$Q_I > 2.8 \times 10^5$. In an orbital resonance, the amount of energy available to heat the satellites by tidal dissipation is the difference between the rate at which work is done on the satellites by tidal torques and the rate at which the orbital energies increase [see, for example (*36*)]. By equating this difference to the sum of the minimal melting heating rates of Europa and Ganymede, I estimate that a maximum Q_I of 1.4×10^4 or less is required for both satellites to reach the critical melting eccentricities during passage through the 3:1 resonance, if the Io-Europa 2:1 resonance was not yet established.
35. P. Goldreich and S. Soter, *Icarus* **5**, 375 (1966).
36. S. J. Peale, *ibid.* **74**, 153 (1988).
37. K. Ellsworth and G. Schubert, *Icarus* **54**, 490 (1983).
38. It is possible that the Io-Europa 2:1 resonance remained stable during evolution through the weaker Europa-Ganymede 3:1 resonance, despite the presence of a chaotic zone. I have carried out some numerical integrations of Io and Europa through the 2:1 resonance, and the results indicate that for small initial eccentricities, capture into this resonance is very likely. Tittemore and Wisdom (*22*) showed that Ariel and Umbriel, which exhibit similar behavior at the 2:1 resonance, may remain captured in the resonance even if on a chaotic trajectory. Interference from Io would not likely change the chaotic 3:1 resonance behavior to quasi-periodic behavior.
39. A. J. Friedson and D. J. Stevenson, *Icarus* **56**, 1 (1983).
40. S. Mueller and W. B. McKinnon, *ibid.* **76**, 437 (1988).
41. P. Helfenstein and E. M. Parmentier, *ibid.* **53**, 415 (1983).
42. ———, *ibid.* **61**, 175 (1985).
43. G. D. Crawford and D. J. Stevenson, *ibid.* **73**, 66 (1988).
44. This research was supported in part by the NASA Planetary Geology and Geophysics Program under grant NAGW-1720. This paper is contribution #90-29 of the University of Arizona Theoretical Astrophysics Program.

16 April 1990; accepted 6 August 1990

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

MARK J. GUILTINAN, 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 (CACGTGGC) in the ABRE. A 2-bp mutation in this sequence prevented binding of EmBP-1. The same mutation reduced the ability of the ABRE to confer ABA responsiveness on a viral promoter in a transient 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 mammals, 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 embryo 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 desiccation (*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 embryo development and under stress conditions (*5, 13, 19*) in a manner similar to other Lea genes. When the *Em* 5′ regulatory region was linked to the reporter gene, β-glucuronidase (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 (*10*). Two elements (Em1 and Em2) within this 50-bp ABRE are conserved in other ABA-regulated promoters (*10*), including the rice *Rab* 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 ^{32}P-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* coding 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.

