Nucleic Acids Research, 1995, Vol. 23, No. 17 3403–3410

A novel DNA-binding domain that may form a single zinc finger motif

Shuichi Yanagisawa

Department of Life Sciences (Chemistry), Graduate School of Arts and Sciences, The University of Tokyo, Meguro, Tokyo 153, Japan

Received June 12, 1995; Revised and Accepted August 6, 1995

EMBL accession nos X79934 and X79935

ABSTRACT

MNB1a is a DNA-binding protein from maize that interacts with the 35S promoter of cauliflower mosaic virus. This protein did not show significant homologies with any other DNA-binding protein and MNB1a seemed to be a member of a multigene family. In this study, isolation of cDNAs from the gene family to which MNB1a belongs revealed a unique conserved domain, referred to herein as the Dof domain, that contains a novel cysteine-rich motif for a single putative zinc finger. The amino acid sequence of the Dof domain and the arrangement of cysteine residues in this domain differ from those of known zinc finger motifs. However, the Dof domain was shown to be a DNA-binding domain that required Zn^{2+} ions for activity. Mutations at cysteine residues eliminated the DNA-binding activity of MNB1a. Thus, the Dof domain may be classified as a novel zinc finger motif. In addition, Southern blot analysis and a survey of DNA databases suggested that proteins that include Dof domains might exist in other eukaryotes, at least in the plant kingdom.

INTRODUCTION

Studies on the structure of DNA-binding transcription factors have revealed domains that are conserved in many such factors. These domains include the homeobox, zinc finger and leucine zipper, and many transcription factors appear to have been constructed from a variety of combinatorially arranged functional domains (1). Zinc finger motifs, which were originally identified in the transcription factor TFIIIA from *Xenopus*, can be classified into subgroups, according to the number and position of the cysteine and histidine residues available for coordination to zinc (2). Standard Zn finger proteins have multiple copies of the finger motif within a single polypeptide, organized as tandem repeats (2–8), and repeated finger domains have been shown to participate in interactions at a single target site by X-ray crystallographic analysis (9).

MNB1a is a DNA-binding protein of maize that interacts in a sequence-specific manner with AAGG motifs in a region (–276 to –252 relative to the transcription start site) of the cauliflower mosaic virus 35S promoter (10), which is one of the most powerful known promoters in the plant kingdom (11). DNA-binding analysis *in vitro* revealed that a leaf-specific nuclear factor, MNF1 (12,13), which binds to a promoter region of a maize gene for a protein that is involved in photosynthesis [phosphoenolpyruvate carboxylase (PEPC)], interacts with identical AAGG motifs in the 35S promoter (10). MNB1a seemed not to be identical to MNF1, but a family whose members show a similar sequence-specificity in DNA-binding is suggestible since the gene for MNB1a appeared to be a member of a multigene family (10). This factor, MNF1, is a candidate for a transcription factor that regulates the expression of several genes in leaf tissue, since the MNF1-binding site in the promoter of the PEPC gene seemed to be important for expression of the gene in differentiated leaves in a study of transgenic tobacco plants (14). Another study using transgenic tobacco plants suggested that a region (–301 to –208) of the 35S promoter, which contains an MNF1-binding site, might be involved in the expression in immature vascular tissue, although complicated combinational effects of multiple *cis*-elements within the 35S promoter was also reported (15). Thus, a regulatory network for the control of gene expression that is associated with differentiation might involve the products of this multigene family.

The MNB1a protein does not exhibit significant homology to other known proteins (10). However, MNB1a is presumably a transcription factor, since the specific interaction between MNB1a and its target sequence activated transcription slightly (unpublished data), although the detail analysis is now in progress. Thus, comparison among members of this family and the delineation of the functional domains in MNB1a will provide valuable information. In this paper, I describe that a unique DNA-binding domain that is conserved in this multigene family and that contains a novel cysteine-rich motif for a single putative zinc finger. The possibility that proteins containing this motif might exist in other eukaryotes is also discussed.

MATERIALS AND METHODS

Construction of a cDNA library and isolation of cDNAs

cDNA was synthesized using an oligo(dT) primer and poly(A)$^+$ RNA prepared from leaves of 10-day-old maize (an inbred strain of *Zea mays* H84) by the previously described method (10). The cDNA, ligated into the λgt10 vector, was packaged with an *in vitro* packaging kit (Amersham, UK). The resulting cDNA library contained 3.7×10^6 independent clones. The cDNAs encoding proteins with homology to MNB1a were isolated by the plaque hybridization technique from a total of 400 000 recombinant phages. Hybridization was allowed to proceed in a buffer that contained 5× Denhardt's solution, 6× SSC and 0.5% SDS at 63°C, and membranes were washed finally in 0.5× SSC plus 0.1% SDS at 63°C. A *Pst*I fragment (positions 67–610) of the MNB1a cDNA (10) was labeled with the DIG-DNA labeling system (Boehringer Mannheim Biochemica, Germany) and used as a probe DNA.

Table 1. Oligonucleotides used in this study

Name	Sequence[a]
LI1	5′-CATGGCCGGCATG-34
LI2	3′-CGGCC-54
SN1	5′-CCTGGTCTTGCTGGCCAACAT-34
SC1	5′-CCCGGCGGCAACCGAGCGTTCT-34
PC1	5′-CGTcCCGGGACACCAAGTTCTGCT-34
PC2(60S)	5′-CGTcCCGGGACACCAAGTTCTcCTACTACAA-34
PC3(49S)	5′-GGGGAtCCGTcCCCGCGGTGTGCGTCG-34
LWT1	5′-caTGGGCGAGGGGGACCCGTGCCCGCGGTGTGCGTC-34
LWT2	3′-CCGCTCCCCCTGGGCACGGGCGCCACACGCAGgGCC-5′
L52Ala1	5′-caTGGGCGAGGGGGACCCGTGCCCGCGGgcgGCGTC-34
L52Ala2	3′-CCGCTCCCCCTGGGCACGGGCGCCcgcCGCAGgGCC-54
L49Ser1	5′-caTGGGCGAGGGG-34
L49Ser	3′-CCGCTCCCCCTaG-54

[a]Small letters indicate the altered nucleotides for creation of restriction enzyme sites or introduction of mutations on cysteines residues.

