including CG31469, have not been intensively investigated thus far. While Primo-2 was verified to have phosphatase activity toward *para*-nitrophenylphosphate (*p*NPP) (Miller et al., 2000), Primo-1 and CG31469 were previously classified as "pseudophosphatases" without experimental proof (Hatzihristidis et al., 2015), contradicting the previous finding that CG31469 has arginine phosphatase activity (Fuhrmann et al., 2013). Herein, we present the crystal structures of Primo-1 and CG31469, representing the first three-dimensional structures of *Drosophila* LMW-PTPs. Structural and biochemical analysis revealed that Primo-1 and CG31469 are catalytically active toward phosphotyrosine and phosphoarginine, respectively. Based on these data, we consider that both the enzymes should be removed from the list of pseudophosphatases. Notably, we named CG31469 as "DARP", which stands for *Drosophila* ARginine Phosphatase. To our knowledge, the present crystal structure of DARP is the first as the eukaryotic protein arginine phosphatase structure.

MATERIALS AND METHODS

Preparation, crystallization, and structure determination of DARP and Primo-1

The DNA fragments coding for DARP and Primo-1 were amplified by polymerase chain reaction and subcloned into the pET21a plasmid, which were then used as templates for preparing mutant proteins: DARP(C7S), DARP(T11I), DARP(Y127F), DARP(T11I·Y127F), and Primo-1(C9S). The proteins were produced in the *Escherichia coli* BL21(DE3) RIL strain (Novagen, USA) at 18°C and purified using a Ni-NTA column (Qiagen, Germany) and a HiLoad™ 26/600 Superdex™ 75 pg column (GE Healthcare, USA), equilibrated with 20 mM Tris-HCl (pH 7.5), 200 mM NaCl, and 2 mM dithiothreitol. Crystals were obtained by the sitting-drop vapor diffusion method at 18°C by mixing and equilibrating 1 μl protein solution (12 mg/ml DARP(C7S) or 6 mg/ml Primo-1) and precipitant solution containing 1.6 M ammonium phosphate monobasic, 100 mM sodium acetate trihydrate (pH 5.2), and 10 mM calcium chloride for DARP(C7S) or 0.1 M 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)-NaOH (pH 7.47), 1.2 M sodium citrate tribasic, and 9% (w/v) D-sorbitol for Primo-1. Before data collection, the crystals were briefly soaked in a cryoprotectant solution consisting of the precipitant solution supplemented with 20% ethylene glycol in the case of DARP(C7S) or with 5% glycerol in the case of Primo-1. Diffraction data were collected on the beamline 7A at the Pohang Accelerator Laboratory, Korea, and processed using the program *HKL* 2000 (Otwinowski and Minor, 1997). Both structures were determined by molecular replacement using the program Phaser (McCoy et al., 2007) using the *Streptococcus pyogenes* SP-PTP structure as a search model (Ku et al., 2016). Coot (Emsley and Cowtan, 2004) and PHENIX (Adams et al., 2010) were used for model building and refinement, respectively. Crystallographic data statistics are summarized in Table 1.

Dephosphorylation assays using Primo-1

The phosphatase activity of Primo-1 toward *p*NPP was measured by detecting the absorbance at 405 nm for 120 min using a SpectraMax plus 384 spectrophotometer (Molecular Devices, USA). The reaction mixture contained 30 nM recombinant Primo-1, 10 mM *p*NPP, 100 mM Bis-Tris (pH 6.0), 200 mM NaCl, and 2 mM dithiothreitol. Kinetic parameters were calculated from initial velocity data obtained by varying the *p*NPP concentration in from 0 to 20 mM. The EnzChek Phosphate Assay Kit (Thermo Fisher Scientific, USA) was used for detecting the phosphatase activity of Primo-1 toward phosphotyrosine by measuring the absorbance at 360 nm for 60 min. Assays were performed at room temperature using 2.5 μM recombinant Primo-1 and 5 mM phosphotyrosine in a 200 μl reaction mixture containing 0.2 mM 2-amino-6-mercapto-7-methylpurine riboside, 0.2 units purine nucleoside phosphorylase, 50 mM Bis-Tris (pH 7.5), and 1 mM $MgCl_2$. The reaction conditions were optimized by varying the amount of enzyme and substrates and pH.

Dephosphorylation assays using DARP

Substrate specificity was measured at room temperature using 1 μM wild-type or C7S mutant recombinant DARP and four types of 600 μM phosphoamino acids in a 200 μl reaction mixture containing 0.2 mM 2-amino-6-mercapto-7-methylpurine riboside, 0.2 units purine nucleoside phosphorylase, 50 mM Tris-HCl, and 1 mM $MgCl_2$. The EnzChek Phosphate Assay Kit was used for determining the phosphatase activity of DARP against the four phosphoamino acids by detecting the absorbance at 360 nm for 60 min using a SpectraMax plus 384 spectrophotometer. Kinetic parameters were calculated from initial velocity data, which were obtained at room temperature using 800 nM recombinant DARP and 0 to 5 mM phosphoarginine in a 200 μl reaction mixture containing 0.2 mM 2-amino-6-mercapto-7-methylpurine riboside, 0.2 units purine nucleoside phosphorylase, 50 mM Tris (pH 7.0), and 1 mM $MgCl_2$. Specific activities of the wild-type and mutant DARP proteins toward phosphoarginine and phosphotyrosine were calculated from initial velocity data (0-10 min), which were determined at room temperature by using the EnzChek Phosphate Assay Kit using a 200 μl reaction mixture containing 500 μM protein and 800 nM phosphoarginine at pH 7.0 or 10 μM protein and 5 mM phosphotyrosine at pH 7.5. The nonenzymatic hydrolysis

Table 1. Data collection and structure refinement statistics

	DARP(C7S)	Primo-1
Data collection		
Space group	$P2_12_12_1$	$P2_1$
Unit cell dimensions		
a, b, c (Å)	40.2, 62.2, 71.1	45.9, 59.9, 60.9
α, β, γ (°)	90, 90, 90	90, 110.2, 90
Wavelength (Å)	0.9793	0.9793
Resolution (Å)	50.0-2.2 (2.24-2.20)[a]	50.0-2.1 (2.14-2.10)[a]
R_{sym}[b]	7.2 (26.1)	9.5 (25.2)
$I/\sigma(I)$	39.9 (6.2)	23.8 (4.3)
Completeness (%)	98.2 (96.1)	97.4 (94.2)
Redundancy	5.3	4.7
Refinement		
Resolution (Å)	50.0-2.2	50.0-2.1
No. of reflections	9,483	18,040
R_{work}[c] / R_{free}	20.3 / 25.4	18.0 / 24.0
No. of atoms		
Protein	1,247	2,396
Water and ion	48	123
HEPES	-	30
RMSD		
Bond length (Å)	0.008	0.008
Bond angle (°)	1.051	0.954
Ramachandran plot (%)		
Most favored region	97.4	98.0
Additionally allowed region	2.6	2.0
Average B-value (Å^2)		
Protein	38.7	29.4
Water and ion	41.5	32.9
HEPES	-	30.7

[a]The numbers in parentheses are statistics from the shell with the highest resolution.
[b]$R_{sym} = \Sigma |I_{obs} - I_{avg}| / I_{obs}$, where I_{obs} is the observed intensity of individual reflection and I_{avg} is the average over symmetry equivalents.
[c]$R_{work} = \Sigma ||F_o| - |F_c|| / \Sigma |F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure factor amplitudes, respectively. R_{free} was calculated with 10% of the data.

of substrates was also measured for baseline correction using a control sample. Substrate specificity and kinetic constant measurements were performed in triplicate, and specific activity measurements were performed in duplicate.

