The Plant Cell, Vol. 7, 173–182, February 1995

Ethylene-Inducible DNA Binding Proteins That Interact with an Ethylene-Responsive Element

Masaru Ohme-Takagi [1] and Hideaki Shinshi

Plant Molecular Biology Laboratory, National Institute of Bioscience and Human Technology, 1-1 Higashi, Tsukuba, Ibaraki 305, Japan

We demonstrated that the GCC box, which is an 11-bp sequence (TAAGAGCCGCC) conserved in the 5′ upstream region of ethylene-inducible pathogenesis-related protein genes in *Nicotiana* spp and in some other plants, is the sequence that is essential for ethylene responsiveness when incorporated into a heterologous promoter. Competitive gel retardation assays showed DNA binding activities to be specific to the GCC box sequence in tobacco nuclear extracts. Four different cDNAs encoding DNA binding proteins specific for the GCC box sequence were isolated, and their products were designated ethylene-responsive element binding proteins (EREBPs). The deduced amino acid sequences of EREBPs exhibited no homology with those of known DNA binding proteins or transcription factors; neither did the deduced proteins contain a basic leucine zipper or zinc finger motif. The DNA binding domain was identified within a region of 59 amino acid residues that was common to all four deduced EREBPs. Regions highly homologous to the DNA binding domain of EREBPs were found in proteins deduced from the cDNAs of various plants, suggesting that this domain is evolutionarily conserved in plants. RNA gel blot analysis revealed that accumulation of mRNAs for EREBPs was induced by ethylene, but individual EREBPs exhibited different patterns of expression.

INTRODUCTION

Ethylene is an endogenous plant hormone that influences many aspects of plant growth and development, such as germination, senescence, epinasty, abscission, and fruit ripening (Abeles et al., 1992). It is one of the most intensely studied plant hormones. The biosynthetic pathway of ethylene has been established (Yang and Hoffman, 1984), and the genes encoding several of the biosynthetic enzymes have been cloned (Sato and Theologis, 1989; Nakajima et al., 1990; Van Der Straeten et al., 1990; Hamilton et al., 1991; Spanu et al., 1991). Recently, signal transducers for ethylene, CTR1 (constitutive triple response) and ETR1 (ethylene insensitive), were identified by molecular genetic studies of Arabidopsis (Chang et al., 1993; Kieber et al., 1993). Ethylene-regulatory *cis* regions have been identified in several ethylene-inducible genes (Broglie et al., 1989; Deikman et al., 1992; Meller et al., 1993; Vögeli-Lange et al., 1994; H. Shinshi, S. Usami, and M. Ohme-Takagi, submitted manuscript), and several proteins that specifically interact with these *cis* regulatory regions have been identified (Deikman and Fischer, 1988; Cordes et al., 1989; Holdsworth and Laties, 1989; Meller et al., 1993; H. Shinshi, S. Usami, and M. Ohme-Takagi, submitted manuscript). However, the nucleotide sequences involved in ethylene-dependent transcription have not been determined, and no corresponding DNA binding proteins have been purified or cloned.

The rate of biosynthesis of ethylene increases rapidly during plant–pathogen interactions (Yang and Hoffman, 1984), and ethylene subsequently induces transcription of a series of pathogenesis-related (PR) protein genes, such as genes encoding class I β-1,3-glucanase and class I chitinase (Felix and Meins, 1987; Vögeli et al., 1988). Comparisons of the 5′ upstream regions of ethylene-inducible PR protein genes in *Nicotiana* spp identified an 11-bp sequence (TAAGAGCCGCC) (Ohme and Shinshi, 1990; Eyal et al., 1993; Hart et al., 1993), which we refer to as the GCC box. This sequence has been predicted to be a target for the ethylene signal transduction pathway because deletion of the GCC box appears to eliminate the ethylene responsiveness of promoters (Broglie et al., 1989; Meller et al., 1993; Vögeli-Lange et al., 1994; H. Shinshi, S. Usami, and M. Ohme-Takagi, submitted manuscript). The specific interactions of nuclear factors with the GCC box have been identified in a 61-bp enhancer element of the tobacco β-1,3-glucanase B gene (Hart et al., 1993) and in the 146-bp ethylene-responsive region of the tobacco chitinase *Chn48* gene (H. Shinshi, S. Usami, and M. Ohme-Takagi, submitted manuscript).

In this report, we show that the GCC box is the ethylene-responsive element (ERE) that enhances ethylene-dependent transcription from a truncated (−46) 35S promoter of cauliflower mosaic virus (CaMV). We cloned and characterized cDNAs corresponding to ERE binding proteins (EREBPs) of tobacco that specifically interact with the GCC box in the ERE. We show here that EREBPs are novel DNA binding proteins that exhibit no sequence homology with known transcription factors or DNA binding proteins. The DNA binding domain of EREBPs was

[1] To whom correspondence should be addressed.

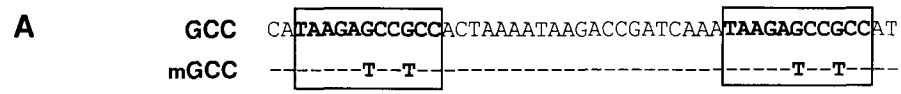
identified within a region of 59 amino acid residues that is conserved among all four EREBPs.

RESULTS

Identification of a *cis*-Acting Element Involved in the Ethylene-Responsive Transcription

Accumulated data suggested that the GCC box plays a role in the regulation of ethylene-inducible transcription of PR protein genes. To investigate whether the GCC box is the *cis* element for ethylene responsiveness, we prepared a 47-bp fragment (designated the GCC fragment) that contained two copies of the sequence of the GCC box in the 5′ upstream region of the tobacco β-1,3-glucanase *Gln2* gene (Ohme and Shinshi, 1990). We also prepared a mutated sequence (mGCC) with two single-base mutations in the sequence of each GCC box (Figure 1A). Chimeric promoter–β-glucuronidase (*Gus*) reporter genes were constructed in which two tandem copies of the GCC or the mGCC fragment were fused upstream of the truncated (−46) CaMV 35S promoter and the *Gus* reporter gene; these were designated 2(GCC)Gus and 2(mGCC)Gus, respectively (Figure 1B). Leaf discs from transgenic tobacco plants harboring the chimeric genes were treated with 1 mM ethephon (2-chloroethylphosphonic acid), an ethylene-releasing chemical, in 50 mM phosphate buffer, pH 7.0, or with the buffer alone and then assayed for GUS activities. The transgenic plants containing 2(GCC)Gus showed an average 115-fold induction of GUS-specific activity with ethephon treatment, whereas plants containing 2(mGGC)Gus exhibited very low GUS activity, and the activity was not induced with ethephon treatment (Figure 1B). These results indicate that the 47-bp fragment containing two copies of the GCC box sequence functioned as an ERE in a gain-of-function assay and that the GCC box sequence is directly involved in ethylene responsiveness of the element, because point mutations in the GCC box eliminated the ability of the element to activate gene transcription in an ethylene-dependent manner.

