Arabidopsis C-terminal domain phosphatase-like 1 and 2 are essential Ser-5-specific C-terminal domain phosphatases

Hisashi Koiwa*†, Stéphane Hausmann‡, Woo Young Bang§, Akihiro Ueda*, Naoko Kondo¶, Akihiro Hiraguri¶, Toshiyuki Fukuhara¶, Jeong Dong Bahk§‖, Dae-Jin Yun§**, Ray A. Bressan††, Paul M. Hasegawa††, and Stewart Shuman‡

*Department of Horticultural Sciences, Texas A&M University, College Station, TX 77843-2133; ‡Molecular Biology Program, The Sloan–Kettering Institute, New York, NY 10021; §Division of Applied Life Science (BK21 Program) and **Environmental Biotechnology National Core Research Center, Graduate School of Gyeongsang National University, Jinju 660-701, Korea; ‖Agricultural Plant Stress Research Center, Chonnam National University, Gwangju 500-757, Korea; ¶Department of Applied Biological Sciences, Tokyo University of Agriculture and Technology, Saiwaicho 3-5-8, Fuchu, Tokyo 183-8509, Japan; and ††Center for Plant Environmental Stress Physiology, Purdue University, 625 Agriculture Mall Drive, West Lafayette, IN 47907-2010

Edited by Michael F. Thomashow, Michigan State University, East Lansing, MI, and approved August 24, 2004 (received for review May 6, 2004)

Transcription and mRNA processing are regulated by phosphorylation and dephosphorylation of the C-terminal domain (CTD) of RNA polymerase II, which consists of tandem repeats of a $Y^1S^2P^3T^4S^5P^6S^7$ heptapeptide. Previous studies showed that members of the plant CTD phosphatase-like (CPL) protein family differentially regulate osmotic stress-responsive and abscisic acid-responsive transcription in *Arabidopsis thaliana*. Here we report that AtCPL1 and AtCPL2 specifically dephosphorylate Ser-5 of the CTD heptad in *Arabidopsis* RNA polymerase II, but not Ser-2. An N-terminal catalytic domain of CPL1, which suffices for CTD Ser-5 phosphatase activity *in vitro*, includes a signature DXDXT acylphosphatase motif, but lacks a breast cancer 1 CTD, which is an essential component of the fungal and metazoan Fcp1 CTD phosphatase enzymes. The CTD of CPL1, which contains two putative double-stranded RNA binding motifs, is essential for the *in vivo* function of CPL1 and includes a C-terminal 23-aa signal responsible for its nuclear targeting. CPL2 has a similar domain structure but contains only one double-stranded RNA binding motif. Combining mutant alleles of *CPL1* and *CPL2* causes synthetic lethality of the male but not the female gametes. These results indicate that CPL1 and CPL2 exemplify a unique family of CTD Ser-5-specific phosphatases with an essential role in plant growth and development.

Transcriptional induction of genes that encode stress tolerance determinants is an integral part of the survival strategy of plants in adverse environments. The *Arabidopsis thaliana* responsive to dehydration (*RD*) genes are prototypal outputs of stress signal integration activated by low temperature, hyperosmolarity, and the plant hormone abscisic acid (ABA). The stress-inducible promoter of the *RD29a* gene contains dehydration/cold-responsive elements and ABA-responsive elements that are the targets of distinct families of DNA binding transcription factors (1, 2). The plant stress response is also regulated by proteins that impact the core RNA polymerase II (Pol II) transcriptional machinery, the mRNA maturation process, and chromatin structure (3–9). Analysis of *Arabidopsis* mutants that display hyperinduction of *RD29a* expression under stress conditions have identified a family of C-terminal domain (CTD) phosphatase-like (CPL) genes that negatively regulate stress-responsive transcription (5, 6). The *CPL1* and *CPL3* genes discovered in the screen for hyperinduction are so named because they encode large polypeptides (967 and 1,241 aa, respectively) with local primary structure similarity to the Fcp1 family of fungal and metazoan protein serine phosphatases, which regulate transcription by dephosphorylating the CTD of the largest subunit of RNA Pol II (10).

The Pol II CTD is composed of a tandemly repeated heptapeptide of consensus sequence $Y^1S^2P^3T^4S^5P^6S^7$. The number of CTD heptad repeats varies widely among species and correlates roughly with evolutionary complexity; e.g., mammals have 52 repeats, *Drosophila* has 42 repeats, fission yeast *Schizosaccharomyces pombe* has 29 repeats, and the microsporidian *Encephalitozoon cuniculi* has 15 repeats. The CTD of *Arabidopsis* Pol II consists of 34 heptad repeats. Studies of CTD function in fungi and metazoans have illuminated its role as a structurally plastic ''landing pad'' for proteins that regulate transcription and cotranscriptional mRNA processing (11). The CTD undergoes waves of phosphorylation and dephosphorylation at positions Ser-5 and Ser-2 during the transcription cycle and in response to environmental stress (10). Remodeling of the CTD phosphorylation array accompanies the transition from initiation to elongation modes and controls the recruitment, activity, and egress of the various mRNA processing machines that act on the nascent transcript (12–15). Fungi and metazoans encode multiple protein serine kinases (16) and protein serine phosphatases that modify the CTD. Among the CTD phosphatases, Fcp1 has been studied most intensively (17–20). Fcp1 orthologs are large polypeptides (723 aa in *S. pombe* and 961 aa in humans) consisting of a conserved central catalytic domain flanked by variable N- and C-terminal segments that are dispensable for CTD phosphatase activity *in vitro*, but are collectively required for Fcp1 function *in vivo*. The minimal phosphatase domain of *S. pombe* Fcp1 is a 425-aa polypeptide composed of two modules: an N-terminal FCP1 homology (FCPH) domain and a C-terminal breast cancer 1 C terminus (BRCT) domain, both of which are essential for Fcp1 phosphatase activity (21, 22). The FCPH domain of Fcp1 belongs to the D*X*D*X*T family of metal-dependent phosphotransferases that act via an acyl-phosphoenzyme intermediate (23). Fcp1 is essential for cell viability in budding and fission yeast (17, 24), and a partial deficiency of human Fcp1 is associated with an autosomal recessive developmental disorder (25). A separate class of human small CTD phosphatases (SCPs), containing the FCPH domain but not the BRCT domain, has been described (26). The effects of overexpressing WT and dominant-negative forms of SCP in mammalian cells suggest that SCP is a negative regulator of transcription (26).