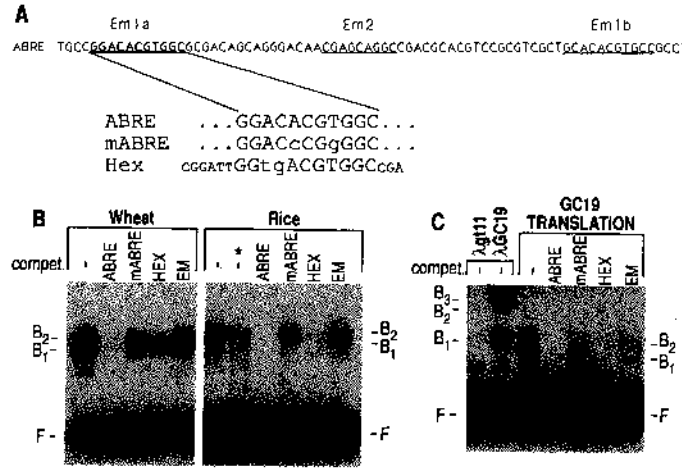

Fig. 1. Nuclear proteins interact specifically with the ABA response element from the wheat *Em* gene. (**A**) Sequences of DNA fragments used in this study: ABRE, 75-bp sequence (1087 to 1161) from the 5′ region of wheat *Em* (*21*), with conserved sequences underlined [Em1a, Em2, and Em1b (*10*)]; mABRE, identical to ABRE except for two single base pair changes (lower case); Hex, 20-bp oligonucleotide derived from the wheat histone gene (−160 to −179, reverse orientation) that contained the HBP-1 binding site (*23*); Em (not shown), a 610-bp DNA fragment (p1015) that contained the wheat *Em* coding sequence (*20*). DNA fragments were cloned and restriction fragments were purified by gel electrophoresis as described (*22*). (**B**) Gel retardation DNA binding assays of native rice and wheat nuclear extracts. Extracts were prepared as described (*22*) from isolated mature wheat embryos imbibed overnight in the presence of 10^{-4} M ABA. Rice extracts were made from suspension cultured cells incubated in the presence or absence (*) of 10^{-4} M ABA. Wheat and rice extracts applied to each lane contained 8.1 and 0.3 μg of protein, respectively. Unlabeled competitor DNA (compet.) was added in 100-fold molar excess 5 min prior to the addition of 0.57 ng of the ^{32}P-labeled ABRE probe (*22*). Binding reactions were incubated at room temperature for 15 min prior to loading. DNA binding reactions and nondenaturing agarose-acrylamide composite gel electrophoresis were as described (*22*). B and F indicate the positions of bound and free DNA probe, respectively. (**C**) Gel retardation DNA binding assay of λGC19 cDNA–encoded proteins. Control (λgt11) and λGC19 lysogen extracts were prepared as described (*25*), and protein (0.42 μg) was added to each binding reaction. In vitro synthesized λGC19 protein (GC19 TRANSLATION) was produced from λGC19 DNA template by a polymerase chain reaction–transcription–translation coupled protocol as described (*22*). From a 50-μl translation reaction, 1 μl was used in each binding reaction. Probe, binding reactions, and electrophoresis conditions were as in (**B**).

Table 1. Transient expression assay of the ABRE and a 2-bp mutant (mABRE). ABRE- and mABRE-containing DNA fragments (Fig. 1A) were tested for their ability to regulate expression in a rice protoplast assay system (*9*). ABRE and mABRE fragments were fused upstream of a CaMV minimal promoter and the GUS reporter gene (*10*, *24*). The plasmid pBM173 was used as a control (none), and activity was measured by production of 4-methylumbelliferyl-β-D-glucuronide (picomole of protein per microgram per hour). Three separate experiments were conducted with similar results. Average activity of triplicate samples (± standard deviation) are shown for each treatment of this single experiment. Averages were used to calculate the ratio.

Element	Activity		Ratio
	−ABA	+ABA	+ABA:−ABA
ABRE	0.76 ± 0.25	8.86 ± 1.86	11.7
mABRE	0.25 ± 0.08	0.52 ± 0.03	2.1
None	0.11 ± 0.16	0.09 ± 0.20	0.8

Extracts 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 *Rab*16A ABRE (*11*). 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 ^{32}P-labeled double-stranded ABRE probe (*25*). From a screen of 120,000 recombinant phages, two overlapping clones were isolated, λGC12 and λGC19, with insert sizes of 528 and 569 bp, respectively. To assess the specificity of the λGC19-encoded DNA binding activity, gel retardation assays were performed. Lysogen extract from the λGC19 clone formed several complexes with the ABRE probe (B_1, B_2, and B_3) that were not formed with proteins from control λgt11 extracts (Fig. 1C). The λGC19 cDNA was transcribed and translated in vitro (*22*), and the resulting polypeptide was tested for binding activity. The in vitro synthesized polypeptides bound specifically to the ABRE probe and formed two major complexes (B_1 and B_2) that had competition profiles that were identical to the native binding proteins in the wheat and rice extracts (Fig. 1C). The λGC19 lysogen extract also showed identical competition profiles (*26*).