It has been reported that the acids released during the breakdown of ethephon (Yang, 1969) resulted in the induction of PR gene expression in Arabidopsis (Lawton et al., 1994). We incubated leaf discs in 50 mM phosphate buffer, pH 7.0, to avoid the effect of the acids and confirmed that the GUS activity of the transformants was not induced when leaf discs were treated



B

CONSTRUCT	GUS activity -Ethephon	GUS activity +Ethephon	Fold induction
a GCC GCC TATA GUS	29.0±6.7	1690±357	115±26
b mGCC mGCC TATA GUS	3.3±0.5	3.9±0.4	1.2±0.2

Figure 1. Ethylene-Responsive Expression Is Conferred on a Minimal CaMV Promoter by the GCC Box Sequence.

(A) DNA sequences of the GCC fragment and the mGCC fragment used in the experiments are shown. The GCC fragment was derived from the 5′ upstream region (−1164 to −1118) of the *Gln2* gene encoding β-1,3-glucanase (Ohme and Shinshi, 1990) with some modifications, as described in Methods. The mutant fragment (mGCC) contained two single-base substitutions (G to T) in the GCC box of the GCC fragment. The sequence of the GCC box is shown in boldface and boxed letters, and only substituted nucleotides are shown for mGCC. Dashes in mGCC indicate nucleotides identical to those in the GCC fragment.

(B) Schematic representations of the promoter region of the chimeric gene constructs and the results of the functional assays are shown. Chimeric gene constructs 2(GCC)Gus (a) and 2(mGCC)Gus (b), in which two tandem copies of the GCC fragment (GCC) or the mGCC fragment (mGCC) were fused upstream of the CaMV (−46) 35S promoter (TATA) and *Gus* with the polyadenylation sequence of the nopaline synthase gene, were constructed as described in Methods. Twenty-one and 28 independent transformants harboring the chimeric gene construct, 2(GCC)Gus and 2(mGCC)Gus, respectively, were analyzed. The average GUS specific activities (picomoles of 4-methylumbelliferone per minute per milligram of protein ± SD) in leaf discs with or without ethephon treatment of independent transformants are shown. Fold induction indicates the average of the ratios of GUS activity of independent transformants with or without ethephon (+ethephon/−ethephon).

with 2 mM hydrochloric acid in the phosphate buffer. Moreover, we observed that silver thiosulfate (50 μM), the ethylene action inhibitor, inhibited ethephon-inducible GUS activity to the basal level (data not shown). These results indicate that the induction by ethephon is due to ethylene release and not the acid's effect.

Identification of Nuclear Factors That Interact with an ERE

The results of the functional analyses suggested that the nuclear proteins that interact with the GCC box should be the transcription factors that regulate ethylene-dependent transcription. To correlate functions in vivo with binding activities in vitro, we investigated the nuclear proteins that interacted with the ERE (GCC fragment) but not with the mutated fragment (mGCC) by gel retardation assays. We identified DNA binding activities to the sequence of the GCC fragment in nuclear extracts, and the binding activity was dramatically decreased by the mutations in the GCC boxes (Figure 2A). These results suggested that the nuclear factors interacted specifically with the GCC box, because the difference in sequence between GCC and mGCC was limited to the two single-base mutations in each GCC box. The binding activity specific to the GCC box of the nuclear factors was further demonstrated in competitive gel retardation assays (Figure 2B). Formation of the DNA–protein complexes between the ERE (GCC probe) and nuclear proteins was subject to specific competition both by the GCC fragment and by a 146-bp ethylene-responsive region from the tobacco chitinase *Chn48* gene (CH48) (Shinshi et al., 1990; H. Shinshi, S. Usami, and M. Ohme-Takagi, submitted manuscript), which contained two copies of the GCC box. However, the mutant sequence mGCC with point mutations in the GCC box (mGCC) was not capable of competing for binding to nuclear factors (Figure 2B). The 146-bp region from the *Chn48* gene is not homologous to the GCC fragment except in its GCC box sequences; this is an indication that the nuclear factors interact specifically with the GCC box in the ERE. Similarly, the specific interaction between the GCC box and a nuclear factor has been identified in the ethylene-responsive region of the *Chn48* gene (H. Shinshi, S. Usami, and M. Ohme-Takagi, submitted manuscript).

The DNA–protein complex formed between the GCC fragment and nuclear proteins was observed when poly(dA-dT)·(dA-dT) was present in the binding reactions. However, the interaction was eliminated when poly(dI-dC)·(dI-dC) was used as a nonspecific competitor (data not shown), as reported previously in a similar system (Hart et al., 1993).

Isolation of cDNAs Encoding EREBPs That Bind to the GCC Box

Because we had demonstrated the presence of nuclear factors that interacted specifically with the GCC box, we were able

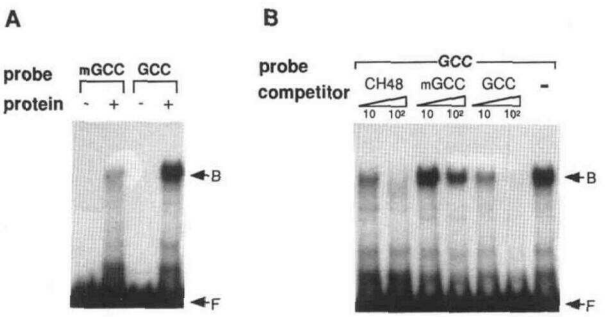


Figure 2. Gel Retardation Assays for the Detection of Binding Activity Specific to the GCC Box in Nuclear Extracts.

(A) Gel retardation assays were performed with nuclear extracts and the GCC or the mGCC fragment to detect proteins that could bind to the ERE. (+), in the presence of or (−), in the absence of nuclear extracts. (B) and (F) indicate DNA–protein complexes and free probe, respectively. Reaction mixtures (10 μL) containing 4 fmol of probe, 2 μg of poly(dA-dT)·(dA-dT), and 10 μg of nuclear protein were loaded on a 4% polyacrylamide gel in 0.25 × TBE after incubation.