Virtually nothing is known about the role of the CTD and the impact of CTD phosphorylation dynamics on gene expression in

PLANT BIOLOGY

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: CTD, C-terminal domain; CPL, CTD phosphatase-like; SCP, small CTD phosphatase; ABA, abscisic acid; Pol II, polymerase II; FCPH, Fcp1 homology; BRCT, breast cancer 1 C terminus; dsRBM, double-stranded RNA binding motif; NLS, nuclear localization signal; TAP, tandem affinity purification.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AY557186 and AY557187).

†To whom correspondence should be addressed. E-mail: koiwa@neo-tamu.edu.

plants. The *Arabidopisis* CPL proteins, which function genetically as brakes on stress-induced transcription, provide a starting point to fill this void. The predicted *Arabidopsis* proteome includes four FCPH domain-containing CPL proteins, which can be grouped into two paralogous pairs. CPL3 and CPL4 contain a C-terminal BRCT domain, whereas CPL1 and CPL2 do not (5). Here we address the following questions concerning the functionality of *Arabidopsis* CPL1 and CPL2. Do CPL1 and CPL2 have the CTD phosphatase activities imputed to them on the basis of local amino acid sequence similarity to Fcp1? If so, do they display any intrinsic preference for dephosphorylation of Ser-2 versus Ser-5? Are the structural features that distinguish CPL1 and CPL2 from CPL3 and CPL4 relevant for activity? What roles might CPL1 and CPL2 play in plant development, either individually or in concert?

Materials and Methods

Detailed information is provided in *Supporting Text*, which is published as supporting information on the PNAS web site. Sequences of oligo-DNA primers are shown in Table 2, which is published as supporting information on the PNAS web site.

CTD Phosphatase Assay. Preparation of synthetic phosphorylated CTD tetraheptad substrates and dephosphorylation assays were performed as described (21). GST-AtCTD was prepared as described (27), phosphorylated by activated ERK2 (Calbiochem) (28), and dialyzed overnight against storage buffer (10 mM Hepes, pH 7.9/10% glycerol/0.1% Triton X-100). Phosphatase reaction mixtures (10 μl) containing 50 mM Tris-acetate buffer as specified, 2.0 μg of phosphorylated GST-AtCTD, 10 mM $MgCl_2$ where specified, and CPLs as specified were incubated for 3 h at 37°C. Ser-2 and Ser-5 phosphorylation status was analyzed by immunoblotting with H5 and H14 mAbs (Covance), respectively, and chemiluminescence (Pierce).

Transient Expression of GFP-Fusion Proteins. GFP-CPL1 plasmids (30 μg) were introduced into *Arabidopsis* protoplasts by polyethylene glycol-mediated transformation as described (29). Transformed protoplasts were incubated at 22°C in the dark. Expression of the fusion protein was evaluated 2 and 3 days after transformation, and images were obtained on an Olympus AX-70 epifluorescence microscope equipped with standard FITC and rhodamine filters. Images from a Peltier-cooled 1.3 M-pixel Sensys camera (Roper Scientific, Trenton, NJ) were captured with the MACPROBE version 4.2.3 digital image system (Applied Imaging, San Jose, CA) and were then processed by using PHOTOSHOP software (Adobe Systems, San Jose, CA).

Genetic Analysis. Diheterozygous *fry2–1/+ cpl2–2* plants were prepared by pollinating *fry2–1* plants with *cpl2–2/+* plant pollens and subsequent PCR screening of F_1 plants for T-DNA insertion in *CPL2* locus as described in *Supporting Text*. The diheterozygous plants were self-pollinated to obtain F_2 progeny. The genotypes of the *CPL1* and *CPL2* loci of viable F_2 seedlings were determined as described in *Supporting Text*. For reciprocal cross analysis, a *fry2–1/+ cpl2–2/+* plant and a *fry2–1* plant were used as parents. F_1 seedlings with homozygous *fry2–1* genotype were selected by using *RD29a-LUC* expression after cold and ABA treatments as described (5, 6). The *fry2–1* phenotype of the F_1 plants indicated that the gametes from the diheterozygous parent contained the *fry2–1* chromosome. Subsequently, genomic DNA was extracted from each *fry2–1* F_1 plant and scored by PCR for the *cpl2–2* T-DNA insertion.

Results

CPL1/2-Like Proteins Are Specific to Plants. CPL1 (967 aa) and CPL2 (771 aa) are CPLs initially identified in *Arabidopsis*. They contain at their N termini an FCPH domain, characteristic of fungal and metazoan CTD phosphatases, but they lack the BRCT domain located downstream of the FCPH domains in the fungal, microsporidian, and metazoan Fcp1 enzymes. Instead, the CPL1 and CPL2 proteins contain putative double-stranded RNA binding motifs (dsRBMs) downstream of their FCPH domains. CPL1 has two dsRBMs, whereas CPL2 has one dsRBM. The other two *Arabidopsis* CPLs (CPL3 and CPL4) resemble the fungal and metazoan Fcp1 phosphatases in that they have a BRCT domain near their C termini. Deletion analyses have shown that the BRCT domain is essential for the phosphatase activity of yeast and metazoan Fcp1 proteins. The distinctive domain architecture of *Arabidopsis* CPL1 and CPL2 raises questions about their evolutionary origins and whether they are bona fide CTD phosphatases.