To precisely define the nucleotides within the ABRE that were important for binding of the nuclear factors in rice and wheat extracts, and to assess whether the cloned λGC19 protein bound to the same region, methylation interference footprinting was performed (*22*). Methylation of guanine residues only within the region of an 8-bp sequence in Em1a (Fig. 1A) interfered with binding of nuclear proteins from both cultured rice cells and wheat embryos (Fig. 2). The two complexes formed with the wheat extract (B1 and B2, Fig. 1) gave identical interference patterns (*26*). The same patterns were obtained with either the λGC19 lysogen extract or the in vitro synthesized GC19 polypeptide (Fig. 2). No other residues within the ABRE showed detectable methylation interference, including the conserved sequence motifs Em1b and Em2 (Fig. 1A).

Taken together, these data suggest that λGC19 encodes the nuclear protein detected in the native extracts. We therefore refer to this protein as EmBP-1. No other nuclear factors that bound to the ABRE were detected by methylation interference. Our results showed that EmBP-1 interacted only with the Em1a sequence, not with Em2 or Em1b (56 bp 3′ to Em1a), which differs from Em1a by only 2 bp (Fig. 1A).

The EmBP-1–encoding cDNA fragment from λGC19 was ^{32}P-labeled and hybridized to wheat embryo RNA and wheat genomic DNA (*26*). Approximately ten bands of hybridization were observed when hexaploid wheat genomic DNA was digested with Hind III. Hybridization to wheat RNA indicated that EmBP-1 is encoded by a polyadenylated [(poly(A)$^+$] mRNA of 1.8 kilobases (kb).

The nucleotide sequence of the λGC19 cDNA (Fig. 3) revealed a single open reading frame of 569 bp and represented about one-third of the full-length EmBP-1 mRNA. The deduced protein sequence of GC19 contains the two motifs present in the leucine zipper family of transcription factors: a highly basic domain adjacent to a

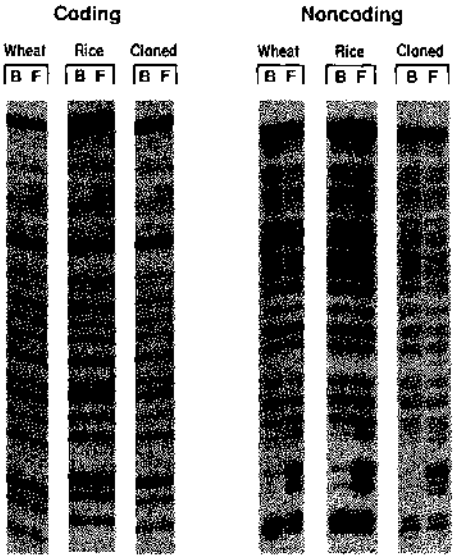

Fig. 2. Methylation interference analysis of native wheat and rice nuclear proteins and of the cloned λGC19-encoded protein (*22*). The ABRE probe (Fig. 1) was 3′ end-labeled by filling in the Eco RI site with Klenow enzyme and [^{32}P]dTTP (*22*). Each strand was labeled separately with the ABRE cloned in opposite orientations within the pUC19 polylinker. Probes were partially methylated and bound to either wheat, rice, λGC19 lysogen extracts (cloned, coding strand), or in vitro synthesized λGC19-encoded proteins (cloned, noncoding strand) as in Fig. 1, with a tenfold increase in all components. After electrophoresis as in Fig. 1, ^{32}P-labeled DNA was purified from free (F) and bound (B) bands as described (*22*). For wheat and rice extracts and the λGC19 translation, complexes B1 and B2 were isolated together. B2 and B3 were isolated together from the λGC19 extract. DNA was then cleaved with piperidine, lyophilized three times, and analyzed on 6% polyacrylamide sequencing gels. Positions of methylated guanines that interfered with binding are indicated with triangles (open for partially interfering bases, closed for more completely interfering bases). Doublet bands in the cloned, noncoding strand lanes were produced as a result of nuclease activity in the translation mix. However, the protected bases were clearly visible and identical to the native extracts.

Table 2. Comparison of promoter elements in plant genes that respond to different signals.

