The Plant Cell, Vol. 9, 611–625, April 1997

Evidence for Direct Activation of an Anthocyanin Promoter by the Maize C1 Protein and Comparison of DNA Binding by Related Myb Domain Proteins

Manuel B. Sainz,[a] Erich Grotewold,[b] and Vicki L. Chandler[a,1]

[a] Institute of Molecular Biology, University of Oregon, Eugene, Oregon 97403
[b] Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724

The enzyme-encoding genes of two classes of maize flavonoid pigments, anthocyanins and phlobaphenes, are differentially regulated by distinct transcription factors. Anthocyanin biosynthetic gene activation requires the Myb domain C1 protein and the basic helix-loop-helix B or R proteins. In the phlobaphene pathway, a subset of C1-regulated genes, including *a1*, are activated by the Myb domain P protein independently of B/R. We show sequence-specific binding to the *a1* promoter by C1 in the absence of B. Activation is decreased by mutations in the C1 DNA binding domain or in *a1* sequences bound by C1, providing direct evidence for activation of the anthocyanin biosynthetic genes by C1. The two C1 binding sites in the *a1* promoter are also bound by P. One site is bound with higher affinity by P relative to C1, whereas the other site is bound with similar lower affinity by both proteins. Interestingly, either site is sufficient for C1 plus B/R or P activation in vivo, demonstrating that differences in DNA binding affinities between P and C1 are insufficient to explain the differential requirement for B. Results of DNA binding site–selection experiments suggest that C1 has a broader DNA binding specificity than does P, which may help C1 to activate a more diverse set of promoters.

INTRODUCTION

Regulation of flavonoid pigment biosynthesis has been an excellent system for the study of combinatorial regulation of an entire set of biosynthetic genes (reviewed in Dooner et al., 1991; van der Meer et al., 1993). Two classes of maize flavonoid pigments are the red and purple anthocyanins and the red phlobaphenes. The anthocyanin and phlobaphene biosynthetic pathways share at least three enzymatic steps and then diverge. Maize genes encoding most of the flavonoid pigment biosynthetic enzymes as well as several regulatory transcription factors have been identified and cloned. Numerous experiments have demonstrated that flavonoid biosynthetic gene activation requires developmental and tissue-specific expression of the transcription factors regulating the anthocyanin and phlobaphene biosynthetic pathways.

Two classes of transcription factors are required for the regulation of anthocyanin biosynthesis in maize. A functional allele of *c1* or *pl* together with a functional allele of *b* or *r* are required to coordinately induce mRNA levels of the anthocyanin biosynthetic genes (Cone et al., 1986, 1993; Chandler et al., 1989; Ludwig et al., 1989). The *c1* (Paz-Ares et al., 1987) and *pl* (Cone et al., 1993) loci encode functionally equivalent proteins that have 80% identity but are expressed in different tissues. These proteins share an N-terminal Myb motif, originally identified as the DNA binding domain of the v-Myb oncoprotein (Biedenkapp et al., 1988) and consisting of two (or three) amino acid repeats. Each Myb repeat forms a helix-helix-turn-helix structure (Ogata et al., 1994). The C terminus of C1 functions as an independent transcriptional activation domain in maize and in yeast when fused to the DNA binding domain of the yeast GAL4 protein (Goff et al., 1991; Sainz et al., 1997). The *b* and *r* genes encode functionally duplicate proteins that share 78% identity and are expressed in diverse tissues (Dellaporta et al., 1988; Chandler et al., 1989; Ludwig et al., 1989; Radicella et al., 1991). B and R have a basic helix-loop-helix (b/HLH) motif characteristic of a large family of transcriptional activators (reviewed in Ludwig and Wessler, 1990). In various proteins of this class, the basic region functions in DNA binding, whereas the HLH domain mediates homodimerization or heterodimerization (Weintraub et al., 1991).

Previous studies have demonstrated that the Myb domain and b/HLH factors induce transcription of the anthocyanin biosynthetic genes. In the presence of functional *pl* and *b* alleles, both steady state mRNA levels and transcription rates of the anthocyanin genes increase (Cone et al., 1993; Patterson et al., 1993). *cis*-Acting sequences that are necessary and sufficient for transcriptional activation by C1 plus B/R map to upstream promoter regions of the biosynthetic genes, which are 5′ of the start of transcription (Roth et al., 1991; Bodeau and Walbot, 1992; Grotewold et al., 1994; Tuerck

[1] To whom correspondence should be addressed. E-mail chandler@molbio.uoregon.edu; fax 541-346-5011. Address after May 1: Department of Plant Sciences, 303 Forbes Hall, University of Arizona, Tucson, AZ 85721. E-mail chandler@ag.arizona.edu; fax 520-621-7186.

and Fromm, 1994). Nevertheless, it remained an open question whether activation is a consequence of the Myb domain and b/HLH factors directly interacting with anthocyanin promoter sequences or whether they act indirectly by activating a gene that encodes the direct activator of the anthocyanin genes.