Construction of expression vectors for deletion mutants of MNB1a

A *Sph*I–*Hin*cII fragment of the MNB1a cDNA was subcloned into of M13 RF to yield RF-MNB1a. The first ATG codon was at a *Sph*I site and a unique *Hin*cII site was located in the 3′ non-coding region. Since this construction created a new *Sal*I site at the ligation site of two *Hin*cII ends, a *Sal*I–*Eco*RI DNA fragment (302 bp) of pCaMVCN (Pharmacia) was inserted into RF-MNB1a to create RF-MNB1a-T. A plasmid, pAX4a$^+$-MNB1a, was constructed with the *Sph*I–*Eco*RI fragment from RF-MNB1a-T and the pAX4a$^+$ vector (Mo Bi Tec, Wagenstieg, Germany) that had been digested with *Nco*I and *Eco*RI, for expression of a fusion protein of β-galactosidase and MNB1a. Two oligonucleotides, LI1 and LI2 (see Table 1) were used to ligate the *Nco*I end and the *Sph*I end. Several plasmids were also constructed to produce fusion proteins that contained different regions of MNB1a. The pCdel series was constructed from pAX-MNB1a that had been digested with *Eco*RI and *Kpn*I, with further digestion by exonuclease III and mung bean nuclease. Another series of plasmid, pNdel, was constructed by direct cloning of the DNA fragments that had been produced from RF-MNB1a by sequential treatments with *Hin*dIII, exonuclease III, mung bean nuclease and *Eco*RI into pAX4a$^+$ that had been digested with *Nru*I and *Eco*RI. The nucleotide sequences around the ligation sites of each construct were confirmed by direct sequencing.

Construction of expression vectors for derivatives of MNB1a with site-specific point mutations

In order to construct a plasmid [pWT(44–147)] for expression of the region from the 44th amino acid residue to the 147th amino acid residue of MNB1a, PCR was performed using two oligonucleotides, SC1 and PC1 (Table 1), and pCdel 529, which was one of the plasmids in the deletion series (Fig. 3), as template. Oligonucleotide SC1 binds 17 base pairs downstream of the stop codons in the expression vector, and oligonucleotide PC1 binds to the nucleotide sequence of the cDNA that corresponds to the amino acid sequence around Cys^{60}. Since oligonucleotide PC1 carried a mutation in a restriction site (*Xma*I) site that did not alter the encoded amino acid sequence, the product of PCR was digested by *Xma*I and *Sal*I and inserted into pAX4a$^+$, which had been digested with *Nco*I and *Sal*I. In this ligation reaction, two oligonucleotides, LWT1 and LWT2 (Table 1), which were complementary to each other, were used. A plasmid (pCys52>Ala) for expression of a derivative of MNB1a with a point mutation was constructed in a similar way. The same PCR product digested with *Xma*I and *Sal*I was cloned into pAX4a$^+$ with two oligonucleotides L52Ala1 and L52Ala2 (Table 1). This set of oligonucleotides carried mutations to encode Ala in place of Cys^{52}. Two plasmids (pCys49>Ser and pCys60>Ser) for

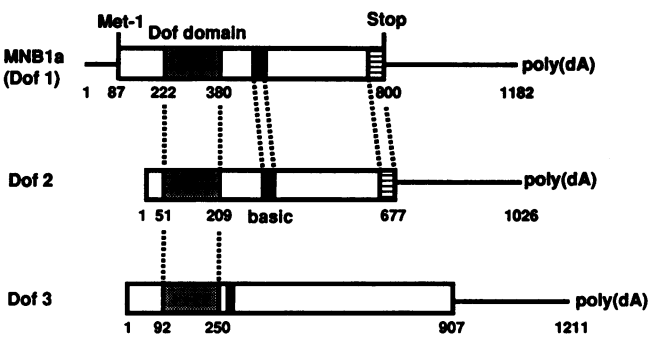

Figure 1. Schematic representation of three cDNA clones of the Dof family. The positions of Dof domains, regions rich in basic amino acid residues, and a homologous region in only MNB1a cDNA and Dof 2 cDNA are indicated by three kinds of box. The basic regions of MNB1a and Dof 2 are homologous to each other, but the basic region of Dof 3 is different.