Signal	Gene	Sequence	Reference
ABA	*Em*(1a)	CACGTGGC*	Fig. 1
ABA	*Triticin*	gACGTGGC	(*30*)
ABA	*Rab*/Lea	tACGTGGC*	(*11*)
LIGHT	G-box†	CACGTGGC*	(*33, 34, 35*)
WOUND	*PI-II*	CACGTGGa	(*31*)

*Indicates sequences that were functionally assayed and footprinted. †Genes include: *ribulose bisphosphate carboxylase* (small subunit), *chalcone synthase*, *alcohol dehydrogenase*.

series of five leucine heptad 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 (GCF13) was obtained that appeared to be full length (*26*). The DNA sequence of GCF13 and GC19 share 98% sequence identity and differ by five amino acids (97% identity). GCF13 contains one large open reading frame that encodes 354 amino acids with a predicted molecular size of 36.2 kilodaltons. GCF13 also contains a 3′-untranslated region of 250 bp with a poly(A)$^+$ tract of 15 bp. Also present in the NH_2-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 ABRE from wheat *Em* (Em1a) was protected by EmBP-1. This sequence is conserved in ABA-regulated genes from wheat (*Em, triticin*) (*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 ABRE 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 (*16, 17*). ABA application can independently activate the wound-induced proteinase inhibitor II (PI-II) genes from tomato and potato (*17*). The ABRE consensus was found at −577 in the PI-II promoter (*31*) (Table 2). Similar sequences were found in the wound-inducible genes *wun* 1, *win* 1, and *win* 2 from potato, and *win* 3 from poplar trees (*32*), suggesting that the wound response of these genes may be mediated by ABA.

In addition, the 8-bp ABRE 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 anaerobically induced ADH-1 promoter from maize (*35*). This conserved sequence is important for transcription of some of these genes (*33, 34*), but none 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 ABRE/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 ABRE (Fig. 1). The 8-bp recognition sequence of EmBP-1 is an imperfect palindrome (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 NH_2-terminal to the conserved basic region of EmBP-1 contains a second

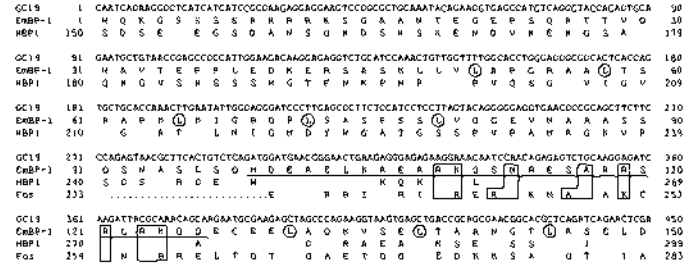

Fig. 3. Nucleotide sequence of the wheat λGC19 cDNA and its deduced EmBP-1 amino acid sequence. The sequence was determined for each strand from subcloned fragments with the dideoxy termination method (Sequenase, US Biochemicals), and sequence compilation and analysis were performed with the University of Wisconsin GCG software package (*43*). Numbering is from the first nucleotide and encoded amino acid, respectively. A comparison is made with a portion of the wheat HBP-1 amino acid sequence (*23*) and the residues in HBP-1 that differ are indicated. The basic and leucine zipper segments of the Fos protein are also indicated (*43*). Amino acids are shown by their single letter abbreviations (*43*). The basic domain is underlined, and periodic leucines and a single methionine are circled. Residues conserved among most leucine zipper proteins are in boxes (*27*).

located in the promoters of the CaMV 35S gene and the *Agrobacterium* nopaline synthase (*nos*) gene (*36*), is similar to the ABRE and competed with ABRE in gel retardation experiments (Table 2 and Fig. 1, B and C). The Hex sites in the CaMV, *nos*, and histone gene promoters can compete for binding to similar factors in nuclear extracts from plants (*37*), but none of these genes are known to be directly regulated by ABA (*9, 38*). A similar sequence (E-box: GGCCACGTGACC) is also found in the major late promoter of adenovirus and in certain mammalian promoters (*39*), and can compete with the G-box element for binding to plant nuclear extracts (*40*).

