

A novel class of plant-specific zinc-dependent DNA-binding protein that binds to A/T-rich DNA sequences

Yukio Nagano*, Hirofumi Furuhashi, Takehito Inaba and Yukiko Sasaki

Graduate School of Agricultural Sciences, Nagoya University, Nagoya 464-8601, Japan

Received July 27, 2001; Revised and Accepted August 21, 2001

DDBJ/EMBL/GenBank accession no. AB045222

ABSTRACT

Complementary DNA encoding a DNA-binding protein, designated PLATZ1 (plant AT-rich sequence- and zinc-binding protein 1), was isolated from peas. The amino acid sequence of the protein is similar to those of other uncharacterized proteins predicted from the genome sequences of higher plants. However, no paralogous sequences have been found outside the plant kingdom. Multiple alignments among these paralogous proteins show that several cysteine and histidine residues are invariant, suggesting that these proteins are a novel class of zinc-dependent DNA-binding proteins with two distantly located regions, C-x₂-H-x₁₁-C-x₂-C-x₍₄₋₅₎-C-x₂-C-x₍₃₋₇₎-H-x₂-H and C-x₂-C-x₍₁₀₋₁₁₎-C-x₃-C. In an electrophoretic mobility shift assay, the zinc chelator 1,10-*o*-phenanthroline inhibited DNA binding, and two distant zinc-binding regions were required for DNA binding. A protein blot with ⁶⁵ZnCl₂ showed that both regions are required for zinc-binding activity. The PLATZ1 protein non-specifically binds to A/T-rich sequences, including the upstream region of the pea GTPase *pra2* and plastocyanin *petE* genes. Expression of the *PLATZ1* repressed those of the reporter constructs containing the coding sequence of luciferase gene driven by the cauliflower mosaic virus (CaMV) 35S90 promoter fused to the tandem repeat of the A/T-rich sequences. These results indicate that PLATZ1 is a novel class of plant-specific zinc-dependent DNA-binding protein responsible for A/T-rich sequence-mediated transcriptional repression.

INTRODUCTION

DNA-binding proteins mediate important cellular processes, such as DNA replication and the regulation of gene expression. A vast number of genes encoding DNA-binding proteins have

been identified in a wide variety of organisms. Characterization of these proteins has led to the classification of common DNA-binding motifs. Many classes of these proteins have been identified in all kinds of higher organisms. Comparing amino acid sequences with their animal or fungal counterparts has shown that ~55% (about 800 genes) of the *Arabidopsis* transcription factors (about 1500 genes) are not specific to plants (1).

Genome sequences of several model organisms have shown that an increase in biological complexity leads to the appearance of specific groups of DNA-binding proteins. One example is the homeobox family, which is absent in yeast and mold but found in animals and plants. Thus, to meet their various biological requirements, organisms have developed unique mechanisms of transcriptional regulation. Therefore, some classes of DNA-binding proteins appear to be restricted to a specific branch of the evolutionary tree. The nuclear receptor family has, to date, been found only in animals and not in plants. Several classes of DNA-binding proteins have been found only in plants. About 45% of the *Arabidopsis* transcription factors are specific to plants (1). For example, ethylene-responsive element-binding proteins (EREBPs) appear to be specific to plants (2).

Light signals control many aspects of plant development, such as seed germination, hypocotyl growth, leaf expansion and floral induction. Many of these changes are caused by the regulation of gene expression. Therefore, it is important to characterize the *cis*-regulatory elements and the *trans*-acting factors involved in plant-specific light regulation. A dark-inducible and light-repressible small GTPase, *pra2*, is mainly expressed in the growing zone of pea epicotyls (3), and identification of the factors responsible for this gene expression will provide new insights into plant gene expression. We previously identified the 12-bp *cis*-regulatory element DE1, which is necessary for light down-regulated expression of the pea *pra2* gene (4) and is sufficient to confer dark-inducible and light down-regulated expression to a minimal promoter (5). We have recently isolated and characterized a cDNA encoding a DNA-binding protein with specificity for the DE1 element (6). In the course of the study, aimed at identifying genes whose

*To whom correspondence should be addressed. Tel: +81 52 789 4168; Fax: +81 52 789 4296; Email: nagano@agr.nagoya-u.ac.jp

Present address:

Hirofumi Furuhashi, Department of Genetics, The Graduate University for Advanced Studies, Mishima, Shizuoka-ken 411-8540, Japan

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

products bind to the DE1 element, we accidentally isolated the cDNA encoding a novel class of plant-specific zinc-dependent DNA-binding proteins designated as PLATZ1. Although this protein did not bind to the DE1 element, we found that it has several interesting biochemical features and binds to the A/T-rich DNA sequence upstream of the DE1 element. A transient assay showed that the *PLATZ1* gene is responsible for A/T-rich sequence-mediated transcriptional repression.

MATERIALS AND METHODS

Screening of the cDNA library

Total RNA was prepared from the growing zones of pea (*Pisum sativum* var. Alaska) stems (between 0 and 1 cm from the top of the hook) that had been grown in the dark for 7 days at 25°C, as described previously (7). The cDNA library was prepared from this RNA, as described previously (6). The resulting cDNA library contains $\sim 8 \times 10^6$ independent cDNAs. The reporter constructs used in yeast one-hybrid screening contain five and nine tandem repeats of the DE1 and its surrounding sequence, 5'-TGAGGATTTTACAGTAAT-3', upstream of the *HIS3* and *lacZ* reporter genes, respectively (6).

The histidine yeast reporter strain was transformed with a pea cDNA library by the LiAc/polyethylene glycerol method [Technical Tips Online (<http://research.bmn.com/tto>) P01525]. Approximately 1.5×10^7 cDNA plasmids were screened. Based on their large colony size and rapid growth, about 500 histidine-positive clones were selected. Plasmids were recovered and electroporated into the *Escherichia coli* strain DH10B (Life Technologies, Rockville, MD) and then rescreened by transforming the *lacZ* yeast reporter strain. The filter-replica method using X-gal was used to confirm β -galactosidase activities. Three plasmids showed a blue color. One of the resulting plasmids was further characterized. The cDNA insert was subcloned into a pBluescript KS plasmid (Stratagene) and then sequenced. The gene was named *PLATZ1*, and the accession number for the sequence reported in this article is AB045222.

Production of the recombinant PLATZ1 protein

For the production of the glutathione *S*-transferase (GST)-fusion proteins, three fragments (amino acid sequences from 1 to 233, from 1 to 100 and from 102 to 233) of the *PLATZ1* coding region were amplified by PyroBEST DNA polymerase (Takara, Shiga, Japan) using two primers containing the *SalI* and *NotI* sites of their 5' ends. The amplified fragment was subcloned into the *EcoRV* site of the pZER0-2 vector and digested with *SalI* and *NotI*. The DNA fragment was electrophoresed, purified using a DNA extraction kit (Amersham Pharmacia Biotech, Uppsala, Sweden), and subcloned into the *SalI*-*NotI* site of a pGEX-4T-3 vector (Amersham Pharmacia Biotech). Procedures for the production and purification of the fusion protein were carried out as suggested by the manufacturer. The purified protein was dialyzed against the binding buffer, which consisted of 20 mM Tris-HCl pH 8.0, 50 mM KCl, 0.5 mM EDTA, 15 mM MgCl₂, 10% glycerol, 1 mM DTT and 200 μ M ZnCl₂.