Data availability

The atomic coordinates and the structure factors of DARP(C7S) and Primo-1 have been deposited in the Protein Data Bank under the accession codes 7BTG and 7CUY, respectively.

RESULTS

Crystal structures of DARP and Primo-1

Among the four *Drosophila* LMW-PTPs (Fig. 1A), wild-type and catalytic cysteine-substituted mutant DARP, Primo-1, and Primo-2 were successfully expressed in *E. coli* and subsequently purified. The purified recombinant proteins were subjected to crystallization. Crystals of DARP(C7S) in the space group $P2_12_12_1$ and wild-type Primo-1 in the space group $P2_1$ were obtained, and the structures of these two *Drosophila* proteins were determined at resolutions of 2.2 Å and 2.1 Å, respectively (Table 1). One DARP and two Primo-1 molecule(s) were contained in the asymmetric unit of each protein crystal. Both structures adopt a canonical LMW-PTP fold, which comprises a central β-sheet that is composed of four parallel β-strands and five α-helices and one 3_{10}-helix that surround the β-sheet (Fig. 1B). The two crystal structures superimposed well, with a root-mean-square deviation (RMSD) value of 0.93 Å over 150 aligned C_α atoms (Fig. 1C, left). Only slight structural differences were observed; for example, α4 of DARP is longer than that of Primo-1 (Supplementary Fig. S1A). The structures (Fig. 1C, middle and right), as well as amino acid sequences (Supplementary Fig. S2), of DARP and Primo-1 share high similarity with those of previously defined LMW-PTP-folded proteins, such as human LMW-PTP isoform A (Fonseca et al., 2015), *S. pyogenes* SP-PTP (Ku et al., 2016), and *B. subtilis* arginine phosphatase YwlE (Fuhrmann et al., 2013).

Structural analysis of the catalytic loop composition of DARP and Primo-1

Given that DARP and Primo-1 were previously classified as enzymatically inactive pseudophosphatases (Hatzihristidis et

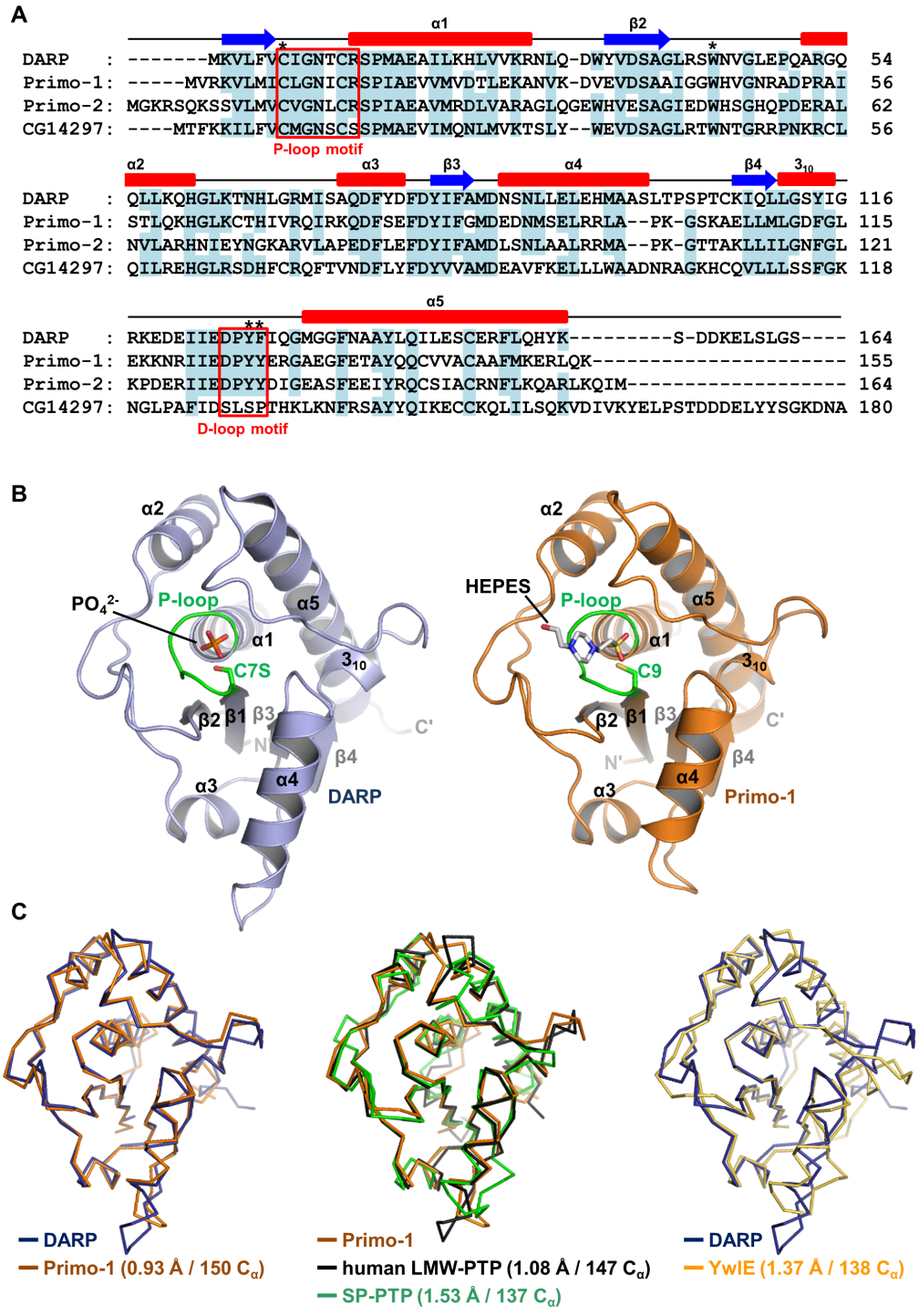

Fig. 1. Crystal structures of DARP and Primo-1. (A) Sequence alignment of four *Drosophila* LMW-PTP proteins. The secondary structure elements of DARP are shown above the sequences. Conserved residues are shaded in cyan. The P-loop and D-loop motif residues are highlighted by red rectangles. Asterisks denote the catalytic cysteine (Cys7 in DARP and Cys9 in Primo-1) and the three bulky aromatic residues in the catalytic pocket region (Trp43, Tyr127, and Phe128 in DARP and Trp45, Tyr126, and Tyr127 in Primo-1). (B) Ribbon drawings of the DARP (left) and Primo-1 (right) structures. Secondary structures are indicated by labels. DARP and Primo-1 are colored navy and orange, respectively, except for the P-loop region, which is colored green. The catalytic residue, C7S in DARP and Cys9 in Primo-1, and the phosphate ion or HEPES molecule bound to the catalytic pocket are displayed in stick representation. (C) Structural alignment. The superimposed structures of DARP, Primo-1, human LMW-PTP isoform A (PDB code 4Z99), SP-PTP (PDB code 5GOT), and YwlE (PDB code 4KK3) are shown in C_α trace representation. The RMSD values and the number of aligned C_α atoms are indicated in parentheses. The figures shown in Figs. B and C were drawn by PyMOL.