The core sequence of ABRE is similar to the cyclic adenosine monophosphate (cAMP) response element (CRE) (*2, 8*). It is possible that genes with an ABRE/G-box that are responsive to different signals may include a common intermediate in a second message pathway, like Ca^{2+} or cAMP. These intermediates may be important in hormone and light response pathways in plants (*41*). In animal systems, Ca^{2+} and cAMP act by stimulating protein kinases, which in turn phosphorylate target proteins, including transcription factors (*42*). This family of DNA-binding proteins in plants may mediate responses to wounding, light, and ABA through their activation or modification by different second messenger pathways.

Methylation interference data indicate that the nuclear factors from plants, humans, and yeast that recognize the G-box (*33*), Hex (*23*), and E-box (*39*) core sequence, bind specifically to regions similar to those in the ABRE (Fig. 2). Yeast factor(s) can activate chimeric genes with promoters that contain the G-box sequence from the ribulose bisphosphate carboxylase small subunit promoter (*33*). These results suggest that there is a conserved family of DNA-binding proteins with very similar binding specificities whose members, including EmBP-1, are integrated into different response pathways. However, this highly conserved regulatory element may also be recognized by a different class of transcription factors. For example, a partial cDNA clone from humans, whose protein product (TFEB) binds to the E-box in immunoglobulin gene, contains the helix-loop-helix structural motif (*39*). Interestingly, while the basic regions of TFEB from humans and EmBP-1 from wheat share no amino acid similarity, they bind to very similar DNA sequences and produce nearly identical methylation interference footprints. Comparisons of the structural and functional properties of transcription factors from yeast, plants, and mammals that recognize the conserved ABRE/G-box/E-box sequence may help identify highly conserved recognition domains in these regulatory proteins.