With respect to the first question, we were able to discern convincing homologs of CPL1 in the rice proteome (GenBank accession nos. CAD41201 and BAB63701) and among ESTs from plants including *Nicotiana benthamiana* and *Medicago truncatura* (GenBank accession nos. CK288050 and BI26702). An alignment of the full-length sequences of *Arabidopsis* CPL1 and CPL2 with rice CPL1 and CPL2 is shown in Fig. 5, which is published as supporting information on the PNAS web site. All four plant polypeptides contain an N-terminal FCPH domain that includes the acylphosphatase signature motif D*X*D*X*T (residues 161–165 in CPL1). Both of the aspartates of the D*X*D*X*T motif are essential for the activity of fungal, microsporidian, and metazoan CTD phosphatases (21, 23, 30). The overall identity between AtCPL1 and OsCPL1 is 45%, whereas AtCPL2 and OsCPL2 have 38% identity. The FCPH domain is especially well conserved (>50% identity among the four plant CPL1 and CPL2 proteins). In contrast, FCPH domains of AtCPL1 and AtCPL3 are only 18% identical, whereas AtCPL1 and human SCP1 are 26% identical in the FCPH domains. These comparisons indicate that CPL1 and CPL2 are a distinct lineage of plant-specific CTD phosphatase-like proteins with unique domain architecture.

***Arabidopsis* CPL1 and CPL2 Specifically Dephosphorylated Ser-5 of CTD Heptad Motif.** To determine whether CPL1 or CPL2 have CTD phosphatase activity, we produced them in bacteria as 6×His-tagged NusA fusions and isolated the recombinant fusion proteins from soluble bacterial lysates by Ni-affinity chromatography. (We found previously that 6×His-CPL1 was completely insoluble when produced in *Escherichia coli*; thus, we resorted here to the 6×His-NusA fusion strategy.) SDS/PAGE analysis showed that the affinity-purified material consisted of a heterogeneous array of polypeptides, the largest of which were consistent with the sizes expected of 6×His-NusA-CPL1 and 6×His-NusA-CPL2, respectively (Fig. 1*A*, denoted by arrows), the most abundant of which was similar in size to 6×His-NusA. Neither the putative full-sized 6×His-NusA-CPL proteins nor the collection of ≈60- to 130-kDa polypeptides migrating between them and 6×His-NusA were present in the preparation of recombinant 6×His-NusA that was purified in parallel (Fig. 1*A*). We surmise that the CPL1 and CPL2 fusion proteins were susceptible to proteolysis in *E. coli*, such that only a minor fraction of the 6×His-NusA fusions recovered contained intact CPL1 or CPL2. Nonetheless, we assayed the preparations for CTD phosphatase activity, using purified 6×His-NusA and purified recombinant *S. pombe* CTD phosphatase Fcp1 as negative and positive controls, respectively.

To gauge whether the CPLs have CTD phosphatase activity, and whether they display any preference for the position of phosphoserine within the CTD heptad, we used synthetic 28-aa peptides consisting of four tandem heptad repeats phosphorylated exclusively at Ser-2 or Ser-5 of each heptad. Reaction of the CPL1 and CPL2 preparations with the Ser-5-PO_4 peptide resulted in P_i release proportional to input protein, whereas the 6×His-NusA protein was inactive (Fig. 1*B*). Addition of 5 μg of

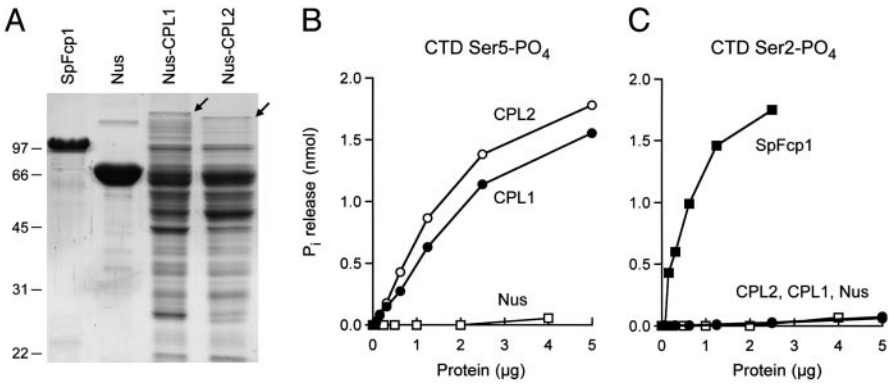

Fig. 1. CTD phosphatase activity of AtCPL proteins. (*A*) Protein preparations. Aliquots of the nickel-agarose preparations of *S. pombe* Fcp1 (SpFcp1; 6 μg), NusA (8 μg), NusA-CPL1 (12 μg), and NusA-CPL2 (12 μg) were analyzed by SDS/PAGE. The polypeptides were visualized by staining with Coomassie blue dye. The positions and sizes (in kDa) of marker proteins are indicated on the left. The polypeptides corresponding to the full-length NusA-CPL fusion proteins are indicated by arrows. (*B* and *C*) CTD phosphatase activity. Reaction mixtures (25 μl) containing 50 mM Tris-acetate (pH 5.5), 10 mM $MgCl_2$, either 22 μM CTD Ser-5-PO_4 phosphopeptide $(YSPTSPS)_4$ (containing 2.2 nmol of input Ser-PO_4) or 25 μM CTD Ser-2-PO_4 phosphopeptide $(YSPTSPS)_4$ (containing 2.5 nmol of input Ser-PO_4), and recombinant proteins as specified were incubated for 60 min at 37°C. Phosphate release is plotted as a function of input protein.

the CPL1 and CPL2 preparations sufficed to hydrolyze 70% and 82% of the input Ser-5-PO_4 residues, respectively. The instructive finding was that the CPL1 and CPL2 preparations did not catalyze P_i release from the Ser-2-PO_4 peptide (Fig. 1*C*). Note that the Ser-2-PO_4 peptide was an effective substrate for *S. pombe* Fcp1 (Fig. 1*C*).

Dephosphorylation of the Ser-5-PO_4 CTD peptide by CPL1 and CPL2 required a divalent cation (Fig. 6 *A* and *B*, which is published as supporting information on the PNAS web site). (All known D*X*D*X*T phosphotransferases are metal-dependent enzymes.) Magnesium supported optimal activity at 10 mM concentration. Manganese and cobalt (10 mM) were also capable of satisfying the divalent cation requirement, whereas calcium, zinc, and copper (10 mM) were ineffective (Fig. 6 *A* and *B*). CPL1 and CPL2 displayed a bell-shaped pH profile with an optimum pH at 5.5–6.0 (Fig. 6*C*). It is sensible that CPL1 and CPL2, as D*X*D*X*T acyl-phosphatases, would have optimal activity at mildly acidic pH, because the phosphoryl transfer mechanism calls for an unprotonated aspartate nucleophile and a protonated aspartate general acid catalyst. These properties, plus the strong selectivity of the CPL1 and CPL2 preparations for hydrolysis of Ser-5 versus Ser-2, suggest that the observed CTD phosphatase activity is attributable to the recombinant 6×His-NusA-CPL1 and 6×His-NusA-CPL2 proteins, their impurity notwithstanding. This attribution is reinforced by ensuing studies of deleted versions of the CPL1 protein.