al., 2015), we investigated whether the two proteins have any structural feature in their catalytic pocket region that negatively affects their catalytic activity. We first analyzed the P-loop arrangement of the two proteins, which is a key factor restricting phosphatase activity in a number of PTP proteins such as PTPRU, DUSP6, DUSP9, and Tk-PTP (Fjeld et al., 2000; Hay et al., 2020; Hong et al., 2005; Yun et al., 2018; Zhou and Zhang, 1999). The catalytic motifs of DARP and Primo-1 are composed of C^{7}IGNTCR^{13} and C^{9}LGNICR^{15}, respectively, and are located at the β1 α1 loop, which is called the phosphate-binding loop or simply P-loop (Fig. 1A). The P-loop motif contains the catalytic cysteine residue (substituted with serine in the DARP structure), which functions as a nucleophile, and the conserved arginine residue (Arg13 in DARP and Arg15 in Primo-1), which anchors the phosphate group of the substrate during dephosphorylation (Fig. 2A). The center of P-loop is occupied by a single phosphate ion in the DARP structure and by the sulfonic group of a HEPES molecule in the Primo-1 structure. Oxygen atoms of these molecules mediate hydrogen bonds with the protein main chain amides and electrostatic interactions with the guanidinium group of the conserved arginine residue, which not only stabilize their positioning in the catalytic pocket but also maintain the active loop conformation of the *Drosophila*

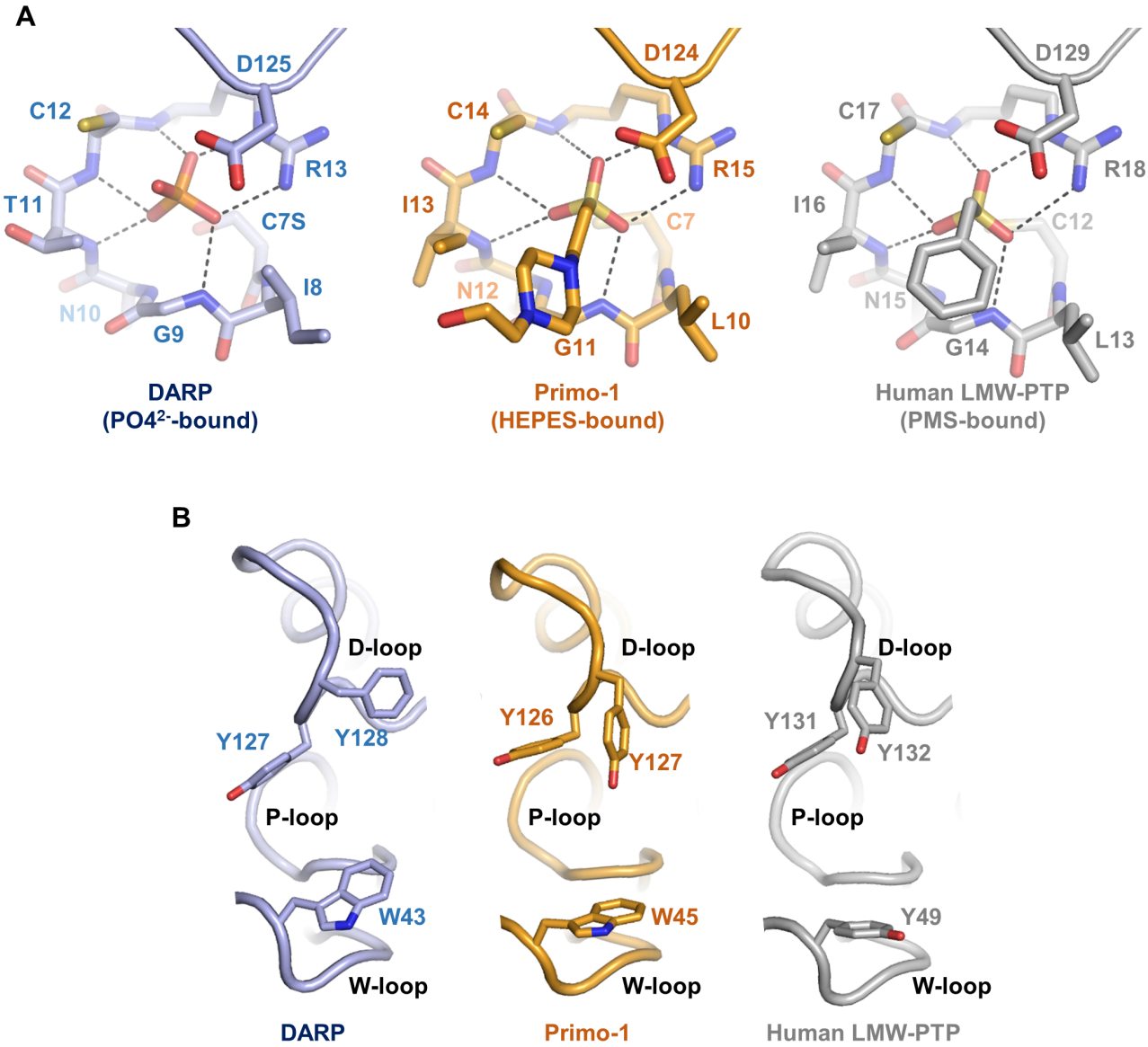

Fig. 2. Catalytic loop arrangement of DARP and Primo-1. (A) Catalytic loop composition. The P-loop residues and general acid/base aspartate residues of the three indicated proteins, together with bound ions or phosphotyrosine mimetics, are shown in stick representation. Hydrogen bonds and electrostatic interactions are indicated by dashed lines. PDB codes: 4Z9A, human LMW-PTP bound to phenylmethanesulfonic acid (PMS). (B) Comparison of the active site regions. The three conserved bulky hydrophobic residues from W-loop and D-loop covering the catalytic pockets are shown in stick representation and labeled. All the figures were drawn by PyMOL.