REFERENCES AND NOTES

1. R. M. Evans, *Science* **240**, 889 (1988); M. Beato, *Cell* **56**, 335 (1989).
2. P. J. Deutsch, J. P. Hoeffler, J. L. Jameson, J. F. Jabener, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 7922 (1988).
3. P. J. Davies, Ed., *Plant Hormones and Their Role in Plant Growth and Development* (Nijhoff, Boston, MA, 1987).
4. A. Theologis, T. V. Huynh, R. W. Davis, *J. Mol. Biol.* **183**, 53 (1985); P. M. Chandler, J. A. Zwar, J. V. Jacobsen, T. J. V. Higgins, A. S. Inglis, *Plant Mol. Biol.* **3**, 407 (1984); J. R. Ecker and R. W. Davis, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 5202 (1987).
5. R. S. Quatrano, B. L. Ballo, J. D. Williamson, M. T. Hamblin, M. Mansfield, in *Plant Molecular Biology*, R. Goldberg, Ed. (Liss, New York 1983), pp. 343–353; R. S. Quatrano, *Oxf. Surv. Plant Mol. Cell Biol.* **3**, 467 (1986).
6. A. K. Huttly, R. A. Martienssen, D. C. Baulcombe, *Mol. Gen. Genet.* **214**, 232 (1988); B. A. McClure, G. Hagen, C. S. Brown, M. A. Gee, T. J. Guilfoyle, *Plant Cell* **1**, 229 (1989).
7. K. E. Broglie, P. Biddle, R. Cressman, R. Broglie, *Plant Cell* **1**, 599 (1989).
8. K. Yamaguchi-Shinozaki, J. Mundy, N.-H. Chua, *Plant Mol. Biol.* **14**, 29 (1990).
9. W. R. Marcotte, Jr., C. C. Bayley, R. S. Quatrano, *Nature* **335**, 454 (1988).
10. W. R. Marcotte, Jr., S. H. Russell, R. S. Quatrano, *Plant Cell* **1**, 969 (1989).
11. J. Mundy, K. Yamaguchi-Shinozaki, N.-H. Chua, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 1406 (1990).
12. A. K. Huttly and D. C. Baulcombe, *EMBO J.* **8**, 1907 (1989).
13. R. S. Quatrano, in *Plant Hormones and Their Role in Plant Growth and Development*, P. Davies, Ed. (Nijhoff, Boston, 1987), pp. 494–514; K. Skriver and J. Mundy, *Plant Cell* **2**, 503 (1990).
14. J. C. Baker, C. Steele, L. Dure, *Plant Mol. Biol.* **11**, 277 (1988); L. Dure III *et al.*, *ibid.* **12**, 475 (1989).
15. J. Gomez *et al.*, *Nature* **334**, 262 (1988); J. Mundy and N.-H. Chua, *EMBO J.* **7**, 2279 (1988).
16. I. E. Henson, *Ann. Bot.* **54**, 569 (1984); H. Jones *et al.*, *Planta* **170**, 257 (1987).
17. H. Pena-Cortes *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 9851 (1989).
18. T. J. Close, A. A. Kortt, P. C. Chandler, *Plant Mol. Biol.* **13**, 95 (1989).
19. J. D. Williamson and R. S. Quatrano, *Plant Physiol.* **86**, 208 (1988).
20. J. C. Litts, G. W. Colwell, R. L. Chakerian, R. S. Quatrano, *Nucleic Acids Res.* **15**, 3607 (1987).
21. J. C. Litts, G. W. Colwell, R. L. Chakerian, R. S. Quatrano, *DNA Sequence*, in press. EMBL accession number X52103.
22. The following oligonucleotides were synthesized, treated with T4 polynucleotide kinase, annealed, and extended with Klenow enzyme. They were then ligated into pUC19 that had been digested with Hind III and filled in with Klenow enzyme (pBM209.11). Oligo 1: 5′-TAGGCGGCACGTGTGCAGCGACGCGGACGTGCGTCGGCCTGCTCG-3′; oligo 2: 5′-TGCCGGACACGTGGCGCGACAGCAGGGACAACGAGCAGGCCGACG-3′. The resulting Hind III fragment from pBM209.11 was purified and filled in with Klenow enzyme. It was then ligated to pUC19 that had been digested with Sal I and filled with Klenow enzyme, resulting in recombinant plasmids designated pMG76.11 and pMG76.155. The inserts in these plasmids are in opposite orientation as verified by nucleotide sequencing. mABRE and Hex oligonucleotides were synthesized, treated with T4 polynucleotide kinase, annealed, and ligated to pUC19 that had been digested with Sal I, and filled in with Klenow enzyme. The resulting recombinant plasmids were designated pMG77.43 and pMG82.122, respectively. *Em* was obtained as a 610-bp Pst I DNA fragment from cDNA clone p1015 (*20*). 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, electroelution, and quantitation by spectrophotometry. Nuclear extracts were prepared as described [O. E. Jensen, K. A. Marcker, J. Schell, F. J. deBruijn, *EMBO J.* **7**, 1265 (1988)]. DNA binding analyses were as described: binding conditions [K. Mikami, T. Tabata, T. Kawata, T. Nakayama, M. Iwabuchi, *FEBS Lett.* **223**, 273 (1987)]; band shift gels and purification of complexes for methylation interference [K. Mikami, T. Nakayama, T. Kawata, T. Tabata, M. Iwabuchi, *Plant Cell Physiol.* **30**, 107 (1989)]. For methylation interference analysis, pMG76.11 and pMG76.155 (for coding and noncoding probes, respectively) were cut with Eco RI, labeled with [^{32}P]dTTP (deoxythymine triphosphate) and Klenow enzyme, excised with Pst I, and the resulting 119-bp Eco RI–Pst I fragments were gel-purified for analysis as described [F. M. Ausubel *et al.*, *Current Protocols in Molecular Biology* (Greene, Brooklyn, NY, 1987), chap. 12, pp. 12.3.1–12.3.4]. λGC19-encoding cDNA was translated and the cDNA insert in λGT11 was amplified by polymerase chain reaction with primers 5 and 6 [B. K. Nisikawa, D. M. Fowlkes, B. K. Kay, *Biotechniques* **7**, 730 (1989)] that carry a T7 promoter and an in frame ATG codon 5′ to the Eco RI cloning site, as well as stop codons 3′ to the Eco RI site. After phenol extraction, the DNA was transcribed in vitro (Stratagene cap analog kit), and the resulting RNA was translated in a rabbit reticulocyte lysate (Promega). The resulting polypeptides (apparent molecular size, 25 kD) were used for binding and footprinting reactions without further purification.
23. T. Tabata *et al.*, *Science* **245**, 965 (1989).
24. ABRE-containing expression plasmids were constructed as follows: The ABRE was isolated from pBM209.11 (*22*) and the mABRE (Fig. 1A) was isolated from a plasmid (pBM234) prepared as for pBM209.11. The 80-bp Hind III fragments differed in only 2 bp and were inserted into the Hind III site of pBM173 (*10*). The orientation of ABRE and mABRE in the resulting plasmids (pBM226 and pBM233, for wild-type and mutant sequences, respectively) was 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 (*9*).
25. A cDNA library was constructed with 10 μg of immature (stage III) ABA-treated wheat embryo poly(A)$^+$ RNA, random primers, and The Librarian Kit (Invitrogen) as described [U. Gubler and B. J. Hofman, *Gene* **25**, 283 (1983)]. Recombinants (2.64×10^6) were amplified, and 120,000 plaques were screened as described [H. Singh, R. G. Clerc, J. H. LeBowitz, *Biotechniques* **7**, 252 (1989)] with minor modifications. Oligonucleotide probes were labeled by annealing the partially complementary ABRE oligonucleotides 1 and 2 (*22*) and extending with Klenow enzyme in the presence of [^{32}P]dCTP (deoxycytidine 5′-triphosphate). The probe was incubated with the filters overnight at 4°C.
26. M. Guiltinan, unpublished data. 6CF13 cDNA sequence for EMBP-1, GenBank accession number M38418.
27. C. R. Vinson, P. B. Sigler, S. L. McKnight, *Science* **246**, 911 (1989).
28. R. J. Schmidt, F. A. Burr, M. J. Aukerman, B. Burr, *Proc. Natl. Acad. Sci. USA* **87**, 46 (1990).
29. P. J. Mitchell and R. Tijan, *Science* **245**, 371 (1989).
30. R. Quatrano, unpublished data.
31. M. Keil, J. Sanchez-Serrano, J. Schell, L. Willmitzer, *Nucleic Acids Res.* **14**, 5641 (1986).
32. B. Siebertz, J. Logemann, L. Willmitzer, J. Schell, *Plant Cell* **1**, 961 (1989); A. Stanford, M. Bevan, D. Northcote, *Mol. Gen. Genet.* **215**, 200 (1989); H. D. Bradshaw, Jr., J. B. Hollick, T. J. Parsons, H. R. G. Clarke, M. P. Gordon, *Plant Mol. Biol.* **14**, 51 (1989).
33. R. G. K. Donald, U. Schindler, A. Batschauer, A. R. Cashmore, *EMBO J.* **9**, 1727 (1990).
34. P. Suhulze-Lefert, J. L. Dangl, M. Beker-Andre, K. Hahlbrock, W. Schulz, *ibid.* **8**, 651 (1989).
35. A. J. DeLisle and R. J. Ferl, *Plant Cell* **2**, 547 (1990).
36. F. Kataagiri, E. Lam, N.-H. Chua, *Nature* **340**, 727 (1989).

