A Plant Leucine Zipper Protein That Recognizes an Abscic Acid Response Element

MARK J. GIULITIANI, WILLIAM R. MARCOTTE, JR., RALPH S. QUATRANO*

The mechanism by which phytohormones, like abscisic acid (ABA), regulate gene expression is unknown. An activity in nuclear extracts that interacts with the ABA response element (ABRE) from the 5′ regulatory region of the wheat Em gene was identified. A complementary DNA clone was isolated whose product is a DNA binding protein (EmBP-1) that interacts specifically with an 8-base pair (bp) sequence (CACGTGGG) in the ABRE. A 2-bp mutation in this sequence prevented binding of EmBP-1. The same mutation also reduced the ability of the ABRE to confer ABA responsiveness on a viral promoter in a transgenic assay. The 8-bp EmBP-1 target sequence was found to be conserved in several other ABA-responsive promoters and in promoters from plants that respond to signals other than ABA. Similar sequences are found in promoters from mamals, yeast, and in the major late promoter of adenovirus. The deduced amino acid sequence of EmBP-1 contains conserved basic and leucine zipper domains found in transcription factors in plants, yeast, and mammals. EmBP-1 may be a member of a highly conserved family of proteins that recognize a core sequence found in the regulatory regions of various genes that are integrated into a number of different response pathways.

RESPONSES OF CELLS TO EXTERNAL stimuli (such as light, hormones, and environmental stress) are mediated in part by the expression of genes whose products contribute to a given physiological effect. Studies with animal hormones have elucidated several response pathways that ultimately converge at the level of gene expression (1, 2). These pathways have been deciphered in part through identification of the protein factors that bind regulatory elements in hormone-responsive genes. The distinct transcriptional regulatory patterns of genes expressed during these responses are determined primarily by specific interactions between protein and DNA (1, 2).

During plant development, hormones influence fruit ripening, seed maturation and germination, shoot and root growth, and responses to environmental and pathogenic stresses (3). However, the response pathway for any one phytohormone has not yet been elucidated. All classes of phytohormones influence the expression of specific genes, at least in part at the level of transcription (4, 5). Related sequences exist in the 5′-upstream regions of genes similarly regulated either by abscisic acid (ABA), gibberellic acid (GA), ethylene, or auxin (6-8). DNA regulatory elements have been functionally identified in promoters of genes responsive to ABA (9-11), GA (12), and ethylene (7).

Like most phytohormones, ABA mediates diverse physiological responses, including promotion of embryogenesis and prevention of precocious germination (5, 13). Late in seed development of a wide variety of plants, a unique set of abundant mRNAs and proteins appear, coincident with high concentrations of endogenous ABA. These late embryonic abundant (Lea) gene products share common physical properties and accumulate in mature embryos (13, 14). If embryos are isolated at earlier developmental stages and exposed to exogenous ABA, some of the Lea class mRNAs and proteins accumulate precociously (5, 13-15). The concentration of ABA also increases when plant tissues are stressed by dessication (16), wounding (17), or high osmoticum (18), resulting in the expression of some of the same Lea genes in nonembryonic tissues (13). We approached the question of how ABA exerts its effect at the level of the gene by defining regulatory elements in the ABA-responsive promoter of the wheat Em gene (9, 10, 19).

Accumulation of mRNA from Em (20, 21) is regulated by ABA during embryoid development and under stress conditions (13, 19) in a manner similar to other Lea genes. When the Em 5′ regulatory region was linked to the reporter gene, β-galactosidase (GUS), and used in transient (9) and transgenic (10) assays, a 646-bp region (−554 to +92) that was essential for response to ABA was identified. Within this region, a 50-bp ABA response element (ABRE) (−152 to −103) has been defined that is capable of conferring ABA inducibility upon a minimal cauliflower mosaic virus (CaMV) promoter (16). Two elements (Em1 and Em2) within this 50-bp ABRE are conserved in other ABA-regulated promoters (10), including the rice Rub and the cotton Lea gene families (8, 11). We now describe the identification of a protein that interacts with the ABRE.