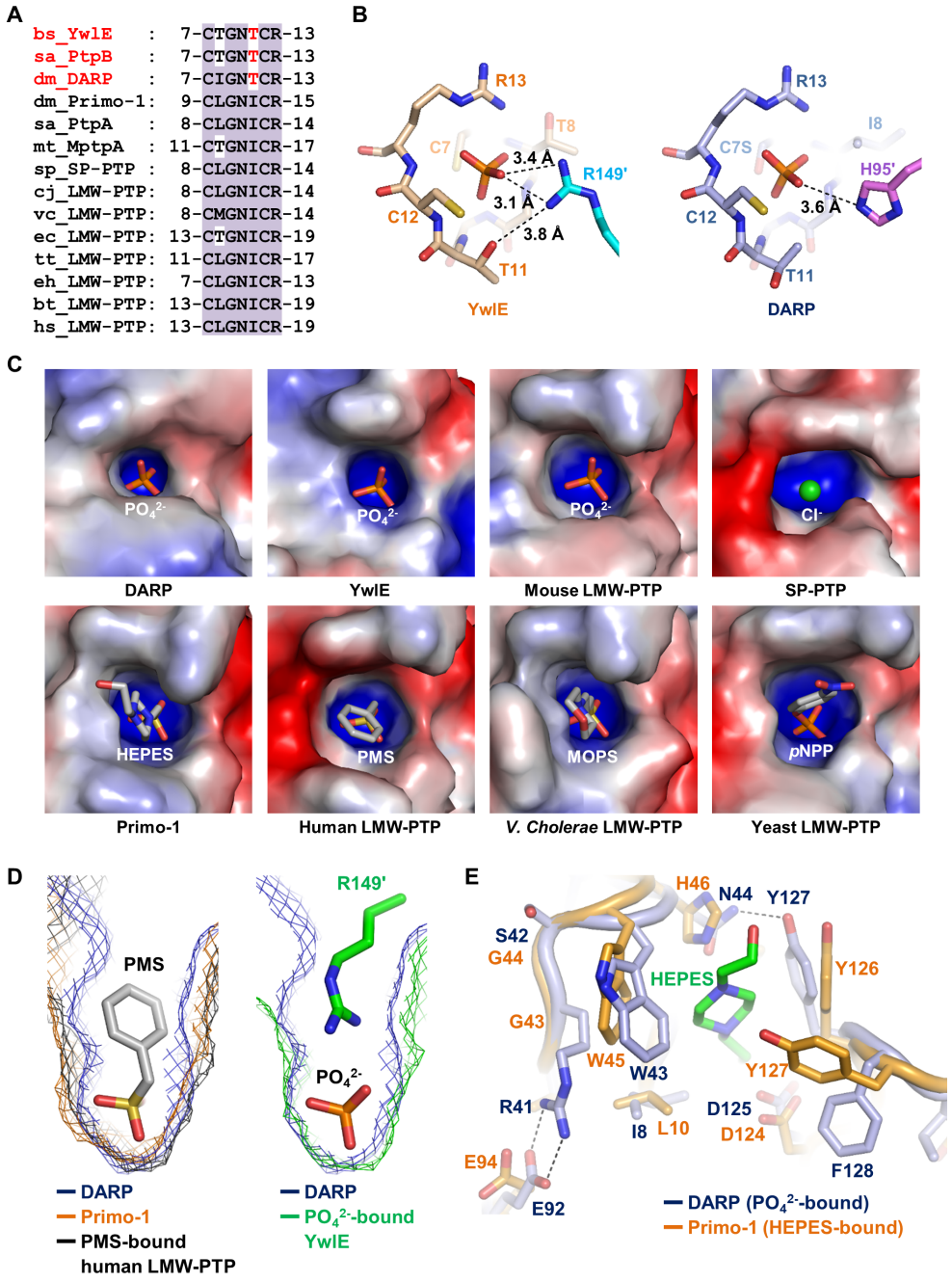

Fig. 3. Structural analysis of the catalytic pocket of DARP and Primo-1. (A) Sequences of the P-loop of three arginine phosphatases (red) and 11 LMW-PTPs (black) are aligned. Conserved residues are shaded in violet. The nonconserved threonine residues, which are only present in arginine phosphatases, are marked in red. bs, *Bacillus subtilis*; sa, *Staphylococcus aureus*; dm, *Drosophila melanogaster*; mt, *Mycobacterium tuberculosis*; sp, *Streptococcus pyogenes*; cj, *Campylobacter jejuni*; vc, *Vibrio cholerae*; ec, *Escherichia coli*; tt, *Thermus thermophilus*; eh, *Entamoeba histolytica*; bt, *Bovine taurus*; hs, *Homo sapiens*. (B) P-loop regions of YwlE and DARP, together with phosphate ions, are shown in stick representation with labels. Hydrogen bonds and electrostatic interactions with residues from crystallographically adjacent molecules (Arg149' in YwlE and His95' in DARP) are displayed as dashed lines. (C) Electrostatic surface representation of the active site pocket of eight phosphatases. Atoms or molecules bound to the active sites are shown as spheres or in stick representation. PMS, phenylmethanesulfonic acid; MOPS, 3-(N-morpholino)propane sulfonic acid. (D) Superimposition of the active sites of different phosphatases. The active site pockets of DARP (navy), Primo-1 (orange), PMS-bound human LMW-PTP (black; PDB code: 4Z9A), and PO_4^{2-}-bound YwlE (green) are shown as colored meshes. Molecules accommodated in the catalytic pocket are presented in stick representation and labeled. (E) Structurally aligned catalytic pockets of DARP (navy) and Primo-1 (orange). Catalytic site-bound molecules and key residues are displayed in stick representation and labeled. The hydrogen bond between Asn44 and Tyr127 and the electrostatic interaction between Arg41 and Glu92 in DARP are indicated by as dashed lines. The figures in Figs. B-E were drawn by PyMOL.

LMW-PTP proteins (Fig. 2A, Supplementary Fig. S1B). These backbone amides are oriented toward the catalytic pocket, and together with the arginine guanidinium group, create a positive electrostatic potential, which is a characteristic feature of catalytically active PTP proteins. This potential is not only complementary to the negatively charged phosphate group of the substitute but also stabilizes the side chain of the catalytic cysteine residue in a thiolate form (Kolmodin and Aqvist, 2001; Zhang et al., 1994). The conserved general acid/base aspartate residues (Asp125 in DARP and Asp124 in Primo-1) that stabilize the phosphoryl-intermediate during the catalytic reaction are also present in both structures at the typical position (Fig, 2A). The DARP and Primo-1 structures also contain two loops surrounding the P-loop region, so-called W-loop and D-loop (Fig. 2B). These loops are conserved among LMW-PTP structures; W-loop carries a tryptophan residue that contributes to substrate accommodation, while D-loop harbors the general acid/base aspartate residue and two substrate recognition-involved hydrophobic/aromatic residues (Fig. 2B). Taken together, we revealed that the P-loops of DARP and Primo-1 are arranged in the canonical form, typically found in enzymatically active PTP proteins, such as human LMW-PTP. Therefore, the proteins are expected to be capable of mediating dephosphorylation reactions.