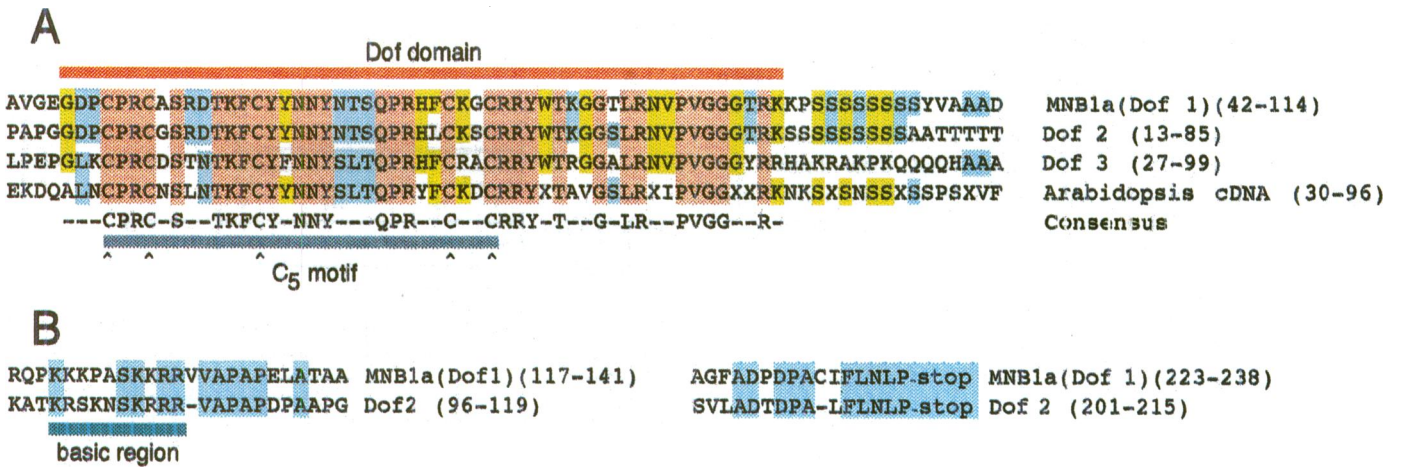

Figure 2. (**A**) Comparison of amino acid sequences in the regions that are highly conserved among three cDNA clones of the Dof gene family. Since, in a survey of databases, nucleotide sequence of a cDNA clone from *Arabidopsis thaliana* (accession number T14116) was found to exhibit significant homology, this sequence is indicated in the comparison. The amino acid sequence of MNB1a (Dof 1) was deduced from the nucleotide sequence, and the amino acid sequences of Dof 2, Dof 3 and cDNA from *Arabidopsis thaliana* are indicated in the same frame as that of Dof 1. Amino acid residues conserved among four clones and amino acid residues conserved among two or three clones are indicated. The positions of the amino acid residues are also indicated. The numbering is given relative to the first methionine residue (in case of MNB1a), the putative first methionine residue of the deduced amino acid sequence (in case of the cDNA from *Arabidopsis thaliana*) or the first amino acid residue encoded by the cDNA (in the case of Dof 2 and Dof 3). Amino acid residues that could not be deduced from the incomplete nucleotide sequence of the *Arabidopsis thaliana* cDNA are indicated by X. (**B**) Comparison of the conserved amino acid sequences encoded MNB1a and Dof 2. Identical amino acid residues and stop codons are indicated. The positions of the amino acids are also indicated.

expression of derivatives of MNB1a with point mutations were also constructed in a similar manner. PCR was performed with pCdel 529 as template using two oligonucleotides, SC1 and PC2. The oligonucleotide PC2 carried mutations for creation of a *Xma*I site and substitution of Cys^{60} by Ser (Table 1). PCR was also performed with pCdel 529 as template using two oligonucleotides SC1 and PC3 (Table 1). The oligonucleotide PC3 carried mutations for creation of a *Bam*HI site and replacement of Cys^{49} by Ser and the mutation for a *Bam*HI site did not alter the encoded amino acid sequence. After the two products of PCR had been digested with *Xma*I and *Sal*I, or with *Bam*HI and *Sal*I, they were cloned into pAX4a$^+$ using oligonucleotides LWT1 and LWT2, or L49Ser1 and L49Ser2, respectively. Both strands of the inserted cDNAs were sequenced by the direct plasmid sequencing technique using oligonucleotides SC1 and SN1 (SN1 binds to the sequence upstream of the multi-cloning site of pAX4a$^+$; Table 1) to confirm the absence of any but the intended mutations.

Southwestern blot analysis

After the plasmids (pAX4a$^+$ or derivatives thereof) had been used to transform cells of *E.coli* strain AD18 [CSH26] (Δ*pro-lac*, *lon*-100, *tsx*::Tn5(*Km*^{r}), F′[*lacI*^{q}, *lacZ* ΔM15, *y*$^+$, *pro*$^+$]), lysates were prepared from each transformant as described previously (10). Proteins in aliquots of lysates were separated by SDS–PAGE on a 7% polyacrylamide gel and transferred to a nitrocellulose filter (Gelman Science, USA). The filter was probed with ^{32}P-labeled DNA that contained MNB1a-binding sites, as described previously (10).

In an experiment to examine dependence of DNA-binding activity on Zn^{2+} ions, strips of a nitrocellulose filter to which proteins had been transferred were incubated for 3 h in buffer B [15 mM HEPES–KOH (pH 7.6), 40 mM KCl, 6% glycerol, 1 mM dithiothreitol] with or without 10 mM 1,10-phenanthroline. After washing in buffer B supplemented with 1 mM EDTA, each strip was incubated for 1 h in a solution (15 mM HEPES, 40 mM KCl, 6% glycerol, 1 mM dithiothreitol) with a final pH of 4.2, with or without 10 mM $ZnCl_2$, 10 mM $MgCl_2$ or 10 mM $CaCl_2$. After strips had been washed in buffer B that contained 1 mM EDTA, they were subjected to Southwestern blot analysis.

Western blot analysis

Proteins in aliquots of lysates from transformants were separated by SDS–PAGE on a 7% polyacrylamide gel and transferred to a nitrocellulose filter (Gelman Science). The filter was blocked in PBS supplemented with 1% skimmed milk and incubated for 1 h with polyclonal antiserum against β-galactosidase (Chemicon International, USA) at a 1:4000 dilution and then for 1 h with horseradish peroxidase-conjugated goat anti-rabbit IgG (1:4000; BioRad, USA). The filter was washed in PBS plus 0.05% Tween-20 between each step and color was developed using 3,3′-diaminobenzidine. The filters used in Southwestern blot analysis were also subjected to Western blot analysis.

Southern blot analysis

Genomic DNA from *Drosophila melanogaster* (Canton S) was purchased from Clontech (USA). Gifts of genomic DNA from *Arabidopsis thaliana* (Columbia ecotype), *Saccharomyces cerevisiae* (YPH 500) and mouse (C3H) were obtained as indicated in Acknowledgements. Southern blot analysis was performed as described previously (16). Genomic DNA, digested with *Eco*RI or *Hin*dIII, was subjected to electrophoresis in a 0.7% agarose gel and blotted onto a nylon membrane (Hybond-N, Amersham). The membrane was probed with appropriate DNA fragments that had been labeled with ^{32}P by random priming. Hybridization was performed at 63°C and the membranes were finally washed in 1 × SSC plus 0.1% SDS at 63°C.