Nuclear extracts from mature wheat embryos and rice cells in suspension cultures contained proteins that bound a 32P-labeled 119-bp DNA fragment (ABRE probe) that contained the 50-bp response element (Fig. 1) (22). Two major protein-DNA complexes (B1 and B2) were specifically competed by unlabeled DNA fragments that contained the ABRE (Fig. 1B). Neither the Em nucleotide sequence nor an ABRE that contained a 2-bp mutation (mABRE) was capable of competing for the binding activity (Fig. 1B). We noticed that the recognition site (Hex) for the wheat transcription factor HBP-1 (23) contains sequences that are similar to the ABRE (Fig. 1A). A DNA fragment that contained the Hex sequence was capable of competing with the ABRE probe for binding activity (Fig. 1B), but did not compete as well as the ABRE fragment (Fig. 1B).

Department of Biology, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3280.

*To whom correspondence should be addressed.

13 OCTOBER 1990  REPORTS 267
Experiments made from ABA-treated rice suspension cells reproducibly contained higher amounts of binding activity than cells cultured without ABA (Fig. 1B, *). In extracts from ABA-treated rice leaves, a slight increase in binding activity relative to untreated tissue was also observed with the Rab16A ABRE (ff). The significance of the differences in binding activity in untreated samples of ABA-responsive tissue is unknown.

In order to correlate in vitro binding activity with in vivo function, transient assays were performed with reporter gene constructs that contained either the ABRE or the mABRE (24). Substitution of two bases within the Em1a box (mABRE) reduced or eliminated the ability of the ABA response element to activate gene expression in an ABA-dependent manner (Table 1). The ABRE-containing construct resulted in approximately a 12-fold increase in expression upon addition of ABA, whereas the mABRE increased twofold in response to ABA.

A cDNA expression library was made from wheat embryo cDNA and screened for proteins capable of binding a 32P-labeled double-stranded ABRE probe (25). From a screen of 120,000 recombinant phages, two overlapping clones were isolated, AGC12 and AGC19, with insert sizes of 528 and 569 bp, respectively. To assess the specificity of the AGC19-encoded DNA binding activity, gel retardation assays were performed.
Table 2. Comparison of promoter elements in plant genes that respond to different signals.

<table>
<thead>
<tr>
<th>Signal</th>
<th>Gene Sequence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABA</td>
<td>Em(1a) CAGCTGGA*</td>
<td>Fig. 1</td>
</tr>
<tr>
<td></td>
<td>Tritium GAGCTGGGC (30)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Arab.1a CAGCTGGA* (30)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Light G-box CAGCTTGC (34, 35)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wound PI-II CAGCTGGA (31)</td>
<td></td>
</tr>
</tbody>
</table>

*Indicates sequences that functionally assayed and footprinted. Genes include: ribulose bisphosphate carboxylase (rubisco), glutamine synthase, alcohol dehydrogenase.

A series of five leucine zipper repeats not found in HBP-1 or other leucine zipper-containing proteins (23, 27, 28). However, each of these repeats is interrupted with a proline residue, which are not found in leucine zipper structures (27). An oligo(dT)-primed wheat embryo cDNA library was screened by hybridization to the GC19 cDNA fragment, and one clone (GC13) was obtained that appeared to be full length (26). The DNA sequence of GC13 and GC19 share 98% sequence identity and differ by five amino acids (97% identity). GC13 contains one large open reading frame that encodes 354 amino acids with a predicted molecular size of 36.2 kilodaltons. GC13 also contains a 3' untranslated region of 250 bp with a poly(A)+ trace of 1.5 bp. Also present in the N-terminal half of the predicted protein sequence is a region relatively rich in proline, a characteristic found in some transcriptional activating domains (29).

Our results show that an 8-bp sequence within the ABA from wheat Em (Em1a) was protected by EmBP-1. This sequence is conserved in ABA-regulated genes from wheat (Em, tritium) (30), rice (Rab) (8), and cotton (Lea) (14) (Table 2). Furthermore, substitution of the 2 bp in Em1a that inhibited EmBP-1 binding, also reduced the ability of the ABA to direct ABA-dependent gene expression in the transient assay. The combination of these expression, footprint, and sequence analyses for several ABA-responsive genes in monocots and dicots suggests that EmBP-1 is involved in the ABA response.