The C-Terminal Region of CPL1 Is Dispensable for CTD Ser-5 Phosphatase Activity. The primary structures of CPL1 and CPL2 diverge in their C-terminal regions except for the conserved dsRBMs (Fig. 5). The *fry2–1* mutant allele of *CPL1* results in hyperactive stress-induced transcription and mimics in most respect the effects of a nonsense mutation in the region of *CPL1* encoding the FCPH domain. The *fry2–1* phenotype is attributed to a point mutation at a splice junction between the eighth exon and the eighth intron (6). Here, we amplified by RT-PCR a CPL1 cDNA from mRNA obtained from the *fry2–1* mutant, using primers in the eighth and 10th exons (Fig. 7, which is published as supporting information on the PNAS web site). The cDNA included the eighth but not the ninth intron, indicating a specific splicing defect. The eighth intron of CPL1 is 101 nt and contains an in-frame stop codon. The predicted CPL1 protein encoded by the *fry2–1* mRNA consists of an N-terminal 676-aa fragment of the WT protein plus a C-terminal tail containing eight extra amino acids encoded by the included intron. The truncated mutant protein lacks the C-terminal 292-aa segment of WT CPL1 that contains the two putative dsRBMs.

To determine whether this deleted region is important for the catalytic activity of CPL1, we expressed a series of truncated CPL1 variants as GST fusion proteins in *E. coli* and isolated them from soluble extracts by glutathione-agarose affinity chromatography. SDS/PAGE analysis of the GST-$CPL1_{1–638}$, GST-$CPL1_{1–563}$, GST-$CPL1_{1–500}$, and GST-$CPL1_{47–638}$ preparations revealed polypeptide constituents of the size expected for each fusion protein (denoted by arrows in Fig. 2*A*). The preparations also contained a prominent polypeptide of ≈60 kDa, which could correspond to the bacterial chaperone GroEL. It is evident that the GST fusions of the truncated CPL1 proteins are purer than the preparations of NusA-CPL1 and NusA-CPL2 analyzed initially. The instructive findings were that 0.1 μg of the $CPL1_{1–638}$ and $CPL1_{1–563}$ proteins catalyzed near quantitative hydrolysis of the input Ser-5-PO_4 substrate, but were virtually inert with the Ser-2-PO_4 substrate (Fig. 2*B*). We estimated a minimal turnover number of 0.8 s^{-1} for the Ser-5 phosphatase activity of the truncated CPL1 proteins, which was comparable to the observed turnover number (0.4 s^{-1}) for hydrolysis of a Ser-2-PO_4 tetraheptad CTD substrate by *S. pombe* Fcp1 (29). Although we cannot compare the activity of these truncated versions with that of full-length CPL1 (because of difficulty purifying the latter), this experiment demonstrates that deletion of as many as 404 aa from the C terminus of CPL1 did not ablate its activity or alter its stringent specificity for dephosphorylation of Ser-5. We infer that the effects of the C-terminal truncation of CPL1 in the *fry2–1* mutant are likely attributable to defects other than simple loss of catalytic activity. A further incremental C-terminal truncation, GST-$CPL1_{1–500}$, resulted in a partial loss of CTD phosphatase activity, whereas the N-terminal deletion mutant GST-$CPL1_{47–638}$ was catalytically inactive (Fig. 2*B*).

CPL1 and CPL2 Dephosphorylate Ser-5 of the Intact *Arabidopsis* RNA Pol II CTD. We tested the ability of CPL1 to dephosphorylate the full-length CTD of *Arabidopsis* Pol II, which consists of 34 heptads (Fig. 3). Phosphorylation of recombinant GST-AtCTD by the ERK2 kinase caused a shift in the electrophoretic mobility of the AtCTD analogous to that seen with metazoan and fungal Pol II_O and GST-CTD_O from human. Accordingly, we designated this phosphorylated low-mobility form as GST-$AtCTD_O$ and the nonphosphorylated form as GST-$AtCTD_A$. GST-$AtCTD_O$ was recognized by mAbs specific to Ser-5-PO_4 (H14) and Ser-2-PO_4 (H5), consistent with previous studies (26, 28).