DNA sequencing and computer analysis

Nucleotide sequences were determined by the dideoxy chain termination method (17) using a TTH sequence kit (TOYOBO, Japan). The FASTA program (18) was used to search for homologies in the SWISS-PROT and EMBL data library. The nucleotide sequence data reported in this paper have been deposited in the EMBL/GenBank/DDBJ databases with accession numbers X79934 (Dof 2) and X79935 (Dof 3). Since errors were found in the originally reported nucleotide sequence of the MNB1a cDNA, due to the presence of very GC-rich regions, the nucleotide sequence previously deposited in the databases was corrected appropriately (accession number X66076).

RESULTS

Isolation of cDNAs that encode proteins with homology to MNB1a

To isolate cDNAs encoding proteins with homology to MNB1a, a cDNA library was constructed in the λgt10 vector with poly(A)$^+$ RNA from maize leaves. The cDNA library was screened using a *Pst*I fragment (positions 67–610) of the MNB1a cDNA, since the members of the multigene family to which MNB1a belongs seemed to have N-terminal regions that are highly homologous (10). Eight cDNA clones were finally isolated as positive clones that significantly hybridized to the cDNA of MNB1a but only two out of the eight clones had restriction maps different from that of MNB1a cDNA (data not shown). Thus, only two cDNA clones were chosen for further analysis. The products of the newly isolated cDNAs were designated Dof 2 and Dof 3, respectively.

Identification of a novel cysteine-rich motif in DNA-binding proteins

The nucleotide sequences of two cDNAs were determined by the dideoxy sequencing technique and were compared with that of MNB1a cDNA. The comparison of these nucleotide sequences allowed clear delineation of the homologous and non-homologous regions in the three clones, as shown schematically in Figure 1. The amino acid sequences corresponding to the nucleotide sequences that were conserved in the three cDNA clones are shown in Figure 2A. The amino acid sequences, referred to as Dof domains, include conserved cysteine residues and a histidine residue that might form only one Zn finger motif. Alignment of the nucleotide sequences of these clones also revealed that these clones did not exhibit significant homology in other regions, with the exception of two short regions that were conserved between only MNB1a and Dof 2 (Fig. 2B). Therefore, the MNB1a cDNA seemed to be divisible into four domains on the basis of homologies among members of the gene family (Fig. 3A).

The Dof domain is distinct in some respects from known Zn finger domains, which have been classified as C_2-C_2 Zn fingers, C_2-H_2 Zn fingers, C_6 Zn fingers and LIM domains (2–8). Standard Zn finger proteins have multiple copies of the finger motif within a single polypeptide, organized as tandem repeats, while the Dof domain contained only a single putative Zn finger motif. The Dof domain consisted of $CX_2CX_7CX_{11}HXCX_2C$ (where X is any amino acid), and was somewhat similar but not identical to the $CX_2CX_{17-19}HX_2CX_2CX_2CX_{7-11}(C)X_8C$ sequence of the LIM domain (5,6). Only a single cDNA clone, from *Arabidopsis*, (accession number T14116) that encodes an amino

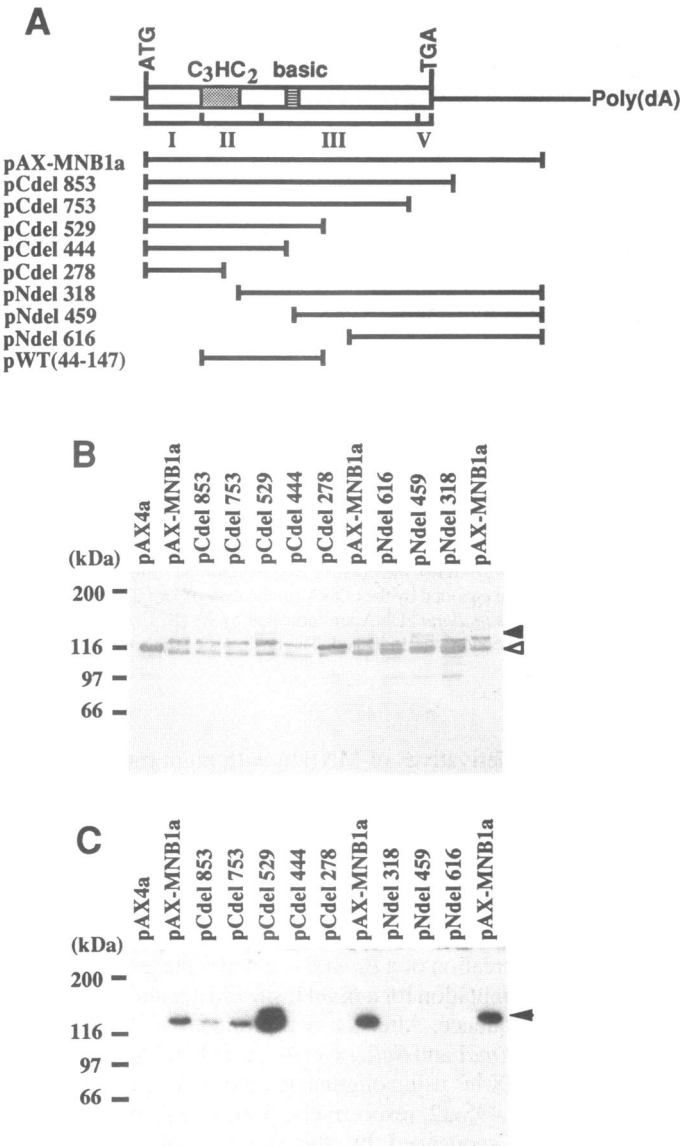

Figure 3. (**A**) The various cDNA fragments that were inserted into the expression vector pAX4a$^+$. Four regions that were recognized from homologies to Dof 2 and Dof 3 are shown. Region II is identical to the Dof domain shown in Figure 2. The C_3HC_2 motif and the basic region in Figure 2 are also indicated. The DNA fragments that were inserted into pAX4a$^+$ for expression of fusion proteins are indicated. The numbers of members of the pCdel series correspond to the nucleotide number at the 3′ end of each inserted cDNA and numbers of members of the pNdel series correspond to that at the 5′ end. The position of a DNA fragment used for construction of plasmid pWT(44–147), as described in Materials and Methods, is also indicated. (**B**) Western blot analysis. After the transformation of *E.coli* cells by plasmids (pAX4a$^+$ or derivatives), proteins (20 μg each) in an aliquot of the lysate from each transformant were resolved by SDS–PAGE and transferred to a nitrocellulose filter. Endogenous β-galactosidase and fusion proteins were visualized by reaction with a polyclonal antiserum. The positions of β-galactosidase and fusion proteins are indicated by open and solid arrowheads, respectively. (**C**) Southwestern blot analysis. Proteins (40 μg each) from each transformant were resolved by SDS–PAGE and transferred to a nitrocellulose filter. The filter was probed with ^{32}P-labeled DNA that contained the MNB1a-binding sites. The positions of fusion proteins are indicated by a solid arrowhead.