Wound- or drought-induced genes (14, 17, 18) may be activated by changes in the endogenous amount of ABA, which rises in response to these stresses (15, 17). ABA application can independently activate the wound-induced protease inhibitor II (PI-II) genes from tomato and potato (17). The ABA consensus was found at -577 in the PI-II promoter (31) (Table 2). Similar sequences were found in the wound-inducible genes wun1, wun1, and wun2 from potato, and wun3 from poplar trees (32), suggesting that the wound response of these genes may be mediated by ABA.

In addition, the 8-bp ABAE exactly matches the consensus G-box motif (Table 2), which was found in a number of yeast promoters (33), plant promoters regulated by visible and ultraviolet light (34), as well as in the anacrotically induced ADH-1 promoter from maize (35). This conserved sequence is important for transcription of some of these genes (33, 34), but in others appear to be positively regulated by ABA. It was proposed that the G-box binding factor (GBF), which recognizes the G-box in several unrelated genes in yeast and plants, was directly involved in their expression (33, 34).

Our results suggest that the set of genes that contains this core sequence can be extended to include certain ABA-regulated genes and, perhaps, genes induced by stress that have an ABA-intermediate in their response pathway (Table 2).

A sequence similar to ABAE/G-box is also found in bacterial, viral, and mammalian promoters. The Hex sequence (Fig. 1A),

leucine heptad repeat region (27). The basic region of EmBP-1 contains eight of the ten conserved residues found in other leucine zipper proteins (Fig. 3) (27). Additionally, 26 of 32 residues are conserved between the basic regions of HBP-1 and EmBP-1, consistent with our finding that the Hex recognition sequence for HBP-1 is capable of competing for binding of EmBP-1 to the ABAE (Fig. 1). The 8-bp recognition sequence of EmBP-1 is an imperfect palindromic (Em1a, Table 2), consistent with other leucine zipper protein target sequences and the proposed structural model for leucine zipper protein-DNA complex formation (27). Interestingly, the sequence directly N-terminal to the conserved basic region of EmBP-1 contains a second...
REFERENCES AND NOTES

22. The following oligonucleotides were synthesized, treated with T4 polynucleotide kinase, annealed, and extended with Kl/on enzyme. They were then ligated into the plC19 that had been digested with Hind III and flt and ligated with Kl/on enzyme (pMBS200). The orientation of ABRE and the G-box element was confirmed by sequencing. The resulting plC19 fragment was designated pMB309.1 and was purified and flt with Kl/on enzyme. It was then ligated to plC19 that had been digested with Stu I and flt with Kl/on enzyme, resulting in recombinant plasmids designated pMG76.1 and pMG76.155. The insert in these plasmids is in opposite orientation as verified by nucleotide sequence mapping. The ABRE and Kl/on oligonucleotides were synthesized, treated with T4 polynucleotide kinase, annealed, and ligated to plC19 that had been digested with StuI and flt with Kl/on enzyme. The resulting recombinant plasmids were designated pMG77.43 and pMG82.136. Finally, was obtained as a 6.10 kb Pst I DNA fragment from plC19 (sequence not352). ABRE, mABRE, and Hex competitor fragments were verified by DNA sequencing and purified by excision with Eco RI and Pst I followed by gel electrophoresis, densitometry, and autoradiography at high exposure. Nuclear extracts were prepared as described [O. E. Jensen, K. A. Mandic, J. Seidell, F. L. Dehnen, EMBO J. 7, 236 (1988)]. DNA binding analyses were as described for other systems [K. Miki, T. Taniura, T. Kawaoka, T. Nakamura, M. Ishimaru, FEBS Lett. 223, 273 (1987)] except that ethidium bromide was used to separate bound and unbound competitor fragments.
24. The ABRE-containing expression plasmids were constructed as follows: The ABRE was isolated from plC19 (sequence not352).2) The G-box element was isolated from a plasmid (pMB234) prepared for plC209.1. The 0.8 kb Hind III fragments in both 0.2 and uncut was inserted into the Hind III site of pCM231 (10). The expression plasmid was introduced into bean seedlings. The resulting polyamines (apparent molecular weight, 26 kD) were used for binding and footprinting reaction without further purification. 25. I. Ghebrehiwet et al., Science 245, 366 (1989).
26. The cDNA clone was constructed with 10 μg of plasmid (stage II) ABA-expressed小麦embryonic poly(ADP-ribose) RNA, random primer, and The Librarian Lambda (Invitrogen) as described (U. Gubiher and B. J. Hopkins, Gene 25, 283 (1983)). Recombinants (2.64 x 10^5) were amplified, and 100,000 plaques were screened as described (H. Singh, R. G. Cook, J. F. Lobakiwits, Nature, 400 (1989). The expressed mABRE and mABRE in the resulting plasmids (pMB226 and pMB233), wild-type and mutant sequences, respectively, were determined by DNA sequencing. Transformed rice protoplasts were incubated for 24 hours in the presence or absence of 10^-4 M ABA prior to the assay as described. 27. C. A. Nowak, P. B. Sigler, S. L. McKnight, Science 246, 911 (1999).
30. R. Quatrano, unpublished data.
Cleaving Yeast and Escherichia coli Genomes at a Single Site