Electrophoretic mobility shift assay

An electrophoretic mobility shift assay (EMSA) was performed using a method described previously (4). The fragment

containing the 68-bp DNA sequence (from -734 to -667 of the pea *pra2* transcriptional start site) was used as a probe. The probe was labeled by incubation with [γ -³²P]ATP and T4 polynucleotide kinase. For the enhancer element of the pea plastocyanin gene (*petE*), the 31-bp synthetic oligonucleotide, 5'-AATATACTAGTATTATTACTAAAAAAAATC-3', was labeled by incubation with [γ -³²P]ATP and T4 polynucleotide kinase and hybridized with the complementary synthetic nucleotide. The recombinant proteins were mixed in a binding buffer (20 μ l) containing 2 μ g of poly(dI-dC)-poly(dI-dC), bovine serum albumin (500 μ g/ μ l), and competitor DNA. The protein-DNA complex was formed by incubating this mixture at 25°C for 20 min with the ³²P-labeled probe. The competitor DNAs are shown schematically (see Fig. 3C). These mutated competitors were produced by PCR. Electrophoresis was conducted at 4°C in a 5% polyacrylamide Tris-borate/EDTA gel. The gel was dried and subjected to autoradiography. The image was visualized with a Bio Imaging Analyzer (Model BAS2000, Fujifilm, Tokyo, Japan). The requirement of zinc ions for the binding of PLATZ1 to DNA was tested by the addition of 1,10-*o*-phenanthroline at concentrations from 0 to 5 mM.

Zinc blotting

Zinc blotting was performed by a method described previously (8). Entire or truncated versions of the recombinant PLATZ1 protein were separated by SDS-polyacrylamide gel electrophoresis and electrophoretically transferred to PVDF filters. The filters were preincubated in a metal-binding buffer containing 100 mM Tris-HCl pH 6.8 and 50 mM NaCl for 1 h, and then incubated for 1 h with 4 μ M ⁶⁵ZnCl₂ (63.36 MBq/mg; New England Nuclear Research Products, Wilmington, DE) in 20 ml of the buffer and washed three times with the buffer for 10 min each. In competition experiments, divalent metal ions (10 μ M each) were included in the buffer for all steps.

Transient assay

The method for plasmid construction was described previously (4-6). The expression and reporter constructs used are shown (see Fig. 8A). The expression construct p35S-*PLATZ1* contains the cauliflower mosaic virus (CaMV) 35S promoter and the entire coding region of *PLATZ1* cDNA. A plasmid, p35S, carrying only the 35S promoter was used as a negative control. The reporter constructs *pra2A/T*-p35S90 and *petEA/T*-p35S contain the CaMV 35S90 promoter (a CaMV 35S promoter of 90 bp from the transcriptional start site) and the luciferase (LUC) gene fused to the three-tandem repeat of the A/T-rich sequences upstream of *pra2* and *petE*, respectively. The control reporter construct p35S90 contains the CaMV 35S90 promoter and the LUC gene. A plasmid containing a 35S-*GUS* construct was also included as the internal standard.

The method for particle bombardment was described previously (4). Three plasmids, an expression construct, a reporter construct and an internal standard, were simultaneously transfected into the upper 1-cm region of etiolated pea epicotyls by particle bombardment, and LUC and GUS activities were measured (4,5). All LUC values were normalized to the corresponding GUS values.

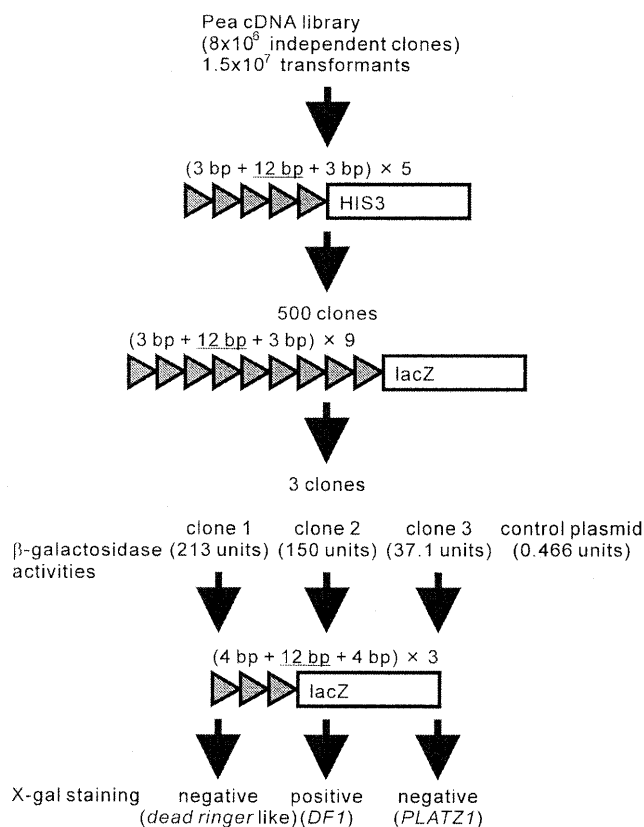


Figure 1. Schematic diagram of one-hybrid screening.

RNA gel blot analysis

RNA gel blot hybridization was performed as described previously (9). Plant materials used in this study were also described elsewhere (9).

RESULTS

Isolation and sequence analysis of a gene encoding a novel DNA-binding protein

The one-hybrid strategy was conducted to identify genes whose products bind to the *cis*-regulatory element DE1 (4,5). The reporter constructs used in this screening contain five and nine tandem repeats of the 18-bp sequence containing the 12-bp DE1 sequence with a neighboring 3 bp at each end, upstream of the *HIS3* and *lacZ* reporter genes, respectively (Fig. 1). From 1.5×10^7 transformants, we selected three clones showing both histidine- and *lacZ*-positive phenotypes. β -galactosidase activities were determined using *o*-nitrophenyl- β -D-galactopyranoside as a substrate (Fig. 1). All candidate clones showed higher activities than the clone containing the control plasmid harboring the *GAL4* gene. We conducted the EMSA using the nine tandem repeats of the 18-bp sequence containing the DE1 sequence as a probe and recombinant proteins from each candidate clone. The EMSA showed that each recombinant protein bound to this probe (data not shown). However, when we conducted the EMSA using the single DE1 sequence as a probe, only the recombinant protein from clone

2 bound to this probe (6). We designated this clone as *DF1* and reported the characterization of this gene previously (6). These EMSA analyses strongly suggest that the recombinant protein from clones 1 or 3 binds only to the tandem copies of the 18-bp sequences. To confirm these possibilities, we conducted a one-hybrid screening using the reporter construct having tandem copies of the 20-bp sequence containing the DE1 sequence with a neighboring 4 bp at each end (Fig. 1). Clones 1 and 3 showed the *lacZ*-negative phenotypes by the conventional X-gal staining. Thus, the recombinant protein from clones 1 or 3 binds only to the tandem copies of the 18-bp sequences.