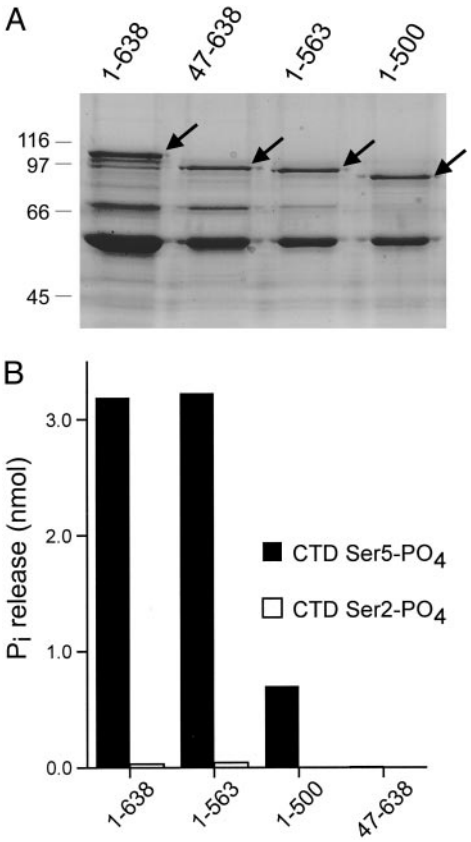

Fig. 2. Effects of N-terminal and C-terminal deletions on AtCPL1 CTD phosphatase activity. (*A*) GST-CPL1 fusions. Aliquots of the affinity-purified GST-AtCPL1 fusion proteins $CPL1_{1–638}$, $CPL1_{47–638}$, $CPL1_{1–563}$, and $CPL1_{1–500}$ containing ≈0.1 μg of the respective GST-CPL1 fusion polypeptides were analyzed by SDS/PAGE. The Coomassie blue-stained gel is shown. The positions and sizes (kDa) of marker proteins are indicated on the left. The polypeptides corresponding to the GST-CPL1 fusions are indicated by arrows. (*B*) CTD phosphatase activity. Reaction mixtures (25 μl) containing 50 mM Tris-acetate (pH 5.5), 10 mM $MgCl_2$, either 32 μM CTD Ser-5-PO_4 phosphopeptide $(YSPTSPS)_4$ (containing 3.2 nmol of input Ser-PO_4) or 36 μM CTD Ser-2-PO_4 phosphopeptide $(YSPTSPS)_4$ (containing 3.6 nmol of input Ser-PO_4), and ≈0.1 μg of the GST-CPL1 fusion polypeptides as specified were incubated for 60 min at 37°C. Representative results from three independent experiments are shown.

Upon treatment with $CPL1_{1–638}$, GST-$AtCTD_O$ was no longer recognized by H14, but it retained its immunoreactivity with H5. The selective elimination of H14 reactivity by CPL1 was magnesium-dependent and was evident when the reactions were performed at its optimum acidic pH (pH 5.5) and neutral pH (pH 7.0). Similar results were obtained for recombinant CPL2 (data not shown). The specificity of CPL1 and CPL2 for dephosphorylation of Ser-5 was underscored by the finding that treatment with the generic phosphatase calf intestine alkaline phosphatase abolished the reactivity of GST-$AtCTD_O$ to both H14 and H5 antibodies and reverted the electrophoretic mobility of GST-$AtCTD_O$ to that of unphosphorylated GST-$AtCTD_A$ (Fig. 3).

CPL1 Localizes to the *Arabidopsis* Nucleus. The subcellular localization of CPL1 and CPL2 was analyzed by transiently expressing tandem affinity purification (TAP)-GFP-CPL fusion proteins in *Arabidopsis* protoplasts (Fig. 4). The DsRed protein fused to the simian virus 40 (SV40) large T antigen nuclear localization signal (NLS) (31) was used as a positive control for nuclear targeting. In protoplasts expressing TAP-GFP-CPL1, green fluorescence

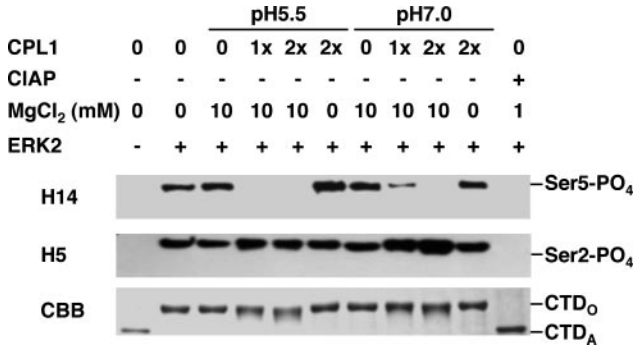

Fig. 3. CPL1 dephosphorylates intact AtCTD. GST-AtCTD protein was phosphorylated by activated mitogen-activated protein kinase ERK2. After overnight dialysis, GST-$AtCTD_O$ was incubated with 0.1 μg (1×) or 0.2 μg (2×) of the GST-$CPL1_{1–638}$ preparation (CPL1) or calf intestine alkaline phosphatase (CIAP) in reaction buffer (50 mM Tris, pH 5.5 or 7.0) without or with 10 mM $MgCl_2$ for 3 h at 37°C. Phosphorylation status was gauged after resolving the products by SDS/PAGE by immunoblotting with mAbs H14 and H5 that detect Ser-5-PO_4 and Ser-2-PO_4, respectively. (*Lower*) The polypeptide compositions were visualized by staining with Coomassie brilliant blue dye (CBB). The amount of GST-AtCTD applied to the gels was 0.1 μg for H14 and 1 μg for H5 and CBB staining.

was detected exclusively in nuclei and colocalized with red fluorescence from NLS_{SV40}-DsRed (Fig. 4*A*). The nuclear localization of CPL1 is consistent with genetic evidence that CPL1 regulates the transcription of stress-responsive genes. In protoplasts expressing TAP-GFP or TAP-GFP-CPL2, a diffuse pattern of green fluorescence was observed in cytosol and nucleus, indicating that neither protein had been targeted to a specific subcellular compartment. To further dissect the signals for nuclear targeting of CPL1, three polypeptide segments were fused individually to the C terminus of GFP (with no N-terminal TAP tag) to generate chimeric proteins GFP-$CPL1_{1–150}$, GFP-$CPL1_{151–639}$, and GFP-$CPL1_{640–967}$. Only GFP-$CPL1_{640–967}$ was specifically targeted to nuclei (Fig. 4*G*).

Localization of a discrete NLS within the $CPL1_{640–967}$ was achieved by fusing a series of CPL1 peptide fragments to GFP (Fig. 8, which is published as supporting information on the PNAS web site). The results show that (*i*) the dsRBM motifs play no apparent role in nuclear targeting; (*ii*) a 23-aa C-terminal peptide, $CPL1_{945–967}$ sufficed for nuclear localization of GFP; and (*iii*) deletion of amino acids 945–957 abolished the NLS

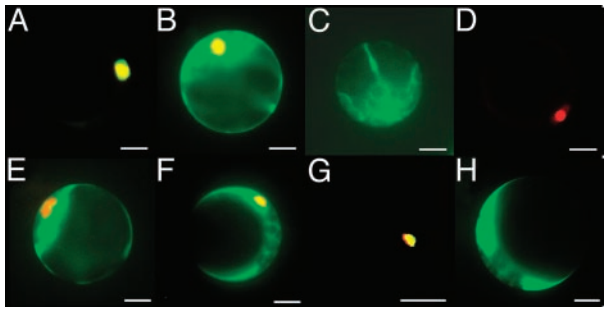

Fig. 4. Subcellular localization of CPL1 and CPL2. TAP-GFP-CPL1 (*A*), TAP-GFP-CPL2 (*B*), TAP-GFP (*C*), GFP-$CPL1_{1–150}$ (*E*), GFP-$CPL1_{151–639}$ (*F*), GFP-$CPL1_{640–967}$ (*G*), and GFP (*H*) were transiently expressed in *Arabidopsis* protoplasts. NLS_{SV40}-DsRed was used as a positive control for nuclear localization (*D*) and was coexpressed with GFP-fusion constructs (*A–C* and *E–G*). Fluorescent signals from GFP and DsRed were obtained by using standard FITC and rhodamine filter sets. The pictures show pseudocolor overlap images of GFP (green) and DsRed (red). Yellow indicates colocalization of GFP and RFP fusion proteins. (Scale bars: 10 μm.)