Structural analysis of the phosphoarginine phosphatase activity of DARP

In contrast to a previous report classifying DARP as a pseudophosphatase (Hatzihristidis et al., 2015), a study by Fuhrmann et al. (2013) identifying *B. subtilis* YwlE as a bacterial protein arginine phosphatase showed that DARP (called CG31469 in their report) also exhibits a phosphoarginine phosphatase activity. We, therefore, analyzed the structure of DARP to account for its nonconventional substrate preference. In YwlE, the threonine residue (Thr11) located at position 5 of the P-loop, which is occupied by isoleucine in most LMW-PTPs (Fig. 3A), is critical for the enzyme's protein arginine phosphatase activity. In the YwlE structure, an arginine residue from a crystallographically adjacent molecule protrudes into the catalytic pocket, mimicking a substrate (Fig. 3B, left). The hydroxyl group of YwlE Thr11 points toward the catalytic pocket and forms hydrogen bonds with the arginine guanidinium group through its hydroxyl group, and therefore enhances the accessibility of the substrate phosphoarginine (Fuhrmann et al., 2013). This key threonine residue is also conserved in the P-loop of DARP at the same position as Thr11 (Fig. 3A). In the DARP structure, however, the hydroxyl group of Thr11 points away from the catalytic cleft due to the absence of a protruding and directing arginine residue. Instead, a histidine residue protrudes from a crystallographically adjacent molecule, but its side chain does not interact with the hydroxyl group of Thr11 (Fig. 3B, right). We assume that the hydroxyl group of Thr11 in DARP reorientates toward the catalytic pocket during the enzymatic reaction and thereby participates in phosphoarginine recognition like the corresponding group of the threonine residue in YwlE.

Next, to analyze the accessibility of the catalytic pocket more precisely, we structurally compared the size of the catalytic pocket between eight LMW-PTPs, including DARP and Primo-1, that contain a bound negatively charged ion or a phosphotyrosine-mimetic molecule. Intriguingly, DARP has a characteristically narrow catalytic pocket compared to that of other LMW-PTPs, while the Primo-1 catalytic pocket is comparable to those of canonical enzymes (Fig. 3C). The distance between the C_γ atom of Trp43 and the C_ζ atom of Tyr127 is 7.8 Å in DARP, which is shorter than the corresponding distances in human LMW-PTP (9.3 Å), Primo-1 (9.6 Å), and SP-PTP (9.8 Å). Consistently, structural alignments demonstrate that phosphoarginine, rather than phosphotyrosine, might fit into the narrow active pocket of DARP (Fig. 3D). Atomic-level structural analysis revealed several features that presumably contribute to the shape of the DARP catalytic pocket: the bulky side chain of Trp43, which is pushed toward the cleft by the aliphatic chain of Arg41; nonconserved electrostatic interactions between Arg41 and Glu92; and a hydrogen bond between Asn44 and Tyr127 that attracts this tyrosine residue toward the pocket (Fig. 3E). These features are unique to DARP, as they are absent in other LMW-PTP structures, including that of Primo-1, YwlE, and human LMW-PTP (Fig. 3E, Supplementary Fig. S1C). In contrast, the catalytic pocket composition of Primo-1 is highly homologous to that of conventional LMW-PTPs such as human LMW-PTP (Supplementary Fig. S1C). Collectively, these data strongly suggest that the accessibility of the catalytic pocket region, which is controlled by the presence of nonconserved threonine in the P-loop region (Figs. 3A and 3B) and the pocket size (Figs. 3C-3E), determines the substrate preference of DARP.

Biochemical characterization of Primo-1 phosphatase activity

Our structural data implied that DARP and Primo-1 might be enzymatically active phosphatases rather than pseudophosphatases. *In vitro* biochemical assays were performed to verify the phosphatase activity of the two proteins. First, the dephosphorylation capability of Primo-1, which was assumed to be a canonical LMW-PTP, was examined using *p*NPP, a widely used phosphatase substrate. Fig. 4A shows that recombinant Primo-1 was active on *p*NPP, confirming this protein as a catalytically active phosphatase. Kinetic constants were calculated using the Michaelis–Menten curves and the Lineweaver–Burk plots. Accordingly, the k_{cat}, K_M, and k_{cat}/K_M values of Primo-1 toward *p*NPP were determined as 59.8 min^{-1}, 0.520 mM, and 115 min^{-1} mM^{-1} (1.92×10^3 s^{-1} M^{-1}), respectively (Fig. 4B). Next, we measured the dephosphorylation activity of Primo-1 toward phosphotyrosine. The reaction of Primo-1 with phosphotyrosine released free phosphates, as evidenced by the absorbance at 360 nm (Fig. 4C), demonstrating that Primo-1 is able to dephosphorylate phosphotyrosine. The dephosphorylation activity of Primo-1 toward *p*NPP (Fig. 4A) and phosphotyrosine (Fig. 4C) was abrogated by substituting the catalytic cysteine residue (Cys9) to serine, further supporting that Primo-1 is a cysteine-based phosphatase.

Biochemical characterization of DARP phosphatase activity

Next, we conducted *in vitro* biochemical assays to verify the phosphatase activity of DARP toward four types of phosphoamino acids: phosphotyrosine, phosphoserine, phos-

phothreonine, and phosphoarginine. Recombinant DARP was only active on phosphoarginine, as evidenced by the absorbance at 360 nm, clearly demonstrating that DARP is an enzymatically active arginine phosphatase (Fig. 5A). Substitution of the catalytic Cys7 residue to serine abrogated the dephosphorylation activity of DARP (Fig. 5A). Initial velocity data were obtained to calculate the kinetic constants. From the Michaelis–Menten curve and Lineweaver–Burk double reciprocal plots, k_{cat}, K_M, and k_{cat}/K_M of DARP toward phosphoarginine were determined as 15.4 min^{-1}, 1.34 mM, and 11.5 min^{-1} mM^{-1} (192 s^{-1} M^{-1}), respectively (Fig. 5B). The kinetic constants of DARP toward phosphotyrosine, phosphoserine, and phosphothreonine were not calculable.