acid sequence homologous to the Dof domain was found in a survey of the databases. However, the product of this cDNA has not been characterized. The amino acid sequence deduced from the *Arabidopsis* cDNA clone was indicated in the comparison of Dof domains (Fig. 2A). Since His^{72} was not found in the sequence of the Dof domain in the *Arabidopsis* cDNA clone, Dof domain did not seem to form a similar structure to Zn_2Cys_6 binuclear cluster (19). Thus, the amino acid sequence of the Dof domain appears to differ from those of known Zn finger motifs.

The Dof domain is involved in the interaction with DNA

Zinc fingers known to be involved in site-specific recognition (20) and repeated finger domains have been shown to participate in interactions at a single target site by X-ray crystallographic analysis (9). An attempt was made to identify the DNA-binding domain of MNB1a (Dof 1). Several plasmids that expressed different regions of the MNB1a cDNA as fusion proteins with β-galactosidase were constructed (Fig. 3A). These plasmids were used to transform cells of *E.coli*. Then, Western blot analysis, using polyclonal antibodies against β-galactosidase and the lysate from each transformant, was performed in order to confirm that almost equal amounts of the various fusion proteins had been expressed from $pAX4a^+$ and each derivative, irrespective of the nature of the construct (Fig. 3B). Since an almost equal amount of fusion protein was detected in each case, with the exception of a relative low amount of fusion protein in the lysate from the transformant that harbored pNdel 459, the results of Southwestern blot analysis, shown in Figure 3C, can be taken to reflect the different DNA-binding activities of the various fusion proteins. The products of the constructs in which the Dof domain had been deleted or disrupted (pNdel 318, pNdel 459, pNdel 616 and pCdel 278) did not show any DNA-binding activity (Fig. 3C). Exposure for a long period revealed low but significant DNA-binding activity of the product of pCdel 444, which was expected to include a Dof domain but to lack a basic region outside the Dof domain (data not shown). By contrast, the product of pCdel 528, which contained the Dof domain and the basic region, had a high DNA-binding activity. These observations suggest that the Dof domain is essential for DNA-binding activity and that a region rich in basic amino acids outside the Dof domain, in which some homology between MNB1a (Dof 1) and Dof 2 was found, reinforces the binding activity.

To confirm that region 1 (the 1st to the 45th amino acid residues) is not involved in the DNA-binding activity of MNB1a, another plasmid, pWT(44–147) was constructed (see Fig. 3A and Materials and Methods). This plasmid was expected to produce a fusion protein that contained a region of MNB1a from the 44th amino acid residue to the 147th amino residues. The DNA-binding activity of this deletion mutant protein of MNB1a was detected [Fig. 4B, lane pWT(44–147)], and the activity of this protein was similar to that of the protein encoded by pCdel 528. This result suggests that two regions, the Dof domain and the basic region, are sufficient for the binding to DNA.

Identification of cysteine residues in the Dof domain that are essential for DNA-binding activity

If the Dof domain forms a Zn finger motif, replacement of cysteine residues that coordinate to zinc by different amino acid residues should eliminate the DNA-binding activity of MNB1a.

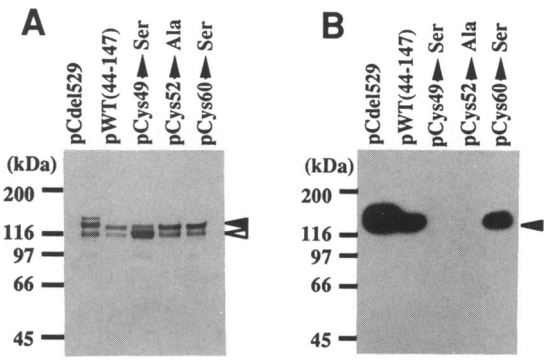

Figure 4. (**A**) Western blot analysis. Proteins (20 μg each) in aliquots of lysates from the transformants obtained with pCdel529, pWT(44–147), pCys49>Ser, pCys52>Ala and pCys60>Ser were resolved by SDS–PAGE and transferred to a nitrocellulose filter. Endogenous β-galactosidase and fusion proteins were visualized by reaction with a polyclonal antiserum. The positions of β-galactosidase and fusion proteins are indicated by open and solid arrowheads, respectively. (**B**) Southwestern blot analysis. Proteins (40 μg each) from each transformant were resolved by SDS–PAGE and transferred to a nitrocellulose filter. The filter was probed with ^{32}P-labeled DNA that contained the MNB1a-binding sites. The positions of fusion proteins are indicated by a solid arrowhead.

The Dof domain includes five cysteine residues and one histidine residue, although the histidine residue was not conserved in a cDNA from *Arabidopsis thaliana* (Fig. 2A). Therefore, three plasmids, based on pWT(44–147), were constructed for expression of proteins with point mutation (see Materials and Methods). Although Western blot analysis demonstrated that the wild-type protein and the mutant proteins were produced at almost equal levels in *E.coli* (Fig. 4A), the DNA-binding activity of each protein was different. The replacement of Cys^{60} by Ser did not affect the DNA-binding activity (Fig. 4B, lane pCys60>Ser). By contrast, the proteins with replacement Cys^{49} by Ser or of Cys^{52} by Ala, did not have any DNA-binding activity. These results suggest that the Dof domain includes a zinc finger motif and that Cys^{49} and Cys^{52} contribute to the zinc finger motif while Cys^{60} did not.