Table 1. Reciprocal cross analysis indicated synthetic lethality in *fry2–1 cpl2–2* male gamete

Cross					*fry2–1* phenotype, genotype	
Female		Male	Total	WT phenotype	*CPL2*	*cpl2–2*/+
fry2–1	×	*fry2–1*/+*cpl2–2*/+	112	76	36	0
fry2–1/+ *cpl2–2*/+	×	*fry2–1*	62	39	12	11

Each F_1 seedling was first scored for *fry2–1* phenotype by using *RD29a-LUC* expression after cold treatment (0°C, 48 h) and then after ABA treatment (100 μM, 3 h). Seedlings with *fry2–1* phenotype were recovered, and their genomic DNA was extracted. The *CPL2* genotype of *fry2-1* F_1 seedlings was determined by detecting T-DNA insertion in *CPL2* by using PCR.

activity (Fig. 8*A*). The C-terminal NLS region of *Arabidopsis* CPL1 is conserved in the rice CPL1 homolog (Fig. 8*B*).

CPL1 and CPL2 Have an Essential Redundant Function in Plant Development. T-DNA insertion or splice-junction mutations in the *CPL1* gene altered ABA sensitivity and freezing sensitivity in *Arabidopsis* seedlings, but did not cause lethality or a severe growth defect (6). This finding is in contrast to the lethality of *fcp1* loss-of-function mutations in fungi (17). To test functional redundancy between CPL1 and CPL2, we gauged the effect of combining the *cpl1/fry2* mutation with the *cpl2–2* allele, which contains a T-DNA insertion in the ninth exon of *CPL2* that disrupts its expression (Fig. 7). The *cpl2–2* mutation by itself causes early flowering in homozygous mutant plants (similar to the effects of the *fry2–1* and *cpl3–1* mutations) and reduced fertility. Nonetheless, homozygous *cpl2–2* plants were able to complete their life cycle and produce viable seeds (unpublished observations). When we analyzed the genotype of F_2 progeny of a *fry2–1*/+ *cpl2–2*/+ F_1 plant, no homozygous *fry2–1 cpl2–2* double mutants were identified in 76 seedlings. The segregation ratio of each locus indicated the synthetic lethality of the *fry2–1 cpl2–2* gamete rather than zygotes (data not shown). To determine the specific gametophytic lethality, the genotypes of F_1 progeny from the reciprocal cross between *fry2–1* and *fry2–1*/+ *cpl2–2*/+ were analyzed (Table 1). Individual F_1 plants were first scored for *fry2–1* phenotype by using luciferase imaging analysis to assay overexpression of an integrated luciferase reporter driven by the *RD29a* promoter (indicative of homozygosity for *fry2–1*) (5). The plants were then tested for the *cpl2–2* T-DNA insertion by PCR. When the *fry2–1* plant was used as a pollen donor, 12 of 23 *fry2–1* F_1 plants contained *cpl2–2* T-DNA, indicating that the *fry2–1* and *cpl2–2* mutations can be cotransmitted through the megagametephyte. In contrast, no *cpl2–2* T-DNA was detected in 36 *fry2–1* F_1 plants when the *fry2–1*/+ *cpl2–2*/+ plant was used as the pollen donor. These results established that the *fry2–1* and *cpl2–2* mutations cannot cotransmit through the pollen, and thus at least one functional *CPL1* family gene is required for normal plant development.

Discussion

Here we report that *Arabidopsis* CPL1 and CPL2 catalyze the metal-dependent hydrolysis of phosphoserine at position 5 of the CTD repeat YSPTSPS, but are apparently incapable of removing phosphoserine at position 2. The N-terminal half of the plant CPL1 protein that suffices for CTD Ser-5 phosphatase activity includes the FCPH module found in all fungal and metazoan CTD-specific phosphatases. Indeed, the plant CPL proteins are notable for conservation of all 11 amino acid side chains that were shown to be essential or important for the CTD phosphatase activity of *S. pombe* Fcp1 (see Fig. 5) (32). Thus, we predict that the architecture of the active site of the plant CPL1 and CPL2 enzymes is similar to that of other CTD phosphatases of the D*X*D*X*T family. In this light, the most striking property of the plant CPL1 and CPL2 phosphatases is their stringent inherent selectivity for Ser-5 versus Ser-2, which can be observed both with a synthetic peptide substrate and full-length AtCTD. This finding contrasts most starkly with the behavior of *S. pombe* Fcp1, which displays a 10-fold preference for dephosphorylating Ser-2-PO_4 versus Ser-5-PO_4 when assayed with the same set of tetraheptad CTD-PO_4 substrate used here to characterize the plant enzymes (21). On the other hand, *E. cuniculi* Fcp1 and the human SCP dephosphorylate both Ser-5 and Ser-2, with a slight (<2-fold) preference for Ser-5. Human Fcp1 dephosphorylates both Ser-5 and Ser-2 (26, 30).

Recently, Krishnamurthy *et al.* (33) reported a specific CTD Ser-5 phosphatase activity associated with *S. cerevisiae* Ssu72, which contains an cysteinyl-phosphatase-type active site and bears no resemblance to the acyl-phosphatase family to which the CPLs belong. Although *Arabidopsis* encodes a putative Ssu72 homolog, it has not been characterized genetically or biochemically. The fact that the *fry2–1 cpl2–2* double-mutation ablates male gamete function suggests that the Ser-5-specific phosphatase activity of the CPL1 and CPL2 proteins does not overlap functionally during plant development with a putative activity contributed by plant Ssu72. Given that multiple CTD kinases differentially phosphorylate the CTD at distinct phases of the transcription cycle in metazoans and fungi, it is reasonable to think that the CPLs and Ssu72 differentially dephosphorylate Pol II in an analogous fashion.