Although the recombinant protein from clone 3 did not bind to the DE1 sequence, we further characterized this clone because it may encode a new class of zinc-dependent DNA-binding protein. We designated the gene as *PLATZ1* (plant AT-rich sequence- and zinc-binding protein 1). The sequence of *PLATZ1* cDNA is shown in Figure 2A. The cDNA contains an ORF of 233 amino acid residues from the first ATG codon located at nucleotide 29 to a termination codon located at nucleotide 728. The polypeptide predicted from the first ATG codon has a calculated relative molecular mass of 26.7 kDa. Coupled *in vitro* transcription and translation of the *PLATZ1* cDNA produced a protein with a molecular mass of ~27 kDa (data not shown). In addition, the upstream region of this putative start codon did not show any sequence similarities to that of the putative *Arabidopsis* ortholog inferred from the genome sequence (accession no. AC000107). Therefore, based on the molecular weight of the *in vitro*-translated product, the sequence similarities and the Kozak rule (10) for the translation-start site, the first ATG codon is most likely the translation-start site.

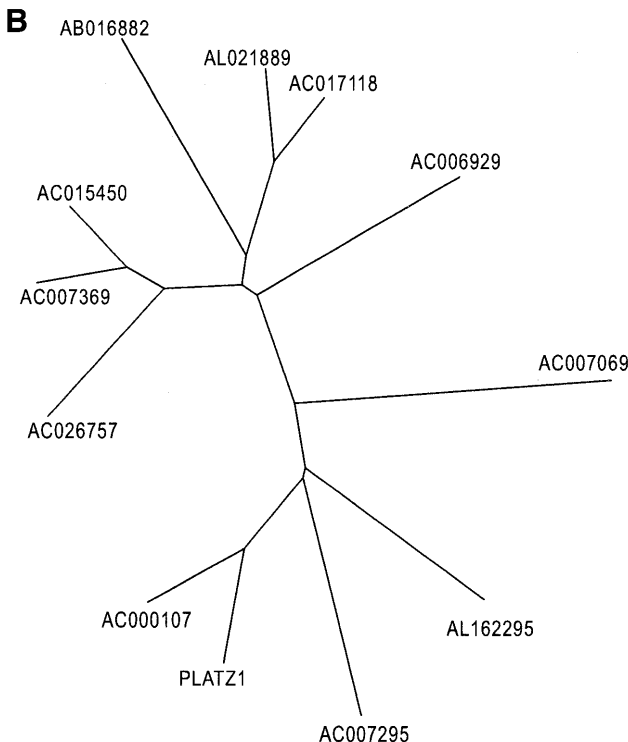
In a BLAST or PSI-BLAST search (11), *PLATZ1* did not show any sequence similarities to the previously reported DNA-binding proteins but showed similarities to 11 uncharacterized proteins predicted from genome sequences of *Arabidopsis*. In addition, *PLATZ1* is similar to proteins encoded by expressed sequence tags isolated from such higher plants as rice, maize, tomato, soybean, pine and poplar (data not shown). These similarities are restricted to sequences of plant species. Figure 2B shows a phylogenetic tree constructed by the neighbor-joining method (12). The tree shows that *PLATZ1* is closely related to the *Arabidopsis* protein predicted from the genome sequence AC000107. The pea *PLATZ1* shows 61% identity and 73% similarity to this *Arabidopsis* protein. Based on the conservation and the fact that the *Arabidopsis* genome was sequenced almost completely (13), this *Arabidopsis* protein could be the ortholog of *PLATZ1*.

Figure 2C shows multiple alignments between *PLATZ1* and predicted *Arabidopsis* proteins. Upon closer inspection of the aligned sequences, it can be seen that five cysteine and three histidine residues are invariant at the N-terminal region (Fig. 2C, shown in gray) and that the consensus signature of this conserved region is C-x₂-H-x₁₁-C-x₂-C-x₍₄₋₅₎-C-x₂-C-x₍₃₋₇₎-H-x₂-H. The conservation of cysteine and histidine residues suggests that these proteins encode zinc-binding proteins (14). This signature is different from that of previously characterized zinc-binding motifs. However, the number of invariant cysteine and histidine residues (CHC₄H₂) is identical to that of the 'double-zinc-finger' domains (15), such as RING (C₃HC₄) and LIM (C₂HC₅). Therefore, this motif may be a new type of

A

```

1 ATCATCATCATCAAGAGGTTGGGGAGAGATGGACATAGACATGAACATGAACATGAACAT 60
   M D I D M N M N M N M
12 GATGAATATGAACCTGCATCGGTAGAGAGGGTTGATGGATGAAACATCTCTCGGAGATTC
   M N M K P A W I E R L M D E T F F G D C
32 TGGGTCACAAAAATCGAAGCAAGAACAGAAACATATTTTTCCTACACTGTTGCTCT
   G V H K N R R K N E K N I F C L H C C L
52 TAGCATCTGCCCCACTGCTTATCTCTCATACCTTCATCTCTCTCTCCAGGTCGAGAG
   S I C P H C L S S H T S H P L L Q V R R
72 GTATGTTTACCACAAATGTAATTCGATGGATGATCGGAAAGCTCATAGATTTCCCAA
   Y V Y H N V I R L D D L E K I I D C S Y N
92 TATTCAGCCCTATACATTAACAGTCGAAAGTGATATTTTAAACAGAGGGCCACAATC
   I Q P Y T I N S A K A V I F L N Q R P Q S
112 AAGGCATGCAAAAGTTACAGCCAAATGTTGCTTCACTTGTGACAGAAATCTCTCAGGATCC
   R S C K V T A N V C F T C D R I L Q D P
132 ATTCACATTTCTCTCTCTCTCATCAAGGTAGAAATACATGGTGTATGAGGGTGAATAAT
   F H F C S L S L S K V E M V Y E G E N L
152 GTCGAACATTTACATCGATTTGATGAATCGGACCTTGCAATATCGCAATTCGAGGGACT
   S N I L H R F D E S D F A I S Q F E G L
172 AAGACTGGATGCACTGAGCTAACTGATGAAGATAAATCAGATTGGCCAAAGTCTCTC
   R V D G S E V I D E D N N Q I G P S S S
192 AAATTAATCCATCATCAATGAAGATCAGGCCACTCAGCAACATAGATTCCTCCCTCTCC
   N Y S I I N E D Q A T S N N R F L P L P
212 TCTCCCTGCCATTTGCTCTCTCTCTGAGTACGAAAGAAAGGCTCTCTCAAGAGGCCCC
   L P A I V L S L I G S R R K G A L P Q R A P
232 ACTCTCCTAATTATCTCTTCCATCAATTTTATATATATAATAATCTCTCTCTCA
   L S
840 TCTTTTTCATCGCTAAATAAGAAATTTCTTCACTACTTTTATTAAGTATATGTTATCTT
900 CTAAGATTATGCTATGCTATGCTGATGGATGGTTTCAACACCAACTTTGTTGAAAA
920 AAAAAAAAAAAAAAAAAAAAA
    
```