MICHAEL KOB and WACLAW SZYBALSKI

The 15-megabase pair Saccharomyces cerevisiae and the 4.7-megabase pair Escherichia coli genomes were completely cleaved at a single predetermined site by means of the Achilles' heel cleavage (AC) procedure. The symmetric lac operator (lacO) was introduced into the circular Escherichia coli genome and into one of the 16 yeast chromosomes. Intact chromosomes from the resulting strains were prepared in agarose microbeads and methylated with Hha I (5'-GGCG-3') methyltransferase (M.Hha I) in the presence of lac repressor (Lacl). All Hae II sites (5'-CAGGCG-3') with the exception of the one in lacO, which was protected by Lacl, were modified and thus no longer recognized by Hae II. After inactivation of M.Hha I and Lacl, Hae II was used to completely cleave the chromosomes specifically at the inserted lacO. These experiments demonstrate the feasibility of using the AC approach to efficiently extend the specificity of naturally occurring restriction enzymes and create new tools for the mapping and precise molecular dissection of multigenebass genomes.

INTEREST IN MAPPING, MANIPULATING, AND SEQUENCING OF LARGE GENOMES (1) has inspired a search for methods of precisely and efficiently cleaving chromosomal DNA into a small number of fragments. Presently available restriction enzymes recognize sites 4 to 8 bp in size and thus cut most genomes into a very large number of fragments. Although two general approaches for generating rare cleavage sites in vitro have been known for several years, they either are limited by the size and diversity of the cleavage sites recognized (2) or use cleaving reagents that cut DNA with low efficiency (3). We have introduced an alternative approach, which we call Achilles' heel cleavage (AC), that allows both very rare and highly efficient cleavage of DNA at predetermined locations (4).

The key to the AC procedure is modification of the DNA substrate so as to "erase" all but a small subset of recognition sites for a restriction enzyme with a methyltransferase (MTase) that recognizes the same DNA sequence. This subset of sites is protected from methylation by a DNA-binding molecule, added just before methylation, that forms sequence-specific complexes capable of excluding the MTase. Thus, cleavable restriction sites remain only at those locations where the recognition sites for the given restriction enzyme and for the DNA-binding molecule overlap.

Previous studies have demonstrated the practicality of using the AC approach to efficiently cleave plasmid DNA when the lac repressor (Lacl) (4), the phage λ repressor (5), or a synthetic oligodeoxynucleotide capable of forming a triple-helix structure were used as blocking molecules (5). Lac-mediated AC has also been shown to specifically and efficiently cleave a λ genome (6) that contained the symmetric lac operator (lacO) (7), which is the ideal Lac-binding site and contains the recognition sequence for both Hae II (5'-AGCGCT) and Hha I (5'-GGCG). Conversion of this operator to an AC site creates a restriction recognition site of ~20 bp, which is large enough to be unique even in the human genome.

Results with these relatively small DNA molecules, through encouraging, did not guarantee that AC could be successfully applied to the dissection of the large genomes for which it was designed. Direct testing of an appropriately modified AC protocol on whole chromosomes was necessary to determine whether the greater sequence complexity of these extremely large DNA molecules would result in decreased specificity and efficiency of cleavage, as was the case with another rare-cutting method (8).

In this report, we show the feasibility of using AC for physically mapping and precisely dissecting chromosomes. Model genomes were generated by introducing lacO into the 4.7-Mb circular genome of Escherichia coli (9, 10) and into one of the 16 chromosomes in the 15-Mb genome of Saccharomyces cerevisiae (11). Conditions were...