Two other remarkable features of the plant CPLs emerge from the present study. First, CPL1 and CPL2 have CTD phosphatase activity, although they lack the essential BRCT domain found in fungal, metazoan, and plant Fcp1 homologs. The catalytic contributions made by the BRCT domain of Fcp1 are not known, but a recent report implicates the BRCT domain in the initial binding of Fcp1 to the phosphorylated CTD (34). The present study underscores the insight from the recent discovery of human SCPs, that not all CTD phosphatases rely on a BRCT domain to achieve either CTD-PO_4 recognition or positional specificity for a particular Ser-PO_4 within the CTD heptad. Second, CPL1 and CPL2 are distinguished by the presence of putative dsRBMs on the carboxyl side of the FCPH domain. The simple notion that a dsRBM functions in lieu of the BRCT domain to promote CTD recognition is vitiated by the finding that deletion of the C-terminal segment containing both dsRBMs did not eliminate CTD phosphatase activity or alter its positional specificity. We conclude that the unique specificity of the plant CPL1 and CPL2 proteins for Ser-5 is intrinsic to the N-terminal catalytic domain. We show here that the CTD of CPL1 plays an ancillary role in nuclear targeting. It is likely that nuclear targeting is critical for CPL1 to function as a transcriptional regulator *in vivo*, insofar as deletion of the NLS-containing C-terminal peptide in the *fry2–1* mutant plants results in a hyperactive transcriptional response to stress. Of course, it is possible that the CTD plays additional roles *in vivo*, e.g., promoting protein–protein or protein–RNA interactions that direct the plant CTD phosphatase to specific genes or otherwise regulate its activity. Indeed, we recently determined that the

PLANT BIOLOGY

CTD of CPL1 can interact with dsRNA *in vitro* (A.H., R. Ito, N.K., Y. Nomura, D. Aizawa, Y. Murai, H.K., M. Seki, K. Shinozoki, and T.F., unpublished work). Precedent for RNA-mediated control of the CTD phosphorylation state emerged from the discovery that the activity of the mammalian CTD kinase PTEF-b is negatively regulated by binding to 7SK RNA (35–37).

The analysis of genetic interactions between the *fry2–1/cpl1* and *cpl2–2* alleles indicated that either CPL1 or CPL2 is required for male gamete function. Although the precise timing of the developmental arrest or demise of the *fry2–1 cpl2–2* double-mutant pollen has not been determined, the fact that the *fry2–1 cpl2–2* double mutation can transmit through female gametes indicates that specific CTD phosphatase isozymes are required in different types of cells. The dispensability of CPL1 and CPL2 in female gametes contrasts with the results obtained with *Xenopus* oocytes, where FCP1 is required to covert the hyperphosphorylated, but transcriptionally quiescent, pool of Pol II_O that is stored in the oocyte to hypophosphorylated Pol II_A, so that it can be recruited to preinitiation complexes upon fertilization (38).

Phosphorylation of CTD Ser-5 is believed to be a key step in the transition from initiation to elongation by fungal and metazoan Pol II and is also implicated in ensuring the timely capping of nascent pre-mRNAs (39). If similar CTD phosphorylation dynamics pertain in plants, then the action of CPL1 as a negative regulator of the transcriptional response to stress could be explained in several ways. CPL1 could decrease elongation complex formation by stripping Ser-5-PO_4 from the initiation complex and thereby affecting promoter clearance. Alternatively, dephosphorylation of Ser-5 in the early elongation complex might negatively affect the efficiency or timing of cotranscriptional mRNA capping. The latter model is appealing in light of the finding that a recessive mutation in the plant *ABH1* gene encoding a nuclear cap binding protein causes attenuation of ABA induction of at least 18 *Arabidopsis* genes (4). In contrast, the predominant molecular phenotype of *ssu72* mutation is the incomplete termination of small nucleolar RNA and some mRNA transcription by Pol II (40).

The specificity of the CPL1/CPL2 paralog pair for hydrolyzing Ser-5-PO_4 raises the prospect that the other plant paralog pair CPL3/CPL4 is responsible for removing Ser-2-PO_4 from the plant CTD. Unfortunately, our efforts to evaluate this idea by purifying and characterizing *Arabidopsis* CPL3 or CPL4 were thwarted by the inability to produce sufficient amounts of soluble recombinant CPL3 or CPL4 proteins in bacteria. Based on the present study of CPL1 and CPL2, we anticipate that further genetic and biochemical dissection of the specificity, expression, localization, and structure–function relationships of all four CPL proteins will provide insights into the role of CTD phosphorylation dynamics in gene-specific transcriptional control during development of a multicellular organism and in adaptation to environmental stress.

We thank Dr. Jian-Kang Zhu for helpful discussion, Dr. David Stelly for assistance in microscopy, and the *Arabidopsis* Biological Resource Center and The Salk Institute for *Arabidopsis* mutant collection and a cDNA clone. This work was supported by National Science Foundation Plant Genome Award DBI-9813360, College of Agriculture and Life Sciences, Texas A&M University; the BK21 program, Gyeongsang National University; Grants-in-Aid for Scientific Research 14390019 and 15028207 from the Ministry of Education, Science, Sports, Culture and Technology of Japan (to T.F.); Grant PF0330401-00 from the Plant Diversity Research Center of the 21st Century Frontier Research Program, Ministry of Science and Technology; and National Core Research Center Project R15-2003-002-01002-0 from the Korea Science and Engineering Foundation (to D.-J.Y.). A.U. is supported by a Research Fellowship from the Japan Society for the Promotion of Science for Young Scientists.