C

```

20 40 60 80
PLATZ1 PAWLERLMDTEFFGDCGVZNRNRNENKTFGLHC-LSLCPHCLP---SHRSHPQLQVRRVYVNVIRLD
AC000107 PAWLEGLMAE7FFSSEGLHETRRRSKNVFLDCC-LSVCPHCLP---SHRSHPQLQVRRVYVNVIRLD
AC007295 PAWLALYAEKFFVQCYEYETAKRNERNVCCDCC-LSLCPHCLP---SHRFHRLQVRRVYVNVIRLD
AL162295 PAWLELLKDFEFNAELDHDDEKNERNLCCDCC-LIICPHCLP---SHRSHPQLQVRRVYVNVIRLD
AC007069 EWELELLKDFEFNAELDHDDEKNERNLCCDCC-LIICPHCLP---SHRSHPQLQVRRVYVNVIRLD
AC007369 EWELELLKDFEFNAELDHDDEKNERNLCCDCC-LIICPHCLP---SHRSHPQLQVRRVYVNVIRLD
AC015450 EWELELLKDFEFNAELDHDDEKNERNLCCDCC-LIICPHCLP---SHRSHPQLQVRRVYVNVIRLD
AC007369 EWELELLKDFEFNAELDHDDEKNERNLCCDCC-LIICPHCLP---SHRSHPQLQVRRVYVNVIRLD
AC026757 EWELELLKDFEFNAELDHDDEKNERNLCCDCC-LIICPHCLP---SHRSHPQLQVRRVYVNVIRLD
AC007369 EWELELLKDFEFNAELDHDDEKNERNLCCDCC-LIICPHCLP---SHRSHPQLQVRRVYVNVIRLD
AL021889 EWELELLKDFEFNAELDHDDEKNERNLCCDCC-LIICPHCLP---SHRSHPQLQVRRVYVNVIRLD
AC017118 EWELELLKDFEFNAELDHDDEKNERNLCCDCC-LIICPHCLP---SHRSHPQLQVRRVYVNVIRLD
AB016882 EWELELLKDFEFNAELDHDDEKNERNLCCDCC-LIICPHCLP---SHRSHPQLQVRRVYVNVIRLD
AC006929 EWELELLKDFEFNAELDHDDEKNERNLCCDCC-LIICPHCLP---SHRSHPQLQVRRVYVNVIRLD
100 120 140
PLATZ1 DLERLDCSNQPYVINSARVFLNRPQSRSC-----KVTANVPTCDRLDLD-PFHFCLSLQKV
AC000107 DLERLDCSNQPYVINSARVFLNRPQSRSC-----KVTANVPTCDRLDLD-PFHFCLSLQKV
AC007295 DLERLDCSNQPYVINSARVFLNRPQSRSC-----KVTANVPTCDRLDLD-PFHFCLSLQKV
AL162295 DLERLDCSNQPYVINSARVFLNRPQSRSC-----KVTANVPTCDRLDLD-PFHFCLSLQKV
AC007069 DLERLDCSNQPYVINSARVFLNRPQSRSC-----KVTANVPTCDRLDLD-PFHFCLSLQKV
AC015450 DLERLDCSNQPYVINSARVFLNRPQSRSC-----KVTANVPTCDRLDLD-PFHFCLSLQKV
AC026757 DLERLDCSNQPYVINSARVFLNRPQSRSC-----KVTANVPTCDRLDLD-PFHFCLSLQKV
AL021889 DLERLDCSNQPYVINSARVFLNRPQSRSC-----KVTANVPTCDRLDLD-PFHFCLSLQKV
AC017118 DLERLDCSNQPYVINSARVFLNRPQSRSC-----KVTANVPTCDRLDLD-PFHFCLSLQKV
AB016882 DLERLDCSNQPYVINSARVFLNRPQSRSC-----KVTANVPTCDRLDLD-PFHFCLSLQKV
AC006929 DLERLDCSNQPYVINSARVFLNRPQSRSC-----KVTANVPTCDRLDLD-PFHFCLSLQKV
    
```

double-zinc-finger motif. In addition to the N-terminal zinc-binding motif, four cysteine residues are conserved at the central region (Fig. 2C, shown in gray), and the consensus signature of this conserved region is C-x₂-C-x₍₁₀₋₁₁₎-C-x₃-C. This signature is somewhat similar to that of the GATA finger (C₂C₂). However, because the consensus sequence of the GATA finger is C-x₂-C-x₍₁₇₋₁₈₎-C-x₂-C, the GATA finger does not fit this motif. Thus, these two putative zinc-binding motifs are not found in the DNA-binding proteins characterized to date. The sequence comparison suggests that the PLATZ1 protein and its paralogous proteins encode a novel class of zinc-dependent DNA-binding protein.

PLATZ1 is a DNA-binding protein that non-specifically binds to A/T-rich DNA sequences

To test whether the PLATZ1 protein is a DNA-binding protein, we conducted an EMSA. The entire coding region of *PLATZ1* was expressed in *E. coli* as a GST-fusion protein. As mentioned above, the recombinant PLATZ1 protein was able to bind to the tandem repeats of DE1 elements (data not shown) but not to the 31-bp DNA fragment (Fig. 3A, underlined) containing the single DE1 element (data not shown). Surprisingly, the EMSA showed that the PLATZ1 protein binds to the 68-bp DNA fragment (Fig. 3A) located at the upstream region of the DE1 element (Fig. 3B, lanes 2 and 3). In competition experiments using the 68-bp unlabeled DNA, increasing amounts of competitor gradually diminished the retarded band (Fig. 3B, lanes 4–6). We further conducted additional competition experiments using various competitors (Fig. 3C) to determine the DNA sequence bound to the PLATZ1 protein. The addition of a wild-type competitor and 400-fold competitors LS6, LS7 and LS8 diminished the retarded bands (Fig. 3D, lanes 3 and 9–11). In contrast, the addition of 400-fold competitors LS1–LS5 diminished the bands even slightly (Fig. 3D, lanes 4–8). Thus, the DNA sequence bound to the PLATZ1 protein is an A/T-rich 30-bp or smaller region, 5'-AACACATATTTTGA-TAAATTTATTACTAAA-3'.

A sequence-similarity search showed that the 30-bp A/T-rich sequence is weakly similar to the 31-bp enhancer element of the pea plastocyanin gene, *petE* (16). Figure 4A shows the similarity between these two sequences. To test whether the PLATZ1 protein binds to the 31-bp A/T-rich enhancer, we conducted an EMSA using this sequence and the recombinant PLATZ1 protein. The recombinant PLATZ1 protein was able to bind to the 31-bp *petE* enhancer sequence (Fig. 4B, lanes 2 and 3). In competition experiments using the unlabeled 31-bp DNA fragment, increasing amounts of the competitor sequence gradually diminished the retarded band (Fig. 4B,

Figure 2. (A) Nucleotide and amino acid sequences of the *PLATZ1* gene. Nucleotides are numbered on the right side, and amino acids are numbered on the left side. (B) Phylogenetic tree of *PLATZ1* and its paralogous *Arabidopsis* proteins. (C) Amino acid sequence alignments of *PLATZ1* and its paralogous *Arabidopsis* proteins. Several *Arabidopsis* proteins were predicted from genome sequences. The left column shows the accession number of the genome sequence. Conserved cysteine and histidine residues are shown in gray. Asterisks indicate positions which have a single, fully conserved residue. Colons indicate positions that have similar residues.