(B) Competitive gel retardation assays to detect binding activities specific to the GCC box are shown. Binding reactions contained a molar excess (10- or 100-fold) of competitor DNA: GCC, GCC fragment; mGCC, mGCC fragment; and CH48, a 146-bp fragment from the 5′ upstream region (−503 to −357) of the *Chn48* gene encoding chitinase (Shinshi et al., 1990) that contained two copies of the GCC box. Formation of DNA–protein complexes between nuclear proteins and the GCC probe was subject to competition by a fragment that contained the GCC box but not the mutant GCC box. (B), DNA–protein complex; (F), free probe; (−), no competitor.

to isolate cDNAs that encoded the binding proteins for characterization of the putative transcription factors that regulate ethylene-responsive transcription. A λgt11 expression library constructed with mRNA from ethephon-treated tobacco leaves was screened by a DNA–ligand binding assay using the GCC fragment as probe. Positive clones were further screened by a differential binding assay with GCC and mGCC probes to isolate clones for proteins with the ability to bind to the GCC box. A total of 22 clones for proteins with such binding activity were isolated, and sequence analysis allowed four different cDNAs encoding EREBPs (EREBP-1 to EREBP-4) to be identified. The specificity of binding of EREBPs to the GCC box in a DNA–ligand binding assay is shown in Figure 3, and it was also confirmed by competitive gel retardation assays with extracts of lysogens (data not shown). The numbers of isolated and partially characterized clones related to EREBP-1 to EREBP-4 were two, one, two, and 11, respectively. The schematic representation of EREBP cDNAs and the nucleotide sequence of the EREBP-2 cDNA are shown in Figures 4 and 5, respectively. Each cDNA sequence for an EREBP contained an open reading frame that began with an ATG codon. The cDNAs encoded proteins of 236, 233, 225, and 291 amino acid residues with predicted molecular masses of 26.4, 25.5, 24.2, and 32.8 kD, respectively, in frame with the *lacZ* gene of λgt11. Each cDNA probably contained the entire

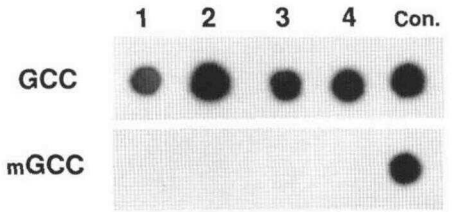


Figure 3. Specificity of Binding to the GCC Box Sequence of EREBPs Encoded by cDNA Clones.

Results are shown for a DNA–ligand binding assay that demonstrates the specific binding of EREBPs to the GCC box. Representative clones that included cDNAs for EREBPs were amplified on a lawn of *E. coli* in duplicate and then transferred to filters. The DNA–ligand binding assay was performed with either the GCC fragment or the mGCC fragment. The numbers 1 to 4 indicate clones (λGC2, λGC9, λGC1, and λGC11) representative of the four group of cDNAs encoding EREBP-1 to EREBP-4. Con. indicates a positive control, λ48, that binds to both the GCC and the mGCC probes and was isolated during the screening procedures.

coding sequence of an EREBP because each EREBP clone hybridized to RNA similar in size to the insert.

Comparison of the amino acid sequences deduced for the four EREBPs from cDNA sequences revealed in each a highly homologous region of 59 amino acid residues that was rich in charged amino acids (Figures 4 and 6). This region was encoded by a sequence near the middle of the coding regions of the cDNAs for EREBP-1, EREBP-2, and EREBP-4. In EREBP-3, this region was located near the N-terminal end. The cDNA for EREBP-4 contained the longest open reading frame among the four cDNAs. EREBP-1 and EREBP-2 shared 68% sequence identity at the amino acid level (Figure 6), and EREBP-1 and EREBP-2 could be assigned to one class. The deduced amino acid sequences of EREBP-1 and EREBP-3 and those of EREBP-1 and EREBP-4 showed no significant sequence identity except in the homologous region, and they can be assigned to different classes.

The DNA Binding Domain

Apart from the 59–amino acid homologous region of EREBPs, no significant sequence homology was found among the four deduced EREBPs. Therefore, it seemed most likely that the DNA binding domain would be found within the homologous region. To examine this possibility, a series of truncated derivatives of an EREBP were synthesized as fusion proteins in *Escherichia coli* by fusing restriction fragments from the region that encoded EREBP-2 (Figures 5 and 7A) to the coding sequence for dihydrofolate reductase in an expression vector. The DNA binding properties of the truncated proteins were then determined. Extracts from *E. coli* cells that expressed the truncated EREBP-2 proteins were fractionated by SDS-PAGE. Proteins were transferred to a nitrocellulose membrane, which was then incubated in the presence of the radiolabeled GCC fragment as probe. As shown in Figure 7B, DNA binding activities were found only in the case of truncated EREBPs that included the conserved region of 59 amino acids (lanes 2, 3, and 5), whereas no binding activity was detected for the truncated EREBPs that lacked the homologous region (lanes 1 and 4). These results indicate that the DNA binding domain of EREBPs is located in the conserved 59–amino acid region of the EREBPs. The specificity of binding to the GCC box in the truncated EREBPs was confirmed with the mutant probe (data not shown).

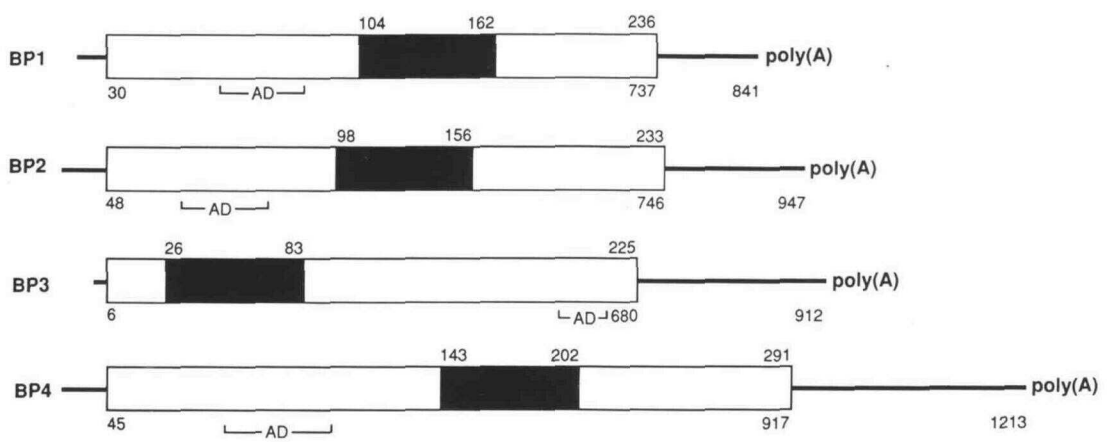


Figure 4. Structure of cDNAs Encoding the Four EREBPs.

A schematic representation of cDNAs for the four EREBPs is shown. BP1 to BP4 indicate cDNAs encoding EREBP-1 to EREBP-4, respectively. Bars indicate encoding open reading frames starting from the first ATG codon, and lines show putative untranslated regions. Filled bars indicate homologous regions found in all four EREBPs. The acidic domains are shown as AD. Numbers above the line indicate numbers of amino acid residues; numbers below the line refer to nucleotides.