The enzymatic property of DARP was further analyzed by comparing the specific activities of wild-type and mutant proteins toward phosphoarginine and phosphotyrosine. Compared to the wild-type, substitution of Thr11 to isoleucine caused a remarkable decrease in the dephosphorylation of phosphoarginine and a noticeable increase in the dephosphorylation of phosphotyrosine (Fig. 5C). These results support the finding of the previous study reporting that the threonine residue at position 5 of the P-loop significantly contributes to the substrate specificity of DARP (Fuhrmann et al., 2013). We also mutated Tyr127 to phenylalanine to abrogate the hydrogen bond between Tyr127 and Asn44 (Fig. 3E), which, according to our crystal structure, is expected to widen the catalytic pocket of DARP and thereby improve phosphotyrosine accessibility (Figs. 3C-3E), without critically altering the active loop arrangement. Activity tests demonstrated that the Y127F mutant was slightly less active on phosphotyrosine than the wild-type enzyme. However, the dephosphorylation activity of the double mutant T11I·Y127F toward phosphotyrosine was noticeably higher than that of the T11I single mutant protein, suggesting that the pocket size of DARP (Fig. 3) is an additional factor controlling the access of phosphotyrosine into the catalytic site (Fig. 5C). Collectively, our biochemical analysis indicates that DARP is a catalytically active arginine phosphatase, and its substrate specificity depends on the amino acid composition of the catalytic pocket region.

DISCUSSION

In this study, we demonstrated that the two *Drosophila* proteins DARP and Primo-1 have not only the canonical LMW-PTP protein fold (Fig. 1) but also well-ordered P-loop and active site-surrounding loops (Fig. 2), implying that their catalytic pocket region are suitable for conducting the dephosphorylation reaction. More critically, our biochemical analysis proved that both enzymes exhibited phosphatase ac-

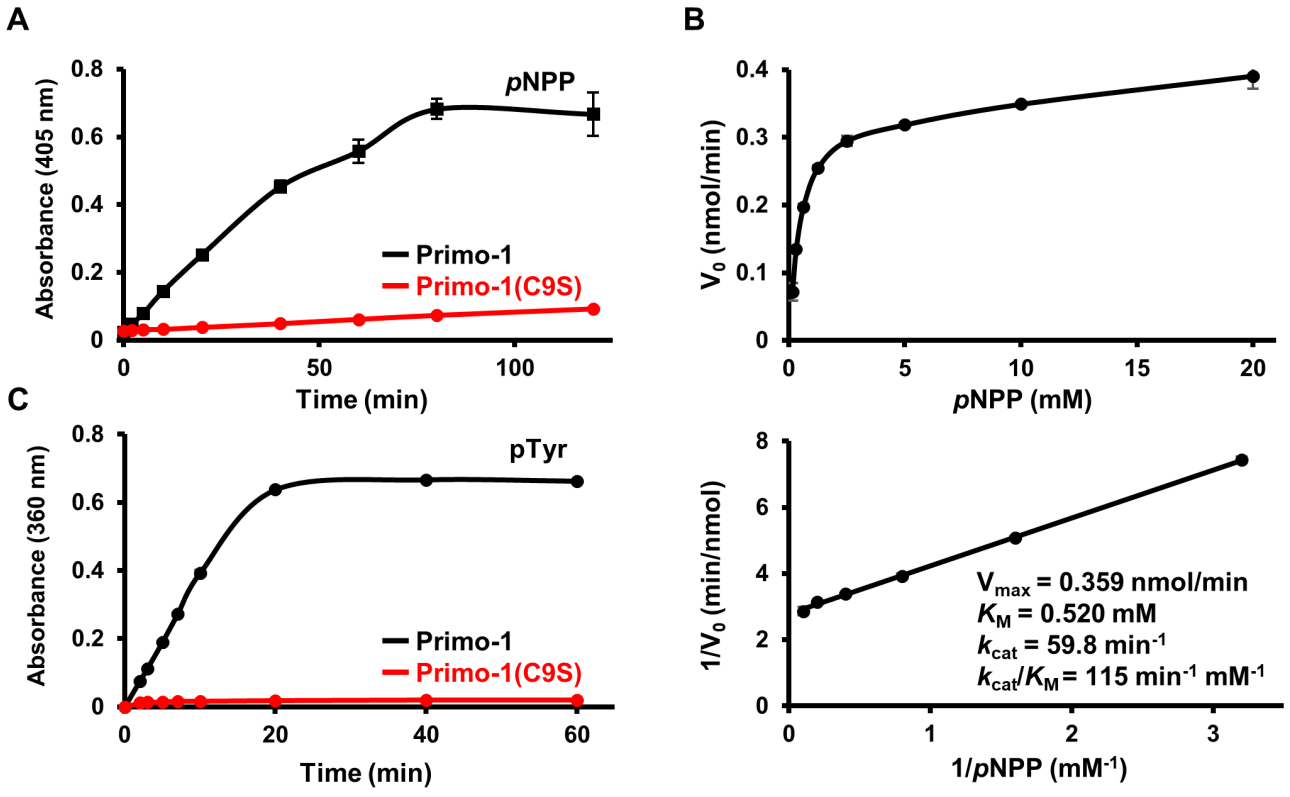

Fig. 4. Characterization of phosphatase activity of Primo-1. Experimental details are described in the Materials and Methods section. (A) Time-dependent detection of *p*NPP hydrolysis measured at 405 nm. (B) Michaelis-Menten curve (top) and Lineweaver-Burk plot (bottom). Graphs were deduced from initial velocity data, which were obtained between 5 min and 20 min after the start of the reaction. V_{max} and K_M values are shown with the Lineweaver-Burk plot. (C) Time-dependent detection of phosphotyrosine hydrolysis using the EnzChek Phosphate Assay Kit measured at 360 nm. pTyr, phosphotyrosine.