The DNA-binding activity of MNB1a requires Zn^{2+} ions

To determine whether MNB1a is a metalloprotein and its activity requires Zn^{2+} ions, increasing amounts of metal-chelating compounds were added to the binding buffer for assays of activity. Both EDTA and 1,10-phenanthroline inhibited the DNA-binding activity of MNB1a but 1,10-phenanthroline was more effective than EDTA (Fig. 5A). Since Zn finger proteins, such as transcription factor TF IIIA of *Xenopus*, are more sensitive to 1,10-phenanthroline than to other chelators (21,22), this observation supports the hypothesis that the novel cysteine-rich motif in the Dof domain forms a zinc finger. To confirm that chelators did not inhibit the binding via a non-specific mechanism, an experiment was performed to determine whether or not the DNA-binding activity could be restored by addition of Zn^{2+} ions after inhibition by a chelator. After strips of a filter to which proteins had been transferred were immersed in a solution that contained 1,10-phenanthroline and subsequently incubated in a solution of a metal salt, the DNA-binding activity associated with each strip was examined. Significant inhibition by 1,10-phenan-

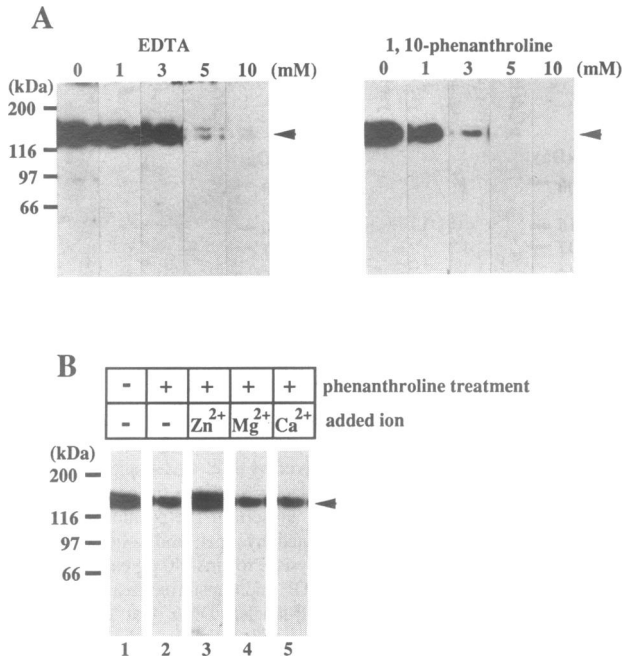

Figure 5. (**A**) Effects of chelators on the DNA-binding activity of MNB1a. Proteins from transformants that carried pCdel 529 were separated by SDS–PAGE and transferred to a nitrocellulose filter. Strips of the filter were incubated in blocking buffer and then in binding buffer with ^{32}P-labeled DNA. Each buffer was supplemented with EDTA or 1,10-phenanthroline at the indicated concentration. (**B**) Zn^{2+}-specific restoration of the DNA-binding activity. Proteins from transformants that carried pWT(44–147) were resolved by SDS–PAGE and transferred onto a nitrocellulose filter. Strips of this nitrocellulose filter were incubated in a buffer without (lane 1) or with 10 mM 1,10-phenanthroline (lanes 2–5) and then incubated in a buffer solution without (lanes 1 and 2) or supplemented with 10 mM $ZnCl_2$ (lane 3), 10 mM $MgCl_2$ (lane 4) or 10 mM $CaCl_2$ (lane 5). The DNA-binding activity of proteins on each strip was analyzed with ^{32}P-labeled DNA.

throline was observed, although the inhibition was weaker than that when 1,10-phenanthroline was added to the binding buffer (Fig. 5B, lanes 1 and 2). The detection of DNA-binding activity in lane 2 may be due to the metal contamination of the equipment. Such technical difficulties have been noted in a previous report (23). The reduction in DNA-binding activity caused by 1,10-phenanthroline was reversed by incubation in a solution that contained Zn^{2+} ions but not in a solution that contained Mg^{2+} or Ca^{2+} ions (Fig. 5B, lanes 3–5). This result suggests that MNB1a is a metalloprotein and that its DNA-binding activity requires Zn^{2+} ions.

Detection of nucleotide sequences homologous to that of the MNB1a cDNA in other eukaryotes

MNB1a, Dof 2 and Dof 3 cDNAs were isolated from maize (monocots). To investigate whether homologs of MNB1a and proteins containing Dof domains exist in other eukaryotes, genomic Southern blot analysis was performed using genomic DNAs from *Arabidopsis thaliana* (dicot), a model plant that is frequently used in recent genetic studies (24), and *Drosophila melanogaster*. When a probe (probe 1 shown in Fig. 6A) covering almost the entire MNB1a cDNA was used, multiple bands were observed in lanes that had been loaded with *Eco*RI-digested or *Hin*dIII digested genomic DNA of *Arabidopsis thaliana* and in lanes loaded with similarly digested genomic DNA from *Drosophila melanogaster* (Fig. 6B and C, panel probe 1). To examine whether a nucleotide sequence corresponding to the Dof domain was responsible for these observed bands, further analysis was carried out with two other DNA probes, which had sequences homologous only in the region that corresponded to the Dof domain. Probe 2 was derived from MNB1a cDNA and probe 3 was derived from the Dof 3 cDNA (Fig. 6A). Both probes yield multiple bands. Some bands indicated by arrowheads in Figure 6B and C were observed with both probe 2 or 3. These bands may reflect DNA fragments that encode Dof domains, and the different intensities of the bands obtained with probes 2 and 3 might depend on the extent of homology between the DNA fragments and the probes.

DISCUSSION

This report describes identification of a novel DNA-binding domain in a maize DNA-binding protein, MNB1a. This domain contains a cysteine-rich motif that presumably forms a single zinc finger, as judged from the following observations: (i) chelators in the binding reaction inhibited the DNA-binding activity; (ii) the DNA-binding activity that was reduced by the presence of a chelator was restored by the addition of Zn^{2+} ions and (iii) mutant derivatives of MNB1a with conversion of Cys^{49} or Cys^{52} by Ser had no DNA-binding activity.