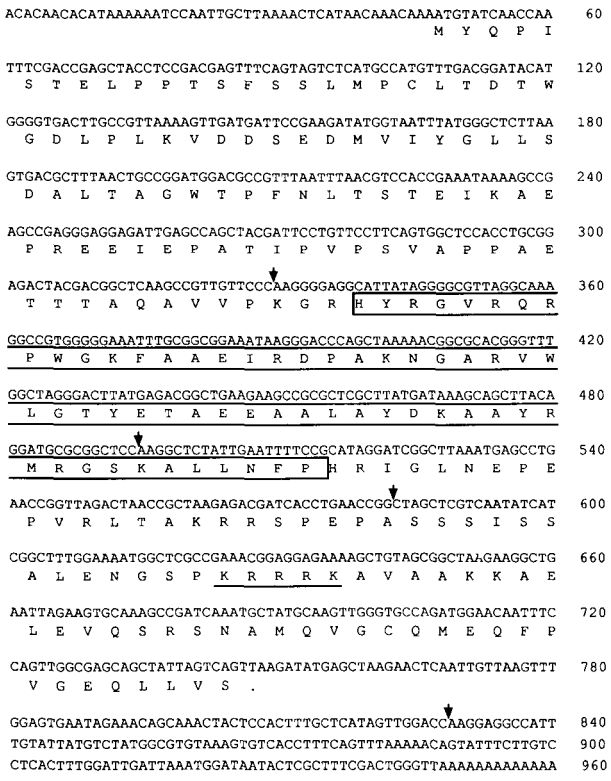


Figure 5. Nucleotide Sequence of the cDNA Encoding EREBP-2 and the Deduced Amino Acid Sequence.

The nucleotide sequence derived from the cDNA for EREBP-2 starts in the 5′ leader region and ends in a poly(A)-containing sequence downstream from a termination codon. The homologous region found in all four EREBPs is boxed. A putative nuclear target signal, KRRRK, is underlined. Restriction sites used for the preparation of truncated proteins (Figure 7) are shown by arrows.

Computer-aided searches of the data bases failed to reveal any significant sequence similarity among EREBPs and any known transcription factors or DNA binding proteins. The EREBPs also did not contain a basic leucine zipper (bZIP) motif or a zinc finger motif, which are common to most plant DNA binding proteins. The secondary structure of the DNA binding domain, as predicted by computer analysis, suggested the possible formation of two α-helical structures, which are separated by regions likely to form a β-strand. However, no structural homology was found with helix-loop-helix or helix-turn-helix structures in the DNA binding domains that have been reported in the data base. We were intrigued to find regions with high homology (53 to 79% identical) to the conserved region of EREBPs in the predicted amino acid sequences for proteins of unknown function from lupine (Perrey et al., 1990), Arabidopsis, and rice (Figure 8), an indication that the DNA binding domain of EREBPs is conserved in plants.

The deduced EREBPs did have some structural features typical of transcription factors. An acidic region, which is probably a transcription activation domain, was found in the EREBP-1, EREBP-2, and EREBP-4 N-terminal regions and in the EREBP-3 C-terminal region (Figures 4 and 6). Putative nuclear target sites composed of a short cluster of basic amino acid residues (KRRRK; Raikhel, 1992) were found in EREBP-1 and EREBP-2 (Figures 5 and 6).

Expression of EREBP mRNAs

Expression of individual EREBP mRNAs was analyzed by RNA gel blot hybridization to compare the expression of genes that encode class I basic β-1,3-glucanase and chitinase (Figure 9). Total RNA was prepared from ethephon-treated or untreated tobacco terminal buds (including the top leaf) and leaves and from root tissues and suspension-cultured cells. After electrophoresis, the RNAs were allowed to hybridize with each of the labeled inserts of cDNAs for EREBPs. Levels of mRNA for all four EREBPs were increased in ethephon-treated plants, but the patterns of expression of individual mRNAs for EREBPs were different. The mRNAs for EREBP-1 and EREBP-2 were expressed at low levels in untreated leaves. The mRNA for EREBP-4 was barely detectable in untreated leaves, but accumulation of this mRNA was strongly induced in leaves by treatment with ethephon. This result corresponds with our observation that cDNAs related to EREBP-4 were in the majority among the clones isolated from the library constructed with mRNA from ethephon-treated leaves. The mRNAs for EREBPs were induced in ethephon-treated terminal buds. This result is in agreement with the results of the gel retardation assays performed by Hart et al. (1993), who found that binding activity of nuclear extracts from the top leaf plus terminal buds to a fragment that included the GCC box was increased by treatment with ethylene.

The accumulation of mRNAs for EREBP-1, EREBP-2, and EREBP-3 was more prominently induced in terminal buds than in leaves. The mRNA for EREBP-3 was expressed in untreated terminal buds (plus the top leaf), in which the level of mRNA for β-1,3-glucanase was extremely low but chitinase mRNA was abundant. The mRNAs for all four EREBPs were expressed in root tissues, in which mRNAs transcribed from genes for class I chitinase and β-1,3-glucanase were expressed at high levels, whereas only mRNA for EREBP-3 was found in tobacco-cultured cells, in which the genes for class I chitinase and β-1,3-glucanase were also expressed at high levels. Chitinase and β-1,3-glucanase have been shown to be coordinately regulated in leaves (Vögeli et al., 1988; Neal et al., 1990). However, in this experiment we observed that levels of mRNAs for chitinase and β-1,3-glucanase were differentially regulated in the terminal buds.

DISCUSSION

The GCC box, which is the conserved sequence in ethylene-inducible PR protein genes (Ohme and Shinshi, 1990; Eyal

Figure 6. Deduced Amino Acid Sequences of Tobacco EREBPs.

Amino acid sequences of the four EREBPs are aligned. The open bar below the sequence data indicates the conserved region of 59 amino acid residues found in all four EREBPs, and filled boxes indicate the identical amino acid residues in the conserved region. Lines below the sequences show the putative acidic domain. Dashes indicate gaps in the amino acid sequences used to optimize the alignment. The nucleotide sequences of the cDNAs encoding EREBPs have been submitted to GenBank, DDBJ, EMBL, and NCBI, with accession numbers D38123 (EREBP-1), D38126 (EREBP-2), D38124 (EREBP-3), and D38125 (EREBP-4).

et al., 1993; Hart et al., 1993), was shown to be the target for the ethylene signal transduction pathway. We showed that a 47-bp fragment that contains two copies of the GCC box can act as a *cis* regulatory element that enhances ethylene-dependent transcription from a heterologous promoter and that point mutations in the GCC box eliminate the ability of the element to activate gene expression in an ethylene-dependent manner. We have conclusively shown by gain-of-function analysis that the GCC box is an ERE, although additional experiments are required to define the minimal sequence that is critical.

It should be noted that other ethylene-inducible genes, such as those related to ripening and senescence (Deikman and Fischer, 1988; Cordes et al., 1989; Cass et al., 1990; Raghothama et al., 1991), do not contain GCC boxes in their 5′ upstream regions. This suggests that a variety of *cis* elements and *trans* factors may be involved in ethylene regulation.