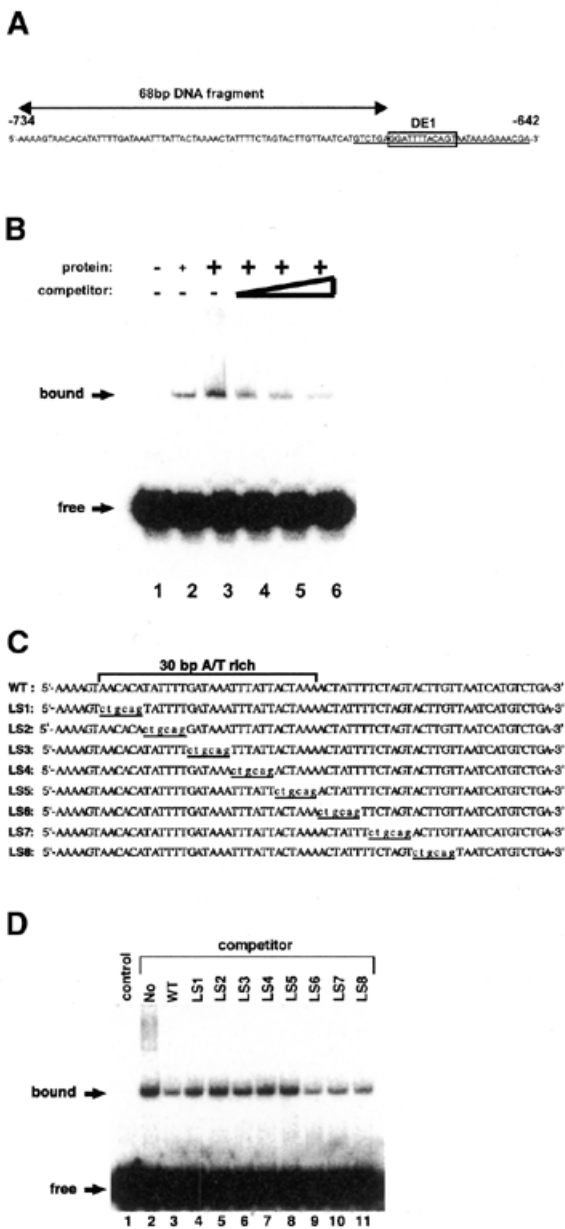


Figure 3. DNA-binding activity of the PLATZ1 protein. (A) Nucleotide sequence of the upstream region (-734 to -642) of the pea *pra2* gene. The DE1 element is boxed. Arrows indicate the 68-bp DNA probe used in the EMSA experiment. The 31-bp DE1 probe is underlined. (B) The recombinant PLATZ1 protein binds to the 68-bp DNA fragment. Binding reactions were carried out with the recombinant PLATZ1 protein and the ³²P-labeled 68-bp DNA fragment (2.5 fmol) in the absence or presence of the unlabeled 68-bp DNA fragment as competitor. Lane 1, no protein; lane 2, 100 ng PLATZ1 protein; lanes 3-6, 200 ng PLATZ1 protein. As indicated, 100, 200 and 400-fold competitors were added (lanes 4, 5 and 6). Arrows indicate the free probe and the protein-bound complex. (C) Nucleotide sequences of various competitors used to define the PLATZ1-binding site. The positions of the mutated nucleotides are underlined. (D) Competition experiments using the 68-bp WT competitor and LS-mutated competitors. Binding reactions were carried out with 200 ng of the PLATZ1 protein and the ³²P-labeled 68-bp DNA fragment (2.5 fmol) in the absence or presence of the unlabeled competitor DNA. A 400-fold competitor was added in each reaction. The control does not contain added protein or competitor DNA.

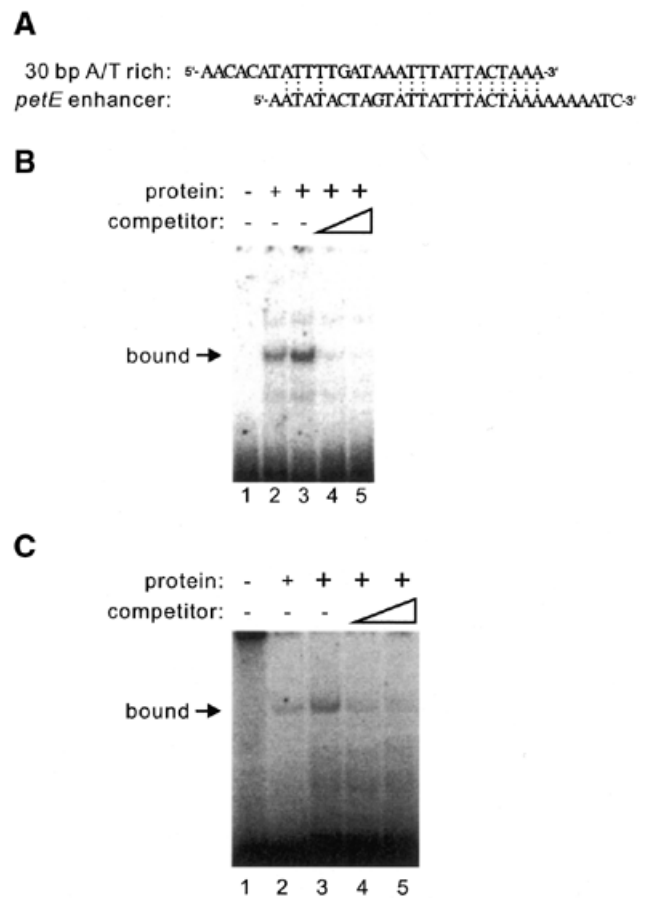


Figure 4. The PLATZ1 protein binds to the enhancer element of the pea phytoalexin gene (*petE*). (A) The sequence similarity between the 30-bp A/T-rich sequence of the *pra2* gene and the 31-bp A/T-rich enhancer of the pea *petE* gene. (B) EMSA using the 31-bp A/T-rich enhancer of the pea *petE* gene. Binding reactions were carried out with the recombinant PLATZ1 protein and the ³²P-labeled 31-bp A/T-rich enhancer fragment (100 fmol) in the absence or presence of the unlabeled 31-bp A/T-rich enhancer fragment as competitor. Lane 1, no protein; lane 2, 100 ng PLATZ1 protein; lanes 3-5, 200 ng PLATZ1 protein. As indicated, 100 and 200-fold competitors were added (lanes 4 and 5). An arrow indicates a protein-bound complex. (C) Competition experiments using oligonucleotides poly(dA-dT)-poly(dA-dT) and the labeled enhancer fragment of the pea *petE* gene. Lane 1, no protein; lane 2, 100 ng PLATZ1 protein; lanes 3-5, 200 ng PLATZ1 protein. As indicated, 50 ng and 100 ng of competitors were added (lanes 4 and 5). An arrow indicates a protein-bound complex.

lanes 4 and 5), indicating that the PLATZ1 protein binds to another type of A/T-rich DNA sequence.