The so-called Dof domain is distinguishable from known zinc finger motifs, since its amino acid sequence and the arrangement of the cysteine residues that probably contribute to formation of a zinc finger are different from those of known zinc finger domains. A proposal for the structure of the Dof domain is shown in Figure 7. The importance of Cys^{74} and Cys^{77} was not confirmed experimentally, but these residues may well contribute to formation of a zinc finger, since His^{72} was not conserved in the Dof domain of the cDNA clone from *Arabidopsis* (Fig. 2). The cysteine residues that are available for coordination to zinc are arranged as $CX_2CX_{21}CX_2C$, and the finger in the Dof domain is larger than other known fingers.

Standard zinc finger proteins contains multiple copies of the finger motif and the repeated motifs participate in interactions at a target site. In the case of members of the GATA multigene family in vertebrates that are assumed to play important roles in development, each member contains two zinc fingers. It has been suggested that the C-terminal finger of GATA-1 plays a dominant role in DNA-binding and that the N-terminal finger plays a minor role for the stability of the DNA–protein complex (25,26). MNB1a contains only one zinc finger motif in the Dof domain. The DNA-binding activity of deletion mutant proteins that contained the Dof domain but not a basic region was very much reduced. The basic region might play a role in reinforcing the DNA–protein complex, as does the N-terminal finger of GATA-1. Incidentally, Dof 2 had a basic region homologous to that of MNB1a, and Dof 3 had a region rich in basic amino acids whose position was shown in Figure 1. However, the sequence of the basic region of Dof 3 was different from that of MNB1a.

Cys-rich motifs have been found not only in DNA-binding proteins but also in other proteins, such as Ser/Thr kinases (27). However, Dof 2 and Dof 3 may also be DNA-binding proteins that recognize sequences that are similar but not identical to that recognized by MNB1a, since a maize nuclear factor (MNF1)

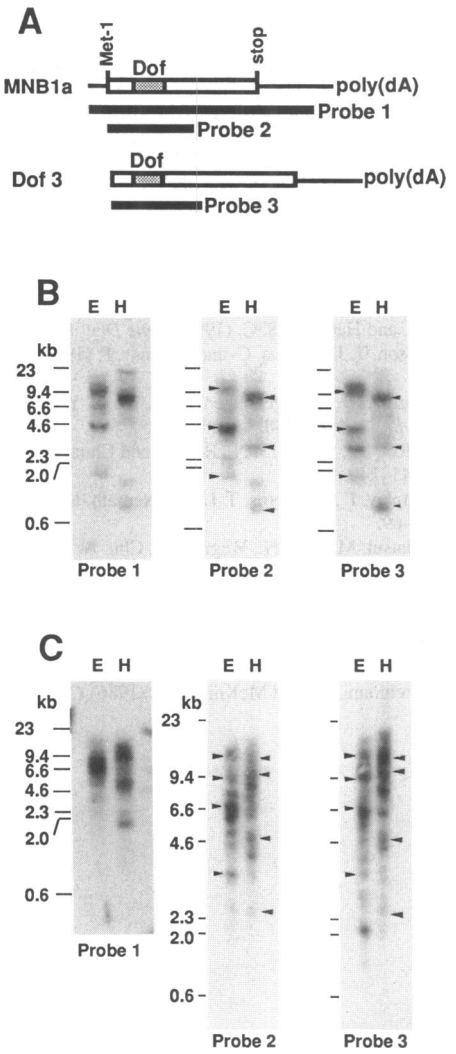

Figure 6. Detection of DNA sequences homologous to that of the MNB1a-coding region in the genomes of a dicot and an insect. Genomic DNAs from *Arabidopsis thaliana* and *Drosophila melanogaster* were digested with *Eco*RI (lane E) and *Hin*dIII (lane H). (**A**) The positions of the probes used for Southern blot analysis. Probe 1 (1–1090) and probe 2 (87–444) were derived from MNB1a cDNA and probe 3 (1–420) was derived from Dof 3 cDNA. (**B**) Detection of sequences homologous to the MNB1a-coding region in the genome of *Arabidopsis thaliana*. (**C**) Detection of sequences homologous to the MNB1a-coding region in the genome of *Drosophila melanogaster*. Arrowheads indicate the positions of the bands that were observed with both probe 2 and probe 3.

bound to the MNB1a-binding sites with a similar but not identical sequence-specificity, suggesting a family of DNA-binding proteins with a similar sequence-specificity (10). The DNA-binding activity of Dof 2 was already detected and its sequence-specificity was similar to that of MNB1a (unpublished data). However, detection of the DNA-binding activity of Dof 3 was not achieved, since I failed in producing Dof 3 in *E.coli*, suggesting that expression of Dof 3 is toxic for *E.coli* (unpublished data). Thus, it is a next question whether all Dof domains recognize similar sequences or not. Analyses on the DNA-binding activity of Dof 3 are now in progress by another assay system.

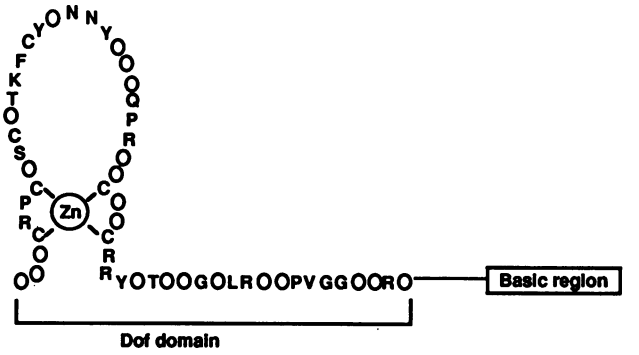

Figure 7. Proposed structure of a novel DNA-binding region containing the Dof domain. Amino acid residues conserved among the deduced products of four cDNA clones that contained the Dof domain are indicated by single letters and non-conserved amino acid residues are indicated by open circles.