In the experiments presented in this study, we determined that the interactions between nuclear factors and the ERE were sensitive to point mutations in the GCC box. Results of assays of binding of proteins to the element in vitro were correlated with results of functional analysis of the element in vivo, indicating that the nuclear factors that interacted with the GCC box are candidates for the *trans*-acting factors through which ethylene-responsive transcription is regulated. DNA–ligand binding screening of a tobacco cDNA library led to the identification of four independent cDNA sequences that encode EREBPs. On the basis of sequence comparisons, EREBP-1 and EREBP-2 can be assigned to one class, and EREBP-3 and EREBP-4 can be assigned to other classes, indicating that multiple classes of the binding proteins that share the same target sequence are encoded in the tobacco genome. The sequence similarity among members of the classes is restricted to the region of 59 amino acids that is the DNA binding domain. The limited similarity outside the conserved region suggests the possibility of a different mode of activation of each EREBP.

EREBPs are novel DNA binding proteins that are unrelated to the bZIP and zinc finger families, and they exhibit no sequence homology with known transcription factors or DNA binding proteins. The DNA binding domain of EREBPs was revealed to be a novel domain that has no structural homology with DNA binding domains that have been reported in the data base. Most of the known plant-derived bZIP proteins share DNA binding specificity for motifs that contain the core sequence ACGT and have a highly conserved basic region that is predicted to be the DNA binding domain (Izawa et al., 1993; Hurst, 1994). The DNA binding domain identified in EREBP

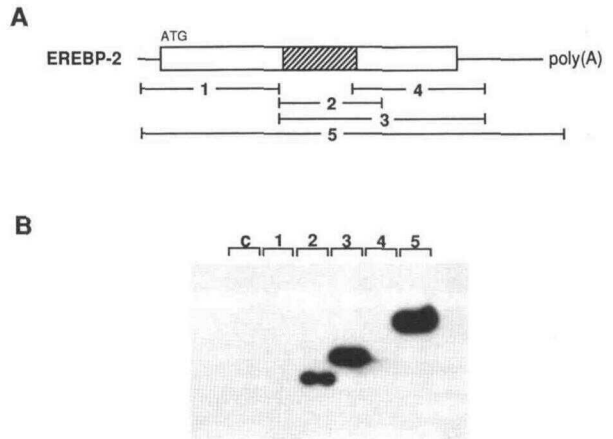


Figure 7. Detection of the DNA Binding Domain.

(A) A schematic representation of the coding regions for the truncated EREBP-2 is shown. Restriction fragments of the EREBP-2 cDNA (1, 1 to 332; 2, 329 to 586; 3, 329 to 831; 4, 494 to 831; and 5, 1 to 960) shown in Figure 5 were inserted into the expression vector to prepare the truncated EREBPs encoded by the regions shown above.

(B) Protein gel blots were probed with DNA for the detection of the DNA binding domain. Extracts of transformed *E. coli* cells were loaded on a polyacrylamide gel, transferred after electrophoresis to a nitrocellulose membrane, and incubated with the GCC probe. Each lane contains a truncated EREBP-2 protein that corresponds to the regions shown in **(A)**. Lane C, control extract from *E. coli* cells that had been transformed with the expression vector without an insert.

was found to be conserved in deduced proteins from various plants; it was found in both dicots and monocots, suggesting that it is evolutionarily conserved in plants and, moreover, that EREBPs represent a new class of DNA binding proteins essential for the regulation of transcription.

The EREBPs exhibited different patterns of expression. The accumulation of all EREBPs was shown to be inducible by ethylene in leaves. The mRNAs for EREBP-1 and EREBP-2 were expressed at low levels in untreated leaves, whereas the mRNA for EREBP-4 was barely detectable. Only mRNA for EREBP-3 was expressed in the cultured cells. The mRNAs for all the EREBPs accumulated to high levels in root tissues. The presence of multiple forms of sequence-specific DNA binding proteins with different patterns of expression raises questions about their respective functions. Tomato genes for heat stress transcription factors have been shown to exhibit different patterns of expression (Scharf et al., 1990). One of these genes is constitutively expressed, and expression of the other two is inducible by heat shock (Scharf et al., 1990). It is likely that changes in the levels and composition of *trans*-acting factors in the cells contribute to the sensitivity, amplitude, and specificity of the transient and sustained transcription of genes in response to exogenous signals. Results for the inducibility by ethylene of the accumulation of the mRNAs for EREBPs raise the question of how the levels of these mRNAs are regulated, which remains to be addressed. One possibility is that the expression of genes for EREBPs is autoregulated at the transcriptional level. Autoregulation has been shown to control the expression of several genes that encode transcription factors (Serfling, 1989).

Currently, it is not known whether transactivation activities of EREBPs are modulated by an ethylene signaling cascade. It is unlikely that DNA binding activity is positively regulated by ethylene signaling because we have demonstrated that EREBPs produced in *E. coli* showed sequence-specific DNA binding activity. Although we cannot rule out other possibilities, it is tempting to speculate that the ethylene-responsive transcription of certain genes in leaves initially involves post-translational modification of factors such as EREBP-1 and EREBP-2 that allows them to activate transcription, whereas in the continuous presence of ethylene, newly formed factors such as EREBP-4 may be required to facilitate sustained transcription.

Recently, two ethylene regulator genes, *ETR1* and *CTR1*, in Arabidopsis were identified and shown to encode putative protein kinases (Chang et al., 1993; Kieber et al., 1993). The

```
EREBP-2     TAQAVVPKGR HYRGVRQRPWGKFAAEIRDPAKNGARVWLGTYETAEEAALAYDKAAYRMRGSKALLNFP HRIGLNEPEP   166
ARABI-1     EPGKRRKRKN V:::I:KH::KW-:::::::R:-:V::::::FN::::::M:::V::KQI::D::K:::: DLHHPPPNYT    80
ARABI-2     EVPNHVDTRK PH::::R:K:R:WV::::E:N:-RS:L:::S.T:DIA::R:::V:VFYL::PSTR:::: DLLLKEEAHL    92
LUPINE          MARPQQ R:::F:::H::SWVS:::HSIL-KT:I:Q::F:S::D::R:::E::RL:C:TR:RT::: YNPNASQSSS    72
RICE-1      GKGGPENSNC A:::::::T:::WV::::PNR--:R.L:::SFP::L:::H:::E::RA:Y:PT:RV::A DNSTDSNDGC    79
RICE-2      AGGGVGAGGP ::::::K::::RY:::::::::-KS::::::::DK:::::R:::D                            99
Consensus                .YRG.R.R.WG...AEIR.P.K-..R.WLGT..TAEEAA.AYD.AA....G..A..NFP
                         ****** ****  *******     ******  *  *** *** **   ** **  ***
```

Figure 8. *The DNA Binding Domain of EREBP Is Conserved in Plants.*

The conserved region of EREBPs represented by EREBP-2 is shown with flanking sequences and is aligned with amino acid sequences deduced from cDNAs derived from Arabidopsis (ARABI), lupine (LUPINE), and rice (RICE). cDNA sequence data for Arabidopsis and rice are from the GenBank, EMBL, and DDBJ data bases (accession numbers: ARABI-1, T0430; ARABI-2, Z2705; RICE-1, D23249; and RICE-2, D23250), and the lupine cDNA sequence is from Perrey et al. (1990). Double dots indicate amino acid residues identical to those in EREBP-2; asterisks, identical amino acid residues in the four EREBPs as shown in Figure 6; dashes, gaps included to optimize the alignment; single dots, the lack of consensus among the amino acid sequences; numbers, the amino acid residues from the first methionine or the first amino acid of the reading frame.