37. K. Mikami, A. Sakamoto, H. Takase, T. Tabata, M. Iwabuchi, *Nucleic Acids Res.* **17**, 9707 (1989).
38. G. An *et al.*, *Plant Cell* **2**, 225 (1990).
39. L. A. Chodish, S. Buratowski, P. A. Sharp, *Mol. Cell. Biol.* **9**, 820 (1989); C. S. Carr and P. A. Sharp, *ibid.* **10**, 4384 (1990).
40. D. Staiger, H. Kaulen, J. Schell, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6930 (1989).
41. C. Ettlinger and L. Lehle, *Nature* **331**, 176 (1988); S. Gallagher, T. W. Short, P. M. Ray, L. H. Pratt, W. R. Briggs, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8003 (1988); M. R. McAinsh, C. Brownlee, A. M. Hetherington, *Nature* **343**, 186 (1990).
42. G. Gonzalez and M. R. Montminy, *Cell* **59**, 675 (1989).
43. J. Devereux, P. Haeberli, O. Smithies, *Nucleic Acids Res.* **12**, 387 (1984). Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. The Fos sequence is from T. Curran, G. Peters, C. Van Beveren, N. M. Teich, I. M. Verma, *J. Virol.* **44**, 674 (1982).
44. We thank L. Brian, R. Brosh and R. Kempainen for their excellent technical assistance and A. Baldwin and T. Thomas for helpful discussions. Thanks are also extended to P. Bedinger for use of her lab facilities during parts of this work, B. Kay for advice and providing the polymerase chain reaction primers, S. Whitfield for photographic assistance, and N. Welsch for preparation of the manuscript. Supported by the USDA Competitive Grants Program (89-37262-4456) (to R.S.Q.).