The Dof domain was found as an 'island' among non-homologous sequences in the three cDNAs. Since the Dof domain is a DNA-binding domain, it is possible that Dof proteins have gained functional diversity as transcription factors, via recombination with other functional domains. The exon-shuffling model (28) might be applicable to the molecular evolution of members of the Dof gene family. Analysis of the intron–exon structure of this family might provide support for this hypothesis. It is worth mentioning that MNB1a is presumably a transcription factor, since specific interaction between MNB1a and its target site activated the transcription of the reporter gene ~2-fold in experiments using maize protoplasts (unpublished data), although detail analyses are necessary to establish the role of MNB1a on transcription and they are now in progress.

In spite of the novel nature of this DNA-binding domain, the Dof domain might be conserved among eukaryotes. The detection of homologous sequences in the genomes of *Arabidopsis thaliana* and *Drosophila melanogaster* suggests the existence of the Dof sequences in these organisms. In fact, a cDNA of *Arabidopsis thaliana* with the sequence for a Dof domain was found in a DNA database. However, the complete nucleotide sequence of this clone has not been determined and the function of the encoded protein is unknown. With two DNA probes (probes 1 and 2 in Fig. 6) derived from MNB1a cDNA, several bands were identified in an analysis of mouse genomic DNA and one band in that of yeast (*Saccharomyces cerevisiae*) genomic DNA (data not shown). However, these bands were not reproduced with probe 3 (Fig. 6) that originated from Dof 3 cDNA. It is unclear whether non-detection with probe 3 was due to the relative low homology between the sequences in the genomes and probe 3, or whether detection with probes 1 and 2 was due to hybridization with sequences outside the Dof domain or other kinds of cross-reaction. However, it is attractive to speculate that the Dof domain might be conserved among eukaryotes. Further analysis is necessary to determine whether Dof proteins are widely conserved in not only the plant kingdom and but also animal kingdom, since no clone encoding the Dof domain is obtained from the animal kingdom.

ACKNOWLEDGEMENTS

The author is indebted to Drs Hiroshi Akanuma (University of Tokyo) for continuous support and to Kiyotaka Okada (National Institute for Basic Biology) for the gift of genomic DNA from *Arabidopsis thaliana*, Yoh Wada (University of Tokyo) for yeast genomic DNA and Kazuyuki Hoshijima (Kyoto University) for genomic DNA from the mouse. This work was supported by a Grant-in-Aid for Scientific Research on Priority Areas ('The Molecular Basis of Flexible Organ Plans in Plants', no. 06278204) from the Ministry of Education, Science and Culture of Japan.

REFERENCES

1 Mitchell, P. J. and Tjian, R. (1989) *Science*, **245**, 371–378.
2 Evans, R. M. and Hollenberg, S. M. (1988) *Cell*, **52**, 1–3.
3 Evans, R.M. (1988) *Science*, **240**, 889–895.
4 Fan, C.-M. and Maniatis, T. (1990) *Genes Dev.*, **4**, 29–42.
5 Freyd, G., Kim, S. K. and Horvitz, H. R. (1990) *Nature*, **344**, 876–879.
6 Karlsson, O., Thor, S. and Norberg, T. (1990) *Nature*, **344**, 879–882.
7 Strich, R., Surosky, R. T., Steber, C., Dubois, E., Messenguy, F. and Esposito, R. E. (1994) *Genes Dev.*, **8**, 796–810.
8 Turner, C. A., Mack, D. H. and Davis, M. M. (1994) *Cell*, **77**, 297–306.
9 Pavletich, N. P. and Pabo, C. O. (1991) *Science*, **252**, 809–816.
10 Yanagisawa, S. and Izui, K. (1993) *J. Biol. Chem.*, **268**, 16028–16036.
11 Benfey, P. N. and Chua, N.-H. (1990) *Science*, **250**, 959–966.
12 Yanagisawa, S. and Izui, K. (1990) *Mol. Gen. Genet.*, **224**, 325–332.
13 Yanagisawa, S. and Izui, K. (1992) *Plant Mol. Biol.*, **19**, 545–553.
14 Brunner, B. (1992) *Expression Analysis of a C4-type Maize Phosphoenolpyruvate Carboxylase Promoter Sequence in Tobacco.* Ph.D. dissertation, Swiss Federal Institute of Technology, Zurich
15 Benfey, P. N., Ren, L. and Chua, N.-H. (1990) *EMBO J.*, **9**, 1685–1696.
16 Yanagisawa, S. and Izui, K. (1989) *J. Biochem.*, **106**, 982–987.
17 Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA*,**74**, 5463–5467.
18 Pearson, W. R. and Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 2414–2448.
19 Marmorstein, R. and Harrison, S. C. (1994) *Gene Dev.*, **8**, 2504–2512.
20 Nardelli, J., Gibson, T. J., Vesque, C. and Charnay, P. (1991) *Nature* **349**, 175–178.
21 Hanas, J. S., Hazuda, D. J., Bogenhagen, D. F., Wu, F. Y.-H. and Wu, C.-W. (1983) *J. Biol. Chem.*, **258**, 14120–14125.
22 Takatsuji, H., Mori, M., Benfey, P. N., Ren, L. and Chua, N.-H. (1992) *EMBO J.*, **11**, 241–249.
23 Vallee, B. T., Rupley, J. A., Cooms, T. L. and Neurath, H. (1960) *J. Biol. Chem.*, **235**, 64–69.
24 Deng, X.-W., Matsui, M., Wei, N., Wagner, D., Chu, A. M., Feldmann, K. A. and Quail, P. H. (1992) *Cell*, **71**, 791–801.
25 Martin, D. I. K. and Orkin, S. H. (1990) *Genes Dev.*, **4**, 1886–1898.
26 Yang, H.-Y. and Evans, T. (1992) *Mol. Cell. Biol.*, **12**, 4562–4570.
27 Quest, A. F. G., Bardes, E. S. G. and Bell, R. M. (1994) *J. Biol. Chem.*, **269**, 2961–2970.
28 Gilbert, W., Marchionni, M. and McKnight, G. (1986) *Cell*, **46**, 151–154.