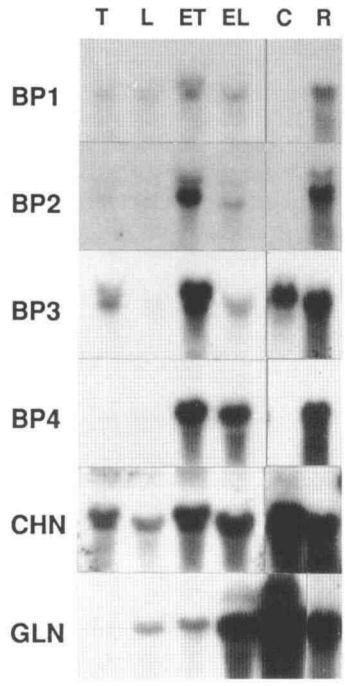


Figure 9. RNA Gel Blot Analysis of EREBP mRNA Expression.

Total RNAs were prepared from tobacco plants with or without ethephon treatment and from tobacco suspension-cultured BY2 cells. Ethephon treatment of tobacco plants was performed by spraying the leaves with a solution of ethephon (1 mM), covering them with a plastic container, and allowing the leaves to stand for 16 hr at 25°C. The RNAs are from the terminal bud plus top leaf (T); lower leaves (L); ethephon-treated terminal bud plus top leaf (ET); ethephon-treated lower leaves (EL); suspension-cultured cells 5 days after subculturing (C); and root tissue of untreated plants (R). The probes are as follows: BP1 to BP4, cDNAs for EREBP-1 to EREBP-4, respectively; CHN, fragment of chitinase cDNA from pCHN50 (Shinshi et al., 1987); and GLN, fragment of β-1,3-glucanase cDNA from pGL43 (Shinshi et al., 1988).

predicted product of *CTR1* was shown to be most similar to members of the Raf family of protein kinases, which activate mitogen-activated protein (MAP) kinase (Kieber et al., 1993). The yeast SLN1 (synthetic lethal of N-end rule) protein, which, like Arabidopsis ETR1, is a eukaryotic example of a constituent of a two-component system, has been suggested to regulate the MAP kinase pathway (Ota and Varshavsky, 1993; Maeda et al., 1994). Among the direct targets of the MAP kinases that have been identified in animal systems are several transcription factors (Egan and Weinberg, 1993). Inhibitors of protein kinase have been shown to block the ethylene-induced transcription of PR genes in tobacco (Raz and Fluhr, 1993). Accumulating evidence suggests that a protein phosphorylation cascade is involved in the ethylene signal transduction pathway. Phosphorylation has been shown to modulate the interaction of transactivation domains of transcription factors with the transcriptional machinery (Hunter and Karin, 1992).

It is now necessary to determine whether EREBPs are signal-activated transcription factors or signal-induced factors, as well as whether their activities are modulated by phosphorylation and/or dephosphorylation. The availability of the cDNA sequences should allow us to address these questions.

METHODS

Preparation of the GCC and mGCC Fragments

The GCC fragment was prepared by synthesizing both strands of the 5′ upstream region of the tobacco *Gln2* gene encoding β-1,3-glucanase (nucleotides −1164 to −1118; Ohme and Shinshi, 1990) with substitutions of T to A (−1161) and A to G (−1128) and with HindIII-BglII and BamHI restriction endonuclease sites at the 5′ and 3′ ends, respectively. The mutated fragment (mGCC) was synthesized with two single-base substitutions in the sequence of the GCC fragment (TAAGA*T*CC*T*CC). The separate fragments were dimerized in tandem and inserted into pUC18 that had been digested with HindIII and BamHI to yield p306.6 (GCC) and p353.1 (mGCC), respectively. The HindIII-BamHI fragments from these clones were used for experiments.

Construction of Chimeric Genes

The TATA region of the cauliflower mosaic virus (CaMV) 35S promoter (−46 to +8) was synthesized with BglII and BamHI restriction endonuclease sites at the 5′ and 3′ ends, respectively, and was fused with the HindIII-BamHI fragment from p306.6 or p353.1 to yield chimeric promoters. Chimeric gene constructs of 2(GCC)Gus and 2(mGCC)Gus were generated by the insertion of the chimeric promoter into pBI101 at the 5′ end of the β-glucuronidase (*Gus*) reporter gene, in which the translation initiation sequence had been replaced with the consensus translation initiator sequence (ACCATGGTC).

Plant Transformation and GUS Assay

Constructs were mobilized into *Agrobacterium tumefaciens* strain LBA4404, and transgenic tobacco (*Nicotiana tabacum* cv BY4) plants were generated by the method of Horsch et al. (1985). Leaf discs from transgenic plants were floated on 50 mM phosphate buffer, pH 7.0, supplemented with or without 1 mM ethephon (2-chloroethylphosphonic acid) and were incubated at 25°C for 16 hr. GUS activity in tissue extracts was assayed by the fluorometric method of Jefferson et al. (1987).

Gel Retardation Assays

Nuclear extracts were prepared essentially as described by Green et al. (1989) from leaves of tobacco (*N. tabacum* cv BY4) that had been sprayed with a solution of ethephon (1 mM), covered with a plastic container, and allowed to stand for 16 hr at 25°C (ethephon treatment). The DNA probe was labeled by filling in with the Klenow fragment of DNA polymerase I. DNA binding reactions were performed in 10 μL of a mixture that contained 4 fmol of probe, 2 μg of poly(dA-dT)·(dA-dT), 10 μg of nuclear protein, and a 10- or 100-fold molar excess of a specific competitor DNA in DNA binding buffer (buffer B: 25 mM

Hepes-KOH, pH 7.5, 40 mM KCl, 0.1 mM EDTA, 10% [v/v] glycerol, 1 mM DTT). After 15 min at room temperature, the reaction mixture was subjected to electrophoresis on a 4% polyacrylamide gel in 0.25 × TBE (22.5 mM Tris-borate, pH 8.0, 0.25 mM EDTA). The gel was then dried and exposed to x-ray film for autoradiography.