17 April 1990; accepted 21 August 1990

Cleaving Yeast and *Escherichia coli* Genomes at a Single Site

MICHAEL KOOB 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 (*lac*O_s) 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′-GCGC) methyltransferase (M·Hha I) in the presence of *lac* repressor (LacI). All Hae II sites ($5'$-$^{A}_{C}GCGC^{T}_{C}$) with the exception of the one in *lac*O_s, which was protected by LacI, were modified and thus no longer recognized by Hae II. After inactivation of M·Hha I and LacI, Hae II was used to completely cleave the chromosomes specifically at the inserted *lac*O_s. 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 multimegabase 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 (LacI) (*4*), the phage λ repressor (*4*), or a synthetic oligodeoxynucleotide capable of forming a triple-helix structure were used as blocking molecules (*5*). LacI-mediated AC has also been shown to specifically and efficiently cleave a λ genome (*6*) that contained the symmetric *lac* operator (*lac*O_s) (*7*), which is the ideal LacI-binding site and contains the recognition sequence for both Hae II (5′-AGCGCT) and Hha I (5′-GCGC). 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.

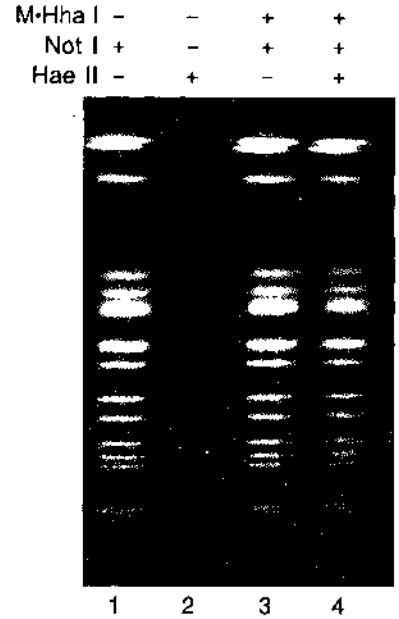

Fig. 1. Complete methylation of the *E. coli* genome. Genomic DNA from *E. coli* strain BNN103 (*9*) was prepared in agarose microbeads (*12*). The microbeads were equilibrated with methylation buffer [50 mM tris-HCl (pH 7.5), 10 mM EDTA, 10 mM dithiothreitol (DTT), 80 μM *S*-adenosyl-methionine, and bovine serum albumin (BSA) (100 μg/ml)], 12.5 U of M·Hha I (New England Biolabs) were added to the reactions indicated, and all samples were incubated for 1 hour at 37°C. The MTase and contaminating nucleases were then inactivated (20 μl of 500 mM EDTA and 1% *N*-lauroylsarcosine, 30 min at 52°C), the microbeads equilibrated with Not I buffer [150 mM NaCl, 10 mM tris-HCl (pH 8.0), 10 mM $MgCl_2$, BSA (100 μg/ml), and 0.01% Triton X-100], and 5 U of the indicated restriction enzyme were added. After incubation at 37°C for 1 hour, the samples were deproteinized as before (15 min at 52°C) and analyzed by PFGE [1% high-strength agarose (Bio-Rad), 150 V, and 25-s switch time for 19 hours at 14°C on a CHEF-DR II system (Bio-Rad)]. Treatments are specified above each lane.

Results with these relatively small DNA molecules, though 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 *lac*O_s 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

McArdle Laboratory for Cancer Research, University of Wisconsin, Madison, WI 53706.