The above analyses showed that the PLATZ1 protein binds to two types of A/T-rich DNA sequences, the upstream region of the pea GTPase *pra2* gene and the enhancer element of the pea *petE* gene. The question raised by these studies is whether the PLATZ1 protein binds specifically to these two types of A/T-rich DNA sequences or non-specifically to A/T-rich DNA sequences. To address this question, we added oligonucleotides poly(dA-dT)-poly(dA-dT) to the EMSA reaction mixture containing 2 ng of the labeled *petE* enhancer fragment, 100 or

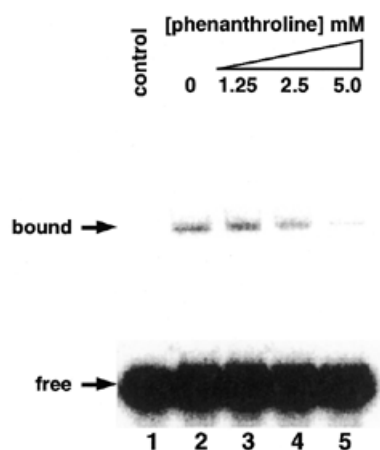


Figure 5. Inhibition of the DNA-binding activity of the PLATZ1 protein by the zinc chelator 1,10-*o*-phenanthroline. Binding reactions were carried out with the PLATZ1 protein (200 ng) and the 32 P-labeled 68-bp DNA fragment (2.5 fmol) in the presence of 1,10-*o*-phenanthroline (1.25–5 mM). Lane 1, no protein; lane 2, no 1,10-*o*-phenanthroline; lanes 3–5, 1.25, 2.5 and 5.0 mM 1,10-*o*-phenanthroline, respectively. Arrows indicate the free probe and the protein-bound complex.

200 ng of the recombinant PLATZ1 protein, and 2 μ g of poly(dI-dC)–poly(dI-dC). The addition of only 50 ng of poly(dA-dT)–poly(dA-dT) diminished the retarded band markedly (Fig. 4C, lane 4). In addition, increasing amounts of the poly(dA-dT)–poly(dA-dT) gradually diminished the retarded band (Fig. 4C, lanes 4 and 5). Thus, the PLATZ1 protein non-specifically binds to A/T-rich DNA sequences.

PLATZ1 is a zinc-dependent DNA-binding protein with two distant regions

To test whether PLATZ1 is a zinc-dependent DNA-binding protein, we added the zinc chelator 1,10-*o*-phenanthroline to the binding reaction. Increasing amounts of 1,10-*o*-phenanthroline gradually inhibited the DNA binding (Fig. 5, lanes 3–5). The results may implicate zinc ion as a cofactor for this DNA-protein complex.

Because these experiments do not exclude the possibility that the PLATZ1 protein binds metal ions other than zinc ions, we used the zinc-blotting technique to test whether the PLATZ1 protein binds zinc ions (8). A protein blot with 65 ZnCl₂ showed that the PLATZ1 protein has the ability to bind zinc ions (Fig. 6). To test whether the PLATZ1 protein binds zinc ions specifically, we conducted competition experiments using various metal ions (Fig. 6). The addition of cold 10 μ M ZnCl₂ diminished the binding signal. In contrast, the addition of 10 μ M CoCl₂, FeCl₂, MnCl₂ or CuCl₂ diminished the signal slightly. These specificities are similar to those of the other zinc-binding proteins (8). Thus, the PLATZ1 protein binds zinc ions specifically.

The consensus sequence in the N-terminal putative zinc-binding motif is C-x₂-H-x₁₁-C-x₂-C-x₍₄₋₅₎-C-x₂-C-x₍₃₋₇₎-H-x₂-H (Fig. 7A). The other consensus sequence in the central region is C-x₂-C-x₍₁₀₋₁₁₎-C-x₃-C (Fig. 7A). The intervening sequence between these two motifs is 56–65 amino acids long. Therefore, these two distant regions may be distinct domains. Alternatively, both regions may be required for zinc binding.

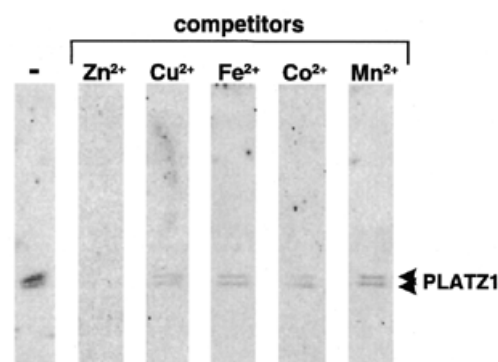


Figure 6. Zinc-binding activity of the PLATZ1 protein. The PLATZ1 protein (1 μ g) was separated by SDS-PAGE and transferred to PVDF filters. The filters were probed with 65 ZnCl₂ in the metal-binding buffer. In competition experiments, divalent metal ions (ZnCl₂, CuCl₂, FeCl₂, CoCl₂ and MnCl₂) were included in the buffer for all steps. Arrowheads indicate the position of the PLATZ1 protein.



Figure 7. Two distantly located zinc-binding motifs are required both for zinc binding and DNA binding. (A) Schematic representation of three recombinant proteins. (B) Zinc-binding activity of the PLATZ1 protein and its truncated versions, PLATZ1-N and PLATZ1-C. Each protein (~1 μ g) was separated by SDS-PAGE and transferred to PVDF filters. One filter was stained with Coomassie Blue (right panel), and the other was probed with 65 ZnCl₂ (left panel). Arrowheads indicate the position of each protein. M, molecular weight markers. (C) DNA-binding activity of the N-terminal (PLATZ1-N) and central regions (PLATZ1-C) of PLATZ1 proteins. The DNA-binding activity of each protein was compared with the binding activity of the entire region of PLATZ1 (PLATZ1). Binding reactions were carried out with different amounts (100 and 200 ng) of each fusion protein and the 32 P-labeled 68-bp DNA fragment (2.5 fmol). The control contains no added protein.

To check these two possibilities, we expressed the two regions separately and conducted zinc blotting. The recombinant

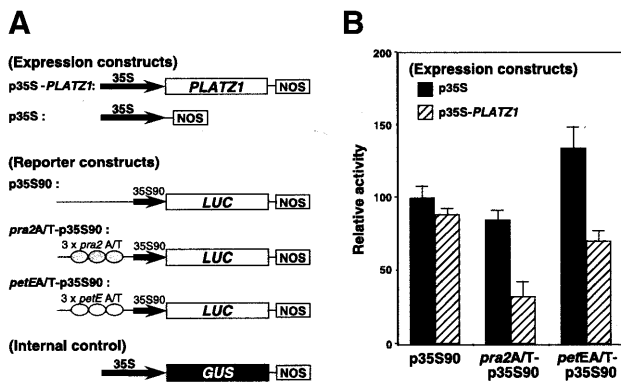


Figure 8. *PLATZ1* represses gene expression in a binding site-dependent manner. (A) Schematic diagrams of the *PLATZ1* expression constructs and LUC reporter constructs used in the transient assay. NOS, nopaline synthase terminator. (B) The indicated reporter and expression constructs were co-bombarded into the growing region of etiolated pea epicotyls. Twelve hours after bombardment, LUC and GUS activities were measured. LUC activity was normalized with respect to GUS activity for each bombardment experiment, and the average of p35S90 bombarded with p35S was taken as 100. Values are the means of at least seven independently bombarded samples, with error bars representing S.E.

proteins used in this study are schematically shown in Figure 7A. The N-terminal (from 1 to 100) and C-terminal (from 102 to 233) regions were expressed as fusion proteins with GST using the pGEX-4T-3 vector and were designated as PLATZ1-N and PLATZ1-C, respectively. We observed two bands of PLATZ1 or PLATZ1-C proteins on SDS-polyacrylamide electrophoresis (Fig. 7B). The degradation of the C-terminus of PLATZ1 possibly causes the appearance of two bands. Zinc blotting showed that both PLATZ1-N and PLATZ1-C proteins bind zinc ions slightly (Fig. 7B), indicating that both zinc-binding motifs are required for efficient zinc-binding activity.