1. Stockinger, E., Gilmour, S. J. & Thomashow, M. F. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 1035–1040.
2. Liu, Q., Kasuga, M., Sakuma, Y., Abe, H., Miura, S., Yamaguchi-Shinozaki, K. & Shinozaki, K. (1998) *Plant Cell* **10,** 1391–1406.
3. Levy, Y. Y., Mesnage, S., Mylne, J. S., Gendall, A. R. & Dean, C. (2002) *Science* **297,** 243–246.
4. Hugouvieux, V., Kwak, J. M. & Schroeder, J. I. (2001) *Cell* **106,** 477–487.
5. Koiwa, H., Barb, A. W., Xiong, L., Li, F., McCully, M. G., Lee, B.-h., Sokolchik, I., Zhu, J., Gong, Z., Reddy, M., *et al*. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 10893–10898.
6. Xiong, L., Lee, H., Ishitani, M., Tanaka, Y., Stevenson, B., Koiwa, H., Bressan, R. A., Hasegawa, P. M. & Zhu, J.-K. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 10899–10904.
7. Xiong, L., Gong, Z., Rock, C. D., Subramanian, S., Guo, Y., Xu, W., Galbraith, D. & Zhu, J.-K. (2001) *Dev. Cell* **1,** 771–781.
8. Vlachonasios, K. E., Thomashow, M. F. & Triezenberg, S. J. (2003) *Plant Cell* **15,** 626–638.
9. Stockinger, E. J., Mao, Y., Regier, M. K., Triezenberg, S. J. & Thomashow, M. F. (2001) *Nucleic Acids Res.* **29,** 1524–1533.
10. Lin, P. S., Marshall, N. F. & Dahmus, M. E. (2002) *Prog. Nucleic Acid Res. Mol. Biol.* **72,** 333–365.
11. Hartzog, G. A. (2003) *Curr. Opin. Genet. Dev.* **13,** 119–126.
12. Licatalosi, D. D., Geiger, G., Minet, M., Schroeder, S., Cilli, K., McNeil, J. B. & Bentley, D. L. (2002) *Mol. Cell* **9,** 1101–1111.
13. Ahn, S. H., Kim, M. & Buratowski, S. (2004) *Mol. Cell* **13,** 67–76.
14. Komarnitsky, P., Cho, E. J. & Buratowski, S. (2000) *Genes Dev.* **14,** 2452–2460.
15. Ho, C. K. & Shuman, S. (1999) *Mol. Cell* **3,** 405–411.
16. Prelich, G. (2002) *Eukaryotic Cell* **1,** 153–162.
17. Archambault, J., Chambers, R., Kobor, M. S., Ho, Y., Bolotin, D., Andrews, B., Kane, C. M. & Greenblatt, J. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 14300–14305.
18. Archambault, J., Pan, G., Dahmus, G. K., Cartier, M., Marshall, N., Zhang, S., Dahmus, M. E. & Greenblatt, J. (1998) *J. Biol. Chem.* **273,** 27593–27601.
19. Chambers, R. S. & Dahmus, M. E. (1994) *J. Biol. Chem.* **269,** 26243–26248.
20. Cho, H., Kim, T.-K., Mancebo, H., Lane, W. S., Flores, O. & Reinberg, D. (1999) *Genes Dev.* **13,** 1540–1552.
21. Hausmann, S. & Shuman, S. (2002) *J. Biol. Chem.* **277,** 21213–21220.
22. Hausmann, S. & Shuman, S. (2003) *J. Biol. Chem.* **278,** 13627–13632.
23. Kobor, M. S., Archambault, J., Lester, W., Holstege, F. C. P., Gileadi, O., Jansma, D. B., Jennings, E. G., Kouyoumdjian, F., Davidson, A. R., Young, R. A., *et al*. (1999) *Mol. Cell* **4,** 55–62.
24. Kimura, M., Suzuki, H. & Ishihama, A. (2002) *Mol. Cell. Biol.* **22,** 1577–1588.
25. Varon, R., Gooding, R., Steglich, C., Marns, L., Tang, H., Angelicheva, D., Yong, K. K., Ambrugger, P., Reinhold, A., Morar, B., *et al*. (2003) *Nat. Genet.* **35,** 185–189.
26. Yeo, M., Lin, P. S., Dahmus, M. E. & Gill, G. N. (2003) *J. Biol. Chem.* **278,** 26078–26085.
27. Umeda, M., Bhalerao, R. P., Schell, J., Uchimiya, H. & Koncz, C. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 5021–5026.
28. Lin, P. S., Dubois, M. F. & Dahmus, M. E. (2002) *J. Biol. Chem.* **277,** 45949–45956.
29. Abel, S. & Theologis, A. (1994) *Plant J.* **5,** 421–427.
30. Hausmann, S., Schwer, B. & Shuman, S. (2004) *Biochemistry* **43,** 7111–7120.
31. Lee, Y. J., Kim, D. H., Kim, Y. W. & Hwang, I. (2001) *Plant Cell* **13,** 2175–2190.
32. Hausmann, S., Erdjument-Bromage, H. & Shuman, S. (2004) *J. Biol. Chem.* **279,** 10892–10900.
33. Krishnamurthy, S., He, X., Reyes-Reyes, M., Moore, C. & Hampsey, M. (2004) *Mol. Cell* **14,** 387–394.
34. Yu, X., Chini, C. C., He, M., Mer, G. & Chen, J. (2003) *Science* **302,** 639–642.
35. Nguyen, V. T., Kiss, T., Michels, A. A. & Bensaude, O. (2001) *Nature* **414,** 322–325.
36. Yang, Z., Zhu, Q., Luo, K. & Zhou, Q. (2001) *Nature* **414,** 317–322.
37. Michels, A. A., Nguyen, V. T., Fraldi, A., Labas, V., Edwards, M., Bonnet, F., Lania, L. & Bensaude, O. (2003) *Mol. Cell Biol.* **23,** 4859–4869.
38. Palancade, B., Dubois, M. F., Dahmus, M. E. & Bensaude, O. (2001) *Mol. Cell Biol.* **21,** 6359–6368.
39. Chiu, Y. L., Ho, C. K., Saha, N., Schwer, B., Shuman, S. & Rana, T. M. (2002) *Mol. Cell* **10,** 585–597.
40. Ganem, C., Devaux, F., Torchet, C., Jacq, C., Quevillon-Cheruel, S., Labesse, G., Facca, C. & Faye, G. (2003) *EMBO J.* **22,** 1588–1598.