Screening of a cDNA Expression Library

A tobacco cDNA library was constructed using the λgt11D expression vector (Pharmacia LKB Biotechnology, Uppsala, Sweden) with an oligo(dT) primer and poly(A)$^+$ RNA prepared from total RNA that had been isolated from tobacco leaves treated with a 1 mM solution of ethephon. One million phages were screened by a DNA–ligand binding assay essentially as described by Singh et al. (1988). The procedure included the addition of a denaturation/renaturation step (Vinson et al., 1988). Filters were incubated for 1 hr in buffer B supplemented with 5% nonfat dry milk; the DNA binding assay was then performed in buffer B that contained the labeled fragment (4 ng/mL), 20 μg/mL poly(dA-dT)·(dA-dT), 20 μg/mL denatured calf thymus DNA, and 0.25% nonfat dry milk. After incubation at 20°C for 1 hr, filters were washed twice for 10 min each with buffer B containing 0.25% nonfat dry milk and then subjected to autoradiography. Positive clones were taken through four rounds of purification and were also screened in duplicate with either the GCC or the mGCC fragment as probe for isolation of clones with binding activity specific for the GCC box sequence. λ clones λGC1 to λGC22, which were positive for binding to the GCC probe and not to the mGCC probe, were further analyzed.

Sequence Analysis

Inserts of λGC clones were digested with NotI and EcoRI or NotI and SfiI and subcloned into the cloning site of pT7T3D (Pharmacia). DNA sequence analysis was performed with an automated sequencer (model 370A; Applied Biosystems, Inc., Foster City, CA) with *Taq*I polymerase. The nucleotide sequence and deduced amino acid sequence were compared with sequences in the GenBank, EMBL, and DDBJ data bases.

Preparation of Truncated Proteins and Protein Gel Blot Analysis

The restriction fragments of the cDNA encoding EREBP-2 shown in Figures 5 and 7 were ligated into the cloning site of the expression vector pQE41 (Qiagen Inc., Chatsworth, CA) in frame with the dihydrofolate reductase gene. Expression of fusion proteins was induced for 3 hr by adding isopropyl β-D-thiogalactopyranoside (final concentration of 2 mM). *Escherichia coli* cells in which truncated ethylene-responsive element binding protein fusion proteins had been expressed were harvested by centrifugation and resuspended in SDS loading buffer (1% SDS, 1% β-mercaptoethanol, 10 mM Tris-Cl, pH 6.8, 20% glycerol); they were then loaded onto an SDS–polyacrylamide gel (10% polyacrylamide). After electrophoresis, proteins were transferred to a nitrocellulose filter in the transfer buffer (100 mM Tris, 192 mM glycine) using a semi-dry blotter (Bio-Rad). Proteins on the filters were subjected to denaturation/renaturation and were incubated with a probe in the same binding buffer as previously described for the screening procedures. After incubation for 1 hr, filters were washed twice for 10 min each with binding buffer and then subjected to autoradiography. After autoradiography, each filter was stained with Coomassie Brilliant Blue R 250 to check for the expression of the fusion proteins.

Analysis of RNA

Isolation of total RNA from tobacco leaves (*N. tabacum* cv BY4) and suspension-cultured cells (BY2) and RNA blot analyses were performed as described elsewhere (Fukuda et al., 1991). Twenty micrograms of total tobacco RNA isolated from different tissues and organs was fractionated by electrophoresis on a formaldehyde gel, transferred to a nylon membrane, and allowed to hybridize to a ^{32}P-labeled fragment. BamHI and EcoRI restriction fragments of pGCs were labeled with a random priming kit (Boehringer Mannheim) and used as probes for hybridization. Equal loading of samples was confirmed by visualizing of RNA by ethidium bromide staining.

ACKNOWLEDGMENT

We thank Dr. Akinori Sarai (RIKEN, Tsukuba) for the computer analysis of EREBPs.

Received September 14, 1994; accepted November 29, 1994.

REFERENCES

Abeles, F.B., Morgan, P.W., and Saltvet, M.E., Jr. (1992). Ethylene in Plant Biology, 2nd ed. (San Diego, CA: Academic Press, Inc.).

Broglie, K.E., Biddle, P., Cressman, R., and Broglie, R. (1989). Functional analysis of DNA sequences responsible for ethylene regulation of a bean chitinase gene in transgenic tobacco. Plant Cell **1**, 599–607.

Cass, L.G., Kirven, K.A., and Christoffersen, R.E. (1990). Isolation and characterization of a cellulase gene family member expressed during avocado fruit ripening. Mol. Gen. Genet. **223**, 76–86.

Chang, C., Kwok, S.F., Bleecker, A.B., and Meyerowitz, E.M. (1993). *Arabidopsis* ethylene-response gene *ETR*1: Similarity of product to two-component regulators. Science **262**, 539–544.

Cordes, S., Deikman, J., Margossian, L.J., and Fischer, R.L. (1989). Interaction of a developmentally regulated DNA-binding factor with sites flanking two different fruit-ripening genes from tomato. Plant Cell **1**, 1025–1034.

Deikman, J., and Fischer, R.L. (1988). Interaction of a DNA binding factor with the 5′-flanking region of an ethylene-responsive fruit ripening gene from tomato. EMBO J. **7**, 3315–3320.

Deikman, J., Kline, R., and Fischer, R.L. (1992). Organization of ripening and ethylene regulatory regions in fruit-specific promoter from tomato (*Lycopersicon esculentum*). Plant Physiol. **100**, 2013–2017.

Egan, S.E., and Weinberg, R.A. (1993). The pathway to signal achievement. Nature **365**, 781–783.

Eyal, Y., Meller, Y., Lev-Yadun, S., and Fluhr, R. (1993). A basic-type PR-1 promoter directs ethylene responsiveness, vascular and abscission zone-specific expression. Plant J. **4**, 225–234.

Felix, G., and Meins, F., Jr. (1987). Ethylene regulation of β-1,3-glucanase in tobacco. Planta **172,** 386–392.

Fukuda, Y., Ohme, M., and Shinshi, H. (1991). Gene structure and expression of a tobacco endochitinase gene in suspension-cultured tobacco cells. Plant Mol. Biol. **16,** 1–10.

Green, P.J., Kay, S.A., Lam, E., and Chua, N.-H. (1989). In vitro DNA footprinting. In Plant Molecular Biology Manual, B11, S.B. Gelvin, R.A. Schilperoort, and D.P.S. Verma, eds (Dordrecht, The Netherlands: Kluwer Academic Publishers), pp. 1–22.

Hamilton, A.J., Bouzayen, M., and Grierson, D. (1991). Identification of tomato gene for ethylene-forming enzyme by expression in yeast. Proc. Natl. Acad. Sci. USA **88,** 7434–7437.

Hart, C.M., Nagy, F., and Meins, F., Jr. (1993). A 61 bp enhancer element of the tobacco β-1,3-glucanase B gene interacts with one or more regulated nuclear proteins. Plant Mol. Biol. **21,** 121–131.