To check whether efficient zinc-binding activity is required for efficient DNA binding, we conducted an EMSA using a labeled 68-bp DNA fragment and truncated versions of the PLATZ1 protein, PLATZ1-N and PLATZ1-C. The addition of the entire length of the PLATZ1 protein to the EMSA reaction resulted in the formation of a retarded band (Fig. 7C, lanes 2 and 3). In contrast, neither PLATZ1-N nor PLATZ1-C alone formed a retarded band (Fig. 7C, lanes 4–7). These results suggest that the loss of zinc-binding activities of these truncated proteins resulted in the loss of DNA-binding activities. Thus, the two distantly located zinc-binding motifs are required for both DNA binding and zinc binding.

Expression of the *PLATZ1* gene represses A/T-rich sequence-mediated transcription in a transient assay

We previously established a particle bombardment procedure using etiolated pea stems (4,5). Using this method, the expression of reporter constructs, containing the coding sequence of luciferase gene driven by CaMV 35S90 promoter fused to the tandem repeat of the A/T-rich sequences, was studied in bombarded stems in the presence of an expressed *PLATZ1* gene. Figure 8A shows the expression and reporter constructs used.

Cotransformation of the control expression construct p35S and the reporter construct *pra2A/T*-p35S90 did not increase

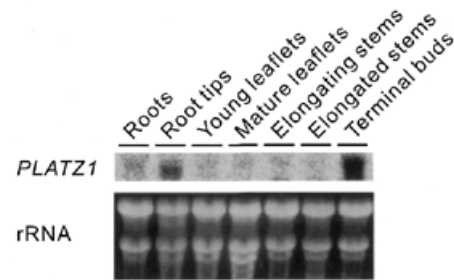


Figure 9. RNA gel blotting of *PLATZ1* mRNA in various organs. Roots, root tips (terminal 5 mm) and terminal buds were collected from 7-day-old pea seedlings. Mature and young leaflets were collected from the third and fifth nodes of 7-day-old plants, respectively. Elongated and elongating stems were collected from the second and fourth internodes of 10-day-old seedlings.

the level of LUC activity compared with that transformed with p35S and the control reporter construct p35S90 (Fig. 8B). Cotransformation of the *PLATZ1*-expressed construct p35S-*PLATZ1* and the reporter construct *pra2A/T*-p35S90 resulted in ~60% decrease in LUC activity compared with that transformed with p35S and the reporter construct p35S90 or *pra2A/T*-p35S90 (Fig. 8B). These results show that the expression of *PLATZ1* can repress gene expression in a binding site-dependent manner.

The results using another reporter construct, *petEA/T*-p35S90, are more complicated. Cotransformation of the control expression construct p35S and the reporter construct *petEA/T*-p35S90 resulted in ~35% increase in LUC activity compared with that transformed with p35S and the control reporter construct p35S90 (Fig. 8B). This A/T-rich sequence is known to be an enhancer element of the pea *petE* gene (16). High mobility group proteins HMG-1 and HMG-1/Y bind to this enhancer sequence (17) and possibly act as a transcription factor responsible for enhanced expression of the *petE* gene. Therefore, endogenous HMG proteins possibly caused this increase. Cotransformation of the *PLATZ1*-expressed construct p35S-*PLATZ1* and the reporter construct *petEA/T*-p35S90 resulted in ~50% decrease in LUC activity compared with that transformed with p35S and the reporter construct *petEA/T*-p35S90 (Fig. 8B). These results again show that the expression of *PLATZ1* can repress gene expression in a binding site-dependent manner.

Expression of the *PLATZ1* mRNA in various organs

To identify an organ-specific expression of the *PLATZ1* gene, we performed gel blot analysis (9) of RNA from roots, root tips, young (expanding) leaflets, mature (expanded) leaflets, elongating stems, elongated stems and terminal buds (Fig. 9). The levels of expression of *PLATZ1* mRNA were much higher in root tips and terminal buds. In contrast, the *PLATZ1* mRNA was barely detectable in the other organs. Thus, the *PLATZ1* mRNA was possibly expressed in the organs undergoing cell division.

DISCUSSION

This paper describes the identification and characterization of a novel class of zinc-dependent DNA-binding proteins that bind to A/T-rich sequences. This class of zinc-dependent

DNA-binding proteins consists of the N-terminal putative zinc-binding motif, C-x₂-H-x₁₁-C-x₂-C-x₍₄₋₅₎-C-x₂-C-x₍₃₋₇₎-H-x₂-H, the 56–65 amino acid intervening sequence, and the central putative zinc-binding motif, C-x₂-C-x₍₁₀₋₁₁₎-C-x₃-C. The presence of distantly located zinc-binding motifs suggests that these two regions are distinct domains. However, the current study shows that two regions are required both for zinc binding and DNA binding (Fig. 7). Thus, a relatively long region is involved in zinc binding and DNA binding. To our knowledge, distantly located zinc-binding regions have not yet been observed in the other zinc-dependent DNA-binding proteins. Therefore, PLATZ1 and its related proteins may have novel mechanisms for zinc binding and DNA binding. Determining the mode of action of two distant regions is an important but challenging problem.

No more than 22% of *Arabidopsis* transcription factors are zinc-coordinating proteins (1). In contrast, most of the transcription factors are zinc-coordinating proteins in the three other eukaryotes whose genomes have been sequenced almost completely: 51% in *Drosophila melanogaster*, 64% in *Caenorhabditis elegans* and 56% in yeast (1). However, here we added a novel class of zinc-dependent DNA-binding proteins in plants. Therefore, future studies of plant transcription factors might find additional classes of zinc-coordinating proteins.

Several classes of DNA-binding proteins have been found only in plants (1). These include the EREBP/AP2 (2), WRKY (18), NAC (19), VP1/ABI3/ARF (20) and other smaller families. For example, ethylene-responsive element-binding proteins and their related proteins, such as AP2, appear to be specific for plants. The solution structure of AtERF1, one of the EREBPs, has a novel mode of DNA recognition by a β -sheet (2). The WRKY proteins are a superfamily of transcription factors with up to 100 representatives in *Arabidopsis*. These proteins have, to date, been found only in plants (18). Members of this family have the conserved amino acid sequence WRKYGQK, together with conserved cysteine and histidine residues. Zinc chelator 1,10-*o*-phenanthroline abolishes *in vitro* DNA binding, suggesting that the WRKY proteins are DNA-binding proteins with a novel zinc-finger-like motif. PLATZ1 and its related proteins are different from these plant-specific DNA-binding proteins. Thus, we have found a novel plant-specific DNA-binding protein. However, some DNA-binding proteins classified as specific to plants might be related to proteins found in other organisms. For example, the trihelix DNA-binding proteins are classified as specific to plants (1). However, they are found in *D.melanogaster*. In addition, they could be distantly related to the MYB proteins, found in all eukaryotes (21). Therefore, future studies might reveal that PLATZ1 and its related proteins are not specific to plants.