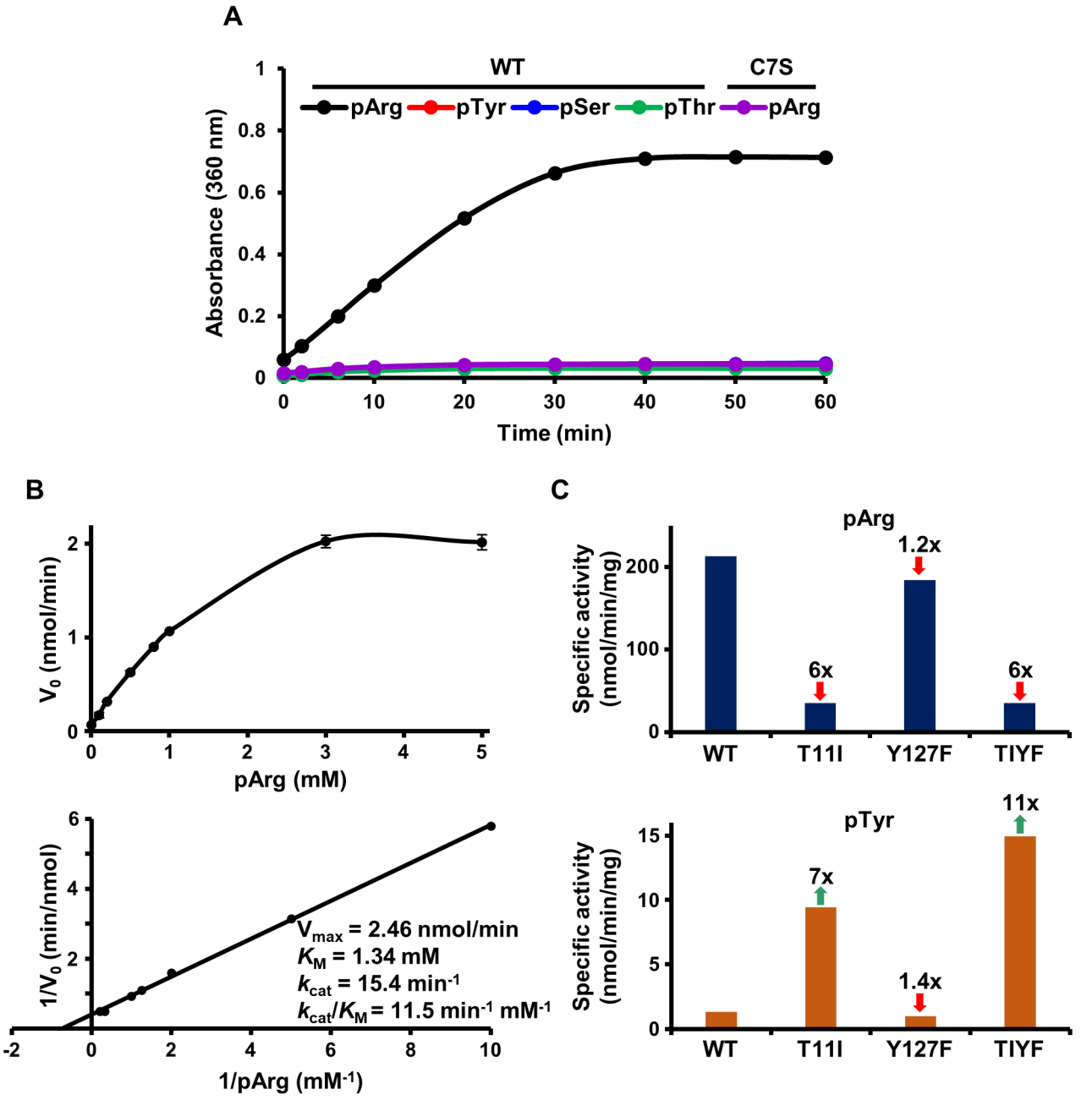

Fig. 5. Characterization of phosphatase activity of DARP. Experimental details are described in the Materials and Methods section. pArg, phosphoarginine; pTyr, phosphotyrosine; pSer, phosphoserine; pThr, phosphothreonine. (A) Substrate specificity. DARP acts only on phosphoarginine. (B) Michaelis-Menten curve (top) and Lineweaver-Burk plot (bottom) of DARP. Graphs were deduced from initial velocity data, which were obtained between 5 min and 8 min after the start of the reaction. V_{max} and K_M values are shown with the Lineweaver-Burk plots. (C) Specific activities of wild-type and three mutant proteins toward phosphoarginine (top) and phosphotyrosine (bottom). The differences in the specific activity between mutant proteins and the wild-type enzyme are highlighted by arrows (green, increase; red, decrease). TIYF, the mutant DARP protein containing T11I and Y127F substitutions.

tivity (Figs. 4 and 5). Taken together, we conclude that DARP and Primo-1 are catalytically active phosphatases rather than pseudophosphatases. One critical difference between the two proteins is that while Primo-1 is assumed to be a conventional phosphotyrosine-targeting phosphatase, DARP has a noncanonical substrate specificity toward phosphoarginine, consistent with the previous report (Fuhrmann et al., 2013). It is known that PTPs belonging to the same subgroup and thus sharing overall folding and P-loop motif sequence can target different amino acid substrates and even distinct molecules.

Differences in one or two amino acids, particularly in the signature P-loop motif, can significantly affect the three-dimensional composition of the catalytic pocket region and thereby alter or even abrogate the enzymatic activity of a PTP protein. Representatively, the secondary PTP domain D2 of all receptor-type PTPs are highly homologous to their primary PTP domain D1; however, D2 domains are enzymatically inactive and known as pseudophosphatases, with their various amino acid alterations in the P-loop motif and/or catalytic pocket region (Andersen et al., 2001). A recent study showed that the D1 domain of PTPRU is also a pseudophosphatase that, due to its atypical P-loop conformation, has no phosphatase activity (Hay et al., 2020). Altered P-loop arrangement is also a critical factor severely restricting the phosphatase activity of several DUSP proteins such as DUSP6, DUSP9, and Tk-PTP (Fjeld et al., 2000; Hong et al., 2005; Yun et al., 2018; Zhou and Zhang, 1999). In this study, we confirmed that DARP prefers phosphoarginine over phosphotyrosine as a substrate (Fig. 5), which relies on its nonconserved and hydrophilic threonine residue instead of hydrophobic isoleucine in the P-loop motif (C^{7}IGN<u>T</u>CR^{13}) and narrow catalytic pocket (Fig. 3). Therefore, the present crystal structure of DARP is the first and only three-dimensional structure of a eukaryotic protein arginine phosphatase.

Arginine, one of the twenty standard amino acids, contains a side chain composed of three-carbon aliphatic straight chain capped by a positively charged guanidinium group where a phosphate moiety can be covalently linked to. To date, phosphorylation and dephosphorylation of this amino acid, which are nonconventional compared to those of serine, threonine, and tyrosine, were mostly investigated in gram-positive bacteria; these reactions are involved in diverse bacterial physiological processes such as stress responses and protein degradation (Elsholz et al., 2012; Fuhrmann et al., 2016; Schmidt et al., 2014; Trentini et al., 2016). Contrastively, even though arginine phosphorylation was first reported in mouse proteins more than 30 years ago (Matthews, 1995; Wakim and Aswad, 1994), its precise biological functionality and associated proteins in eukaryotes remain to be elucidated. Therefore, the physiological role of DARP in *Drosophila* signaling pathways, and whether protein arginine phosphatases exist and function in other species, including human, are important and challenging issues to be addressed.

In summary, we determined the crystal structures of DARP and Primo-1, which are catalytically active phosphatases with an LMW-PTP fold. Structural and biochemical analyses revealed that DARP and Primo-1 dephosphorylate phosphoarginine and phosphotyrosine, respectively. The results of the present study will aid in finding unknown substrates for these proteins and signaling pathways in which the proteins are involved. This, in turn, might expand our understanding of post-translational phosphorylation and dephosphorylation controlling signaling pathways in *Drosophila*.

Note: Supplementary information is available on the Molecules and Cells website (www.molcells.org).

ACKNOWLEDGMENTS

We thank the beamline 7A at the Pohang Accelerator Laboratory in Korea for assistance during the data collection. This work was supported by the National Research Foundation of Korea funded by the Ministry of Science and ICT (MSIT; 2019M3E5D6063955) and by the Korea Research Institute of Bioscience and Biotechnology Research Initiative Program (to B.K.).