Holdsworth, M.J., and Laties, G.G. (1989). Site-specific binding of a nuclear factor to the carrot extensin gene is influenced by both ethylene and wounding. Planta **179,** 17–23.

Horsch, R.B., Fry, J.E., Hoffmann, N.L., Eichholtz, D., Rogers, S.G., and Fraley, R.T. (1985). A simple and general method for transferring genes into plants. Science **227,** 1229–1231.

Hunter, T., and Karin, M. (1992). The regulation of transcription by phosphorylation. Cell **70,** 375–387.

Hurst, H.C. (1994). Transcription factors 1: bZIP proteins. Protein Profile **1,** 133–137.

Izawa, T., Foster, R., and Chua, N.-H. (1993). Plant bZIP protein DNA binding specificity. J. Mol. Biol. **230,** 1131–1144.

Jefferson, R.A., Kavanagh, T.A., and Bevan, M.W. (1987). GUS fusions: β-Glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. **6,** 3901–3907.

Kieber, J.J., Rothenberg, M., Roman, G., Feldmann, K.A., and Ecker, J.R. (1993). *CTR1*, a negative regulator of the ethylene response pathway in Arabidopsis, encodes a member of the Raf family of protein kinases. Cell **72,** 427–441.

Lawton, K.A., Potter, S.L., Uknes, S., and Ryals, J. (1994). Acquired resistance signal transduction in Arabidopsis is ethylene independent. Plant Cell **6,** 581–588.

Maeda, T., Wurgler-Murphy, S.M., and Saito, H. (1994). A two-component system that regulates an osmosensing MAP kinase cascade in yeast. Nature **369,** 242–245.

Meller, Y., Sessa, G., Eyal, Y., and Fluhr, R. (1993). DNA-protein interactions on a *cis*-DNA element essential for ethylene regulation. Plant Mol. Biol. **23,** 453–463.

Nakajima, N., Nakagawa, N., and Imaseki, H. (1990). Molecular size of wound-induced 1-aminocyclopropane-1-carboxylate synthase from *Cucurbita maxima* Duch. and change of translatable mRNA of the enzyme after wounding. Plant Cell Physiol. **29,** 989–998.

Neal, A.D., Wahleithner, J.A., Lund, M., Bonnett, H.T., Kelly, A., Meeks-Wagner, D.R., Peacock, W.J., and Dennis, E.S. (1990). Chitinase, β-1,3-glucanase, osmotin, and extensin are expressed in tobacco explants during flower formation. Plant Cell **2,** 673–684.

Ohme, M., and Shinshi, H. (1990). Structure and expression of a tobacco β-1,3-glucanase gene. Plant Mol. Biol. **15,** 941–946.

Ota, I.M., and Varshavsky, A. (1993). A yeast protein similar to bacterial two-component regulators. Science **262,** 566–569.

Perrey, R., Warskulat, U., and Wink, M. (1990). Molecular cloning of a lupin-specific gene from a cDNA library of suspension-cultured cells of *Lupinus polyphyllus*. Plant Mol. Biol. **15,** 175–176.

Raghothama, K.G., Lawton, K.A., Goldbrough, P.B., and Woodson, W.R. (1991). Characterization of an ethylene-regulated flower senescence-related gene from carnation. Plant Mol. Biol. **17,** 61–71.

Raikhel, N. (1992). Nuclear targeting in plants. Plant Physiol. **100,** 1627–1632.

Raz, V., and Fluhr, R. (1993). Ethylene signal is transduced via protein phosphorylation events in plants. Plant Cell **5,** 523–530.

Sato, T., and Theologis, A. (1989). Cloning the mRNA encoding 1-aminocyclopropane-1-carboxylase, the key enzyme for ethylene biosynthesis in plants. Proc. Natl. Acad. Sci. USA **86,** 6621–6625.

Scharf, K.-D., Rose, S., Zott, W., Schöff, F., and Nover, L. (1990). Three tomato genes code for heat stress transcription factors with a region of remarkable homology to the DNA binding domain of the yeast HSF. EMBO J. **9,** 4495–4501.

Serfling, E. (1989). Autoregulation—A common property of eukaryotic transcription factors? Trends Genet. **5,** 131–133.

Shinshi, H., Mohnen, D., and Meins, F., Jr. (1987). Regulation of a plant pathogenesis-related enzyme: Inhibition of chitinase and chitinase mRNA accumulation in cultured tobacco tissues by auxin and cytokinin. Proc. Natl. Acad. Sci. USA **84,** 89–93.

Shinshi, H., Wenzler, H., Neuhaus, J.M., Felix, G., Hofsteenge, J., and Meins, F., Jr. (1988). Evidence for N- and C-terminal processing of a plant defense-related enzyme: Primary structure of tobacco prepro-β-1,3-glucanase. Proc. Natl. Acad. Sci. USA **85,** 5541–5545.

Shinshi, H., Neuhaus, J.M., Ryals, J., and Meins, F., Jr. (1990). Structure of a tobacco endochitinase gene: Evidences that different chitinase genes can arise by transcription of sequence encoding cysteine-rich domain. Plant Mol. Biol. **14,** 357–368.

Singh, H., LeBowitz, J.H., Baldwin, A.S., Jr., and Sharp, P.A. (1988). Molecular cloning of an enhancer binding protein: Isolation by screening of an expression library with a recognition site DNA. Cell **52,** 415–423.

Spanu, P., Reinhardt, D., and Boller, T. (1991). Analysis and cloning of the ethylene-forming enzyme from tomato by functional expression of its mRNA in *Xenopus laevis* oocytes. EMBO J. **10,** 2007–2013.

Van Der Straeten, D., Van Wiemeersch, L., Goodman, H.M., and Van Montagu, M. (1990). Cloning and sequence of two different cDNAs encoding 1-aminocyclopropane-1-carboxylate synthetase in tomato. Proc. Natl. Acad. Sci. USA **87,** 4859–4863.

Vinson, C.R., LaMarco, K.L., Johnson, P.F., Landschultz, W.H., and McKnight, S.L. (1988). In situ detection of sequence-specific DNA binding activities specified by a recombinant bacteriophage. Genes Dev. **2,** 801–806.

Vögeli, U., Meins, F., Jr., and Boller, T. (1988). Co-ordinate regulation of chitinase and β-1,3-glucanase in bean leaves. Planta **174,** 364–372.

Vögeli-Lange, R., Fründt, C., Hart, C.M., Nagy, F., and Meins, F., Jr. (1994). Developmental, hormonal, and pathogenesis-related regulation of the tobacco class I β-1,3-glucanase B promoter. Plant Mol. Biol. **25,** 299–311.

Yang, S.F. (1969). Ethylene evolution from 2-chloroethylphosphonic acid. Plant Physiol. **44,** 1203–1204.

Yang, S.F., and Hoffman, N.E. (1984). Ethylene biosynthesis and its regulation in higher plants. Annu. Rev. Plant Physiol. **35,** 155–189.