The EMSA showed that the PLATZ1 protein non-specifically binds to A/T-rich sequences, including the upstream region of the pea GTPase *pra2* gene and the enhancer element of the pea *petE* gene (Figs 3 and 4). Transient assays showed that the PLATZ1 gene is able to repress the expression from these two A/T-rich sequences *in planta* (Fig. 8B). These results indicate that PLATZ1 is a DNA-binding protein responsible for A/T-rich sequence-mediated transcriptional repression. However, the interpretation of the case of the *petE* enhancer is more complicated. Endogenous HMG proteins are possible DNA-

binding proteins responsible for *petE* enhancer-mediated transcriptional activation (17). The transient assay showed that the PLATZ1 gene represses this enhancer function. Thus, multiple proteins are involved in *petE* enhancer-mediated transcriptional regulation. These DNA-binding proteins may independently act as transcriptional regulators. Alternatively, these proteins may mediate so-called 'context-dependent' transcriptional regulation (22). Several possible models of the actions are as follows: (i) PLATZ1 triggers the alteration of the local chromatin structure constructed by HMG-I/Y or histone H1 (23), consequently influencing the expression from nearby promoters; (ii) overlapping binding sites influence the response to DNA-binding protein(s); and (iii) sequence-specific protein-protein interactions influence the response to transcription factor(s).

The PLATZ1 mRNA was possibly expressed in the organs undergoing cell division. Although still speculative, this suggests that the function of the PLATZ1 gene is related to cell division or regulation of differentiation in cells undergoing cell division. The cell cycle-regulated gene expression of the PLATZ1 mRNA has not been detected (unpublished results). Therefore, it is important to elucidate the precise physiological and cellular function of the PLATZ1 protein. The A/T-rich DNA sequence of the *petE* gene is known to act as an enhancer element (16). The current study shows that PLATZ1 negatively regulates this enhancer function. Thus, the cellular role of PLATZ1 could be complicated. However, the complicated mechanisms of transcriptional regulation are frequently observed in many genes (22). Possibly, multiple DNA-binding proteins, including the PLATZ1 protein, regulate gene expression cooperatively. Overexpression or underexpression of PLATZ1 may allow us to determine the cellular role of the PLATZ1 gene in plant growth and development.

ACKNOWLEDGEMENTS

We are grateful for the use of the facility at the Radioisotope Research Center of Nagoya University. This work was supported by Grants-in-Aid for the Encouragement of Young Scientists (no. 12760225) and for Priority (no. 13017210) to Y.N. from the Japanese Ministry of Education, Culture, Science, Sports, and Technology. This work was also supported by the Japan Society for the Promotion of Science, Research for the Future Program Grants no. JSPS-RTFT.96L006012 to Y.S. Research fellowships from the Japan Society for the Promotion of Science for Young Scientists were received by T.I.

REFERENCES

- Riechmann, J.L., Heard, J., Martin, G., Reuber, L., Jiang, C., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O.J., Samaha, R.R., Creelman, R., Pilgrim, M., Broun, P., Zhang, J.Z., Ghandehari, D., Sherman, B.K. and Yu, G. (2000) *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. *Science*, **290**, 2105–2110.
- Allen, M.D., Yamasaki, K., Ohme-Takagi, M., Tateno, M. and Suzuki, M. (1998) A novel mode of DNA recognition by a beta-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *EMBO J.*, **17**, 5484–5496.
- Nagano, Y., Okada, Y., Narita, H., Asaka, Y. and Sasaki, Y. (1995) Location of light-repressible, small GTP-binding protein of the YPT/rab family in the growing zone of etiolated pea stems. *Proc. Natl Acad. Sci. USA*, **92**, 6314–6318.

4. Inaba, T., Nagano, Y., Sakakibara, T. and Sasaki, Y. (1999) Identification of a *cis*-regulatory element involved in phytochrome down-regulated expression of the pea small GTPase gene *pra2*. *Plant Physiol.*, **120**, 491–499.
5. Inaba, T., Nagano, Y., Reid, J.B. and Sasaki, Y. (2000) DE1, a 12-base pair *cis*-regulatory element sufficient to confer dark-inducible and light down-regulated expression to a minimal promoter in pea. *J. Biol. Chem.*, **275**, 19723–19727.
6. Nagano, Y., Inaba, T., Furuhashi, H. and Sasaki, Y. (2001) Trihelix DNA-binding protein with specificities for two distinct *cis*-elements, both important for light down-regulated and dark-inducible gene expression in higher plants. *J. Biol. Chem.*, **276**, 22238–22243.
7. Nagano, Y., Murai, N., Matsuno, R. and Sasaki, Y. (1993) Isolation and characterization of cDNAs that encode eleven small GTP-binding proteins from *Pisum sativum*. *Plant Cell Physiol.*, **34**, 447–455.
8. Schiff, L.A., Nibert, M.L. and Fields, B.N. (1988) Characterization of a zinc blotting technique: evidence that a retroviral gag protein binds zinc. *Proc. Natl Acad. Sci. USA*, **85**, 4195–4199.
9. Madoka, Y. and Mori, H. (2000) Two novel transcripts expressed in pea dormant axillary buds. *Plant Cell Physiol.*, **41**, 274–281.
10. Kozak, M. (1999) Initiation of translation in prokaryotes and eukaryotes. *Gene*, **234**, 187–208.
11. Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
12. Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, **4**, 406–425.
13. The *Arabidopsis* Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, **408**, 796–815.
14. Klug, A. and Schwabe, J.W. (1995) Protein motifs 5. Zinc fingers. *FASEB J.*, **9**, 597–604.
15. Schwabe, J.W. and Klug, A. (1994) Zinc mining for protein domains. *Nature Struct. Biol.*, **1**, 345–349.
16. Sandhu, J.S., Webster, C.I. and Gray, J.C. (1998) A/T-rich sequences act as quantitative enhancers of gene expression in transgenic tobacco and potato plants. *Plant Mol. Biol.*, **37**, 885–896.
17. Webster, C.I., Packman, L.C., Pwee, K.H. and Gray, J.C. (1997) High mobility group proteins HMG-1 and HMG-I/Y bind to a positive regulatory region of the pea plastocyanin gene promoter. *Plant J.*, **11**, 703–715.
18. Eulgem, T., Rushton, P.J., Robatzek, S. and Somssich, I.E. (2000) The WRKY superfamily of plant transcription factors. *Trends Plant Sci.*, **5**, 199–206.
19. Xie, Q., Sanz-Burgos, A.P., Guo, H., Garcia, J.A. and Gutierrez, C. (1999) GRAB proteins, novel members of the NAC domain family, isolated by their interaction with a geminivirus protein. *Plant Mol. Biol.*, **39**, 647–656.
20. Ulmasov, T., Hagen, G. and Guilfoyle, T.J. (1997) ARF1, a transcription factor that binds to auxin response elements. *Science*, **276**, 1865–1868.
21. Nagano, Y. (2000) Several features of the GT-factor trihelix domain resemble those of the Myb DNA-binding domain. *Plant Physiol.*, **124**, 491–493.
22. Fry, C.J. and Farnham, P.J. (1999) Context-dependent transcriptional regulation. *J. Biol. Chem.*, **274**, 29583–29586.
23. Zhao, J. and Graf, G. (2000) The high mobility group I/Y protein is hypophosphorylated in endoreduplicating maize endosperm cells and is involved in alleviating histone H1-mediated transcriptional repression. *J. Biol. Chem.*, **275**, 27494–27499.