AUTHOR CONTRIBUTIONS

H.S.L. and Y.M. conceived and analyzed experiments and wrote the manuscript. H.C.S. provided expertise and feedback. S.J.K. provided supervision. B.K. conceived and analyzed experiments, wrote the manuscript, secured funding, and provided supervision.

CONFLICT OF INTEREST

The authors have no potential conflicts of interest to disclose.

ORCID

Hye Seon Lee https://orcid.org/0000-0002-9031-389X
Ho-Chul Shin https://orcid.org/0000-0001-7878-0367
Seung Jun Kim https://orcid.org/0000-0003-0293-6972
Bonsu Ku https://orcid.org/0000-0003-1784-8975

REFERENCES

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213-221.

Andersen, J.N., Mortensen, O.H., Peters, G.H., Drake, P.G., Iversen, L.F., Olsen, O.H., Jansen, P.G., Andersen, H.S., Tonks, N.K., and Moller, N.P. (2001). Structural and evolutionary relationships among protein tyrosine phosphatase domains. Mol. Cell. Biol. *21*, 7117-7136.

Ardito, F., Giuliani, M., Perrone, D., Troiano, G., and Lo Muzio, L. (2017). The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review). Int. J. Mol. Med. *40*, 271-280.

Caselli, A., Paoli, P., Santi, A., Mugnaioni, C., Toti, A., Camici, G., and Cirri, P. (2016). Low molecular weight protein tyrosine phosphatase: Multifaceted functions of an evolutionarily conserved enzyme. Biochim. Biophys. Acta *1864*, 1339-1355.

Elsholz, A.K., Turgay, K., Michalik, S., Hessling, B., Gronau, K., Oertel, D., Mader, U., Bernhardt, J., Becher, D., Hecker, M., et al. (2012). Global impact of protein arginine phosphorylation on the physiology of *Bacillus subtilis*. Proc. Natl. Acad. Sci. U. S. A. *109*, 7451-7456.

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

Fjeld, C.C., Rice, A.E., Kim, Y., Gee, K.R., and Denu, J.M. (2000). Mechanistic basis for catalytic activation of mitogen-activated protein kinase phosphatase 3 by extracellular signal-regulated kinase. J. Biol. Chem. *275*, 6749-6757.

Fonseca, E.M., Trivella, D.B., Scorsato, V., Dias, M.P., Bazzo, N.L., Mandapati, K.R., de Oliveira, F.L., Ferreira-Halder, C.V., Pilli, R.A., Miranda, P.C., et al. (2015). Crystal structures of the apo form and a complex of human LMW-PTP with a phosphonic acid provide new evidence of a secondary site potentially related to the anchorage of natural substrates. Bioorg. Med. Chem. *23*, 4462-4471.

Fuhrmann, J., Mierzwa, B., Trentini, D.B., Spiess, S., Lehner, A., Charpentier, E., and Clausen, T. (2013). Structural basis for recognizing phosphoarginine and evolving residue-specific protein phosphatases in gram-positive bacteria. Cell Rep. *3*, 1832-1839.

Fuhrmann, J., Subramanian, V., Kojetin, D.J., and Thompson, P.R. (2016).

Activity-based profiling reveals a regulatory link between oxidative stress and protein arginine phosphorylation. Cell Chem. Biol. *23*, 967-977.

Hatzihristidis, T., Desai, N., Hutchins, A.P., Meng, T.C., Tremblay, M.L., and Miranda-Saavedra, D. (2015). A *Drosophila*-centric view of protein tyrosine phosphatases. FEBS Lett. *589*, 951-966.

Hay, I.M., Fearnley, G.W., Rios, P., Kohn, M., Sharpe, H.J., and Deane, J.E. (2020). The receptor PTPRU is a redox sensitive pseudophosphatase. Nat. Commun. *11*, 3219.

Hong, S.B., Lubben, T.H., Dolliver, C.M., Petrolonis, A.J., Roy, R.A., Li, Z., Parsons, T.F., Li, P., Xu, H., Reilly, R.M., et al. (2005). Expression, purification, and enzymatic characterization of the dual specificity mitogen-activated protein kinase phosphatase, MKP-4. Bioorg. Chem. *33*, 34-44.

Hunter, T. (1995). Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. Cell *80*, 225-236.

Kolmodin, K. and Aqvist, J. (2001). The catalytic mechanism of protein tyrosine phosphatases revisited. FEBS Lett. *498*, 208-213.

Ku, B., Keum, C.W., Lee, H.S., Yun, H.Y., Shin, H.C., Kim, B.Y., and Kim, S.J. (2016). Crystal structure of SP-PTP, a low molecular weight protein tyrosine phosphatase from *Streptococcus pyogenes*. Biochem. Biophys. Res. Commun. *478*, 1217-1222.

Matthews, H.R. (1995). Protein kinases and phosphatases that act on histidine, lysine, or arginine residues in eukaryotic proteins: a possible regulator of the mitogen-activated protein kinase cascade. Pharmacol. Ther. *67*, 323-350.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658-674.

Miller, D.T., Read, R., Rusconi, J., and Cagan, R.L. (2000). The *Drosophila primo* locus encodes two low-molecular-weight tyrosine phosphatases. Gene *243*, 1-9.

Otwinowski, Z. and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307-326.

Schmidt, A., Trentini, D.B., Spiess, S., Fuhrmann, J., Ammerer, G., Mechtler, K., and Clausen, T. (2014). Quantitative phosphoproteomics reveals the role of protein arginine phosphorylation in the bacterial stress response. Mol. Cell. Proteomics *13*, 537-550.

Tonks, N.K. (2006). Protein tyrosine phosphatases: from genes, to function, to disease. Nat. Rev. Mol. Cell Biol. *7*, 833-846.

Trentini, D.B., Suskiewicz, M.J., Heuck, A., Kurzbauer, R., Deszcz, L., Mechtler, K., and Clausen, T. (2016). Arginine phosphorylation marks proteins for degradation by a Clp protease. Nature *539*, 48-53.

Wakim, B.T. and Aswad, G.D. (1994). Ca^{2+}-calmodulin-dependent phosphorylation of arginine in histone 3 by a nuclear kinase from mouse leukemia cells. J. Biol. Chem. *269*, 2722-2727.

Yun, H.Y., Lee, J., Kim, H., Ryu, H., Shin, H.C., Oh, B.H., Ku, B., and Kim, S.J. (2018). Structural study reveals the temperature-dependent conformational flexibility of Tk-PTP, a protein tyrosine phosphatase from *Thermococcus kodakaraensis* KOD1. PLoS One *13*, e0197635.

Zhang, Z.Y., Wang, Y., and Dixon, J.E. (1994). Dissecting the catalytic mechanism of protein-tyrosine phosphatases. Proc. Natl. Acad. Sci. U. S. A. *91*, 1624-1627.

Zhou, B. and Zhang, Z.Y. (1999). Mechanism of mitogen-activated protein kinase phosphatase-3 activation by ERK2. J. Biol. Chem. *274*, 35526-35534.