A model in which the Myb domain and b/HLH factors physically interact to activate transcription has emerged from several studies. There is no evidence that either factor regulates expression of the other. Synthesis of *pl* mRNA is not dependent on a functional *b* or *r* allele (Cone and Burr, 1989), and *b* mRNA synthesis is not dependent on a functional *pl* allele (V. Chandler, unpublished results). In addition, constitutive expression of either C1 or B does not circumvent a requirement for the other activator (Goff et al., 1990). Furthermore, C1 and B interact in two-hybrid protein–protein interaction assays in yeast and in maize (Goff et al., 1992), consistent with a direct physical association between them. An extensive analysis of the B protein did not identify a transcriptional activation domain in B (Goff et al., 1992), suggesting that B mediates activation via protein–protein interaction with C1.

Unlike the anthocyanin pathway, in which a Myb domain and a b/HLH protein are both required for activation, only the Myb domain P protein has been found to regulate phlobaphene biosynthesis. The N-terminal 114–amino acid Myb domains of C1 and P are 69% identical, whereas the C termini of the two proteins share no similarity. Independently of B or R, P activates pigment biosynthesis (Styles and Ceska, 1977; V. Chandler, unpublished results) and induces the mRNA levels of three genes shared between the phlobaphene and anthocyanin biosynthetic pathways (Grotewold et al., 1991, 1994). In contrast, the C1 protein is absolutely dependent on B or R for transcriptional activation of the same biosynthetic genes.

To determine whether C1 is a direct activator of anthocyanin biosynthetic genes, we asked whether C1 exhibited sequence-specific DNA binding to functionally important sequences of the *a1* promoter and whether B was required for this activity. The *a1* gene encodes dihydroflavonol reductase (Schwarz-Sommer et al., 1987) and is independently regulated by C1 plus B/R and by P (Klein et al., 1989; Grotewold et al., 1994). To gain insight into how the related C1 and P proteins activate some of the same targets yet have different requirements for B or R, we analyzed the ability of these proteins and mutant derivatives to bind to wild-type and mutant *a1* promoter sequences and to activate transcription in maize cells.

RESULTS

C1 Is a Sequence-Specific DNA Binding Protein

The available structures of animal Myb domains demonstrate that each Myb repeat consists of a helix-helix-turn-helix motif and that the third helix in each repeat makes sequence-specific DNA contacts (Ogata et al., 1994). With the exception of a single leucine/isoleucine difference in the first predicted DNA recognition helix, C1 and P are identical to each other within these regions. Thus, a reasonable hypothesis was that C1 would bind to an *a1* promoter sequence to which P binds (Grotewold et al., 1994).

To test whether C1 could bind to the *a1* sequence bound by P, extracts from *Escherichia coli* expressing a protein with the first 253 amino acids of C1 fused to an epitope tag and six histidine residues (C1/6×HIS) were used in a gel retardation assay with a DNA fragment that spans from −76 to −47 in the *a1* promoter. As shown in Figure 1A, a DNA binding activity was observed in C1/6×HIS-containing extracts (lane 2) but not in the negative control (lane 7). Preincubation of the binding reactions with the affinity-purified C1 antibody (lane 5) but not with preimmune serum (lane 6) abolished binding, indicating that the binding activity corresponded to C1/6×HIS. In addition, preincubation with an unrelated purified antibody had no effect on binding, and under certain conditions, a C1 antibody–dependent supershift was observed (data not shown). DNA binding by C1/6×HIS was effectively competed by a large molar excess of unlabeled −76 to −47 *a1* sequence (lane 3) but not by a mutant derivative of the sequence with two base pair changes in the P binding site (lane 4). The substitutions in the mutant derivative reduced transcriptional activation by P and C1 plus R in vivo (Grotewold et al., 1994). Binding by C1/6×HIS to the radiolabeled −76 to −47 *a1* DNA sequence of the mutant was not observed under these conditions (data not shown). The six-histidine residue tag did not influence C1 DNA binding, because when these experiments were repeated with bacterially expressed full-length C1 protein without the six-histidine, similar results were obtained (data not shown). These results demonstrate that C1 exhibits sequence-specific DNA binding to the *a1* promoter and suggest that C1 binds to the same site bound by P (Grotewold et al., 1994). Binding by C1 is independent of B, and the addition of B protein preparations had no effect on C1 binding to the −76 to −47 *a1* sequence (data not shown).

Given that the Myb domain is sufficient for DNA binding by animal Myb proteins (Howe et al., 1990) and P (Williams and Grotewold, 1997), we tested whether the Myb domain of C1 was sufficient for binding to the −76 to −47 region of the *a1* promoter. Extracts from bacteria expressing a protein with a 10-histidine residue tag fused to the N terminus of the C1 Myb domain (amino acids 1 to 119; C1myb) exhibited binding to the −76 to −47 region of the *a1* promoter (lanes 10 to 12), and this activity was absent in extracts from uninduced cells (lane 9), demonstrating that the Myb domain of C1 is sufficient for DNA binding activity.

DNase I footprinting and methylation interference analyses were performed to determine where C1 binds within the *a1* promoter. Figure 1B depicts a DNase I footprint of the −138 to −1 region of the *a1* promoter, using the purified C1/6×HIS protein, which showed protection of bases within the −65 to −54 region of the *a1* promoter. In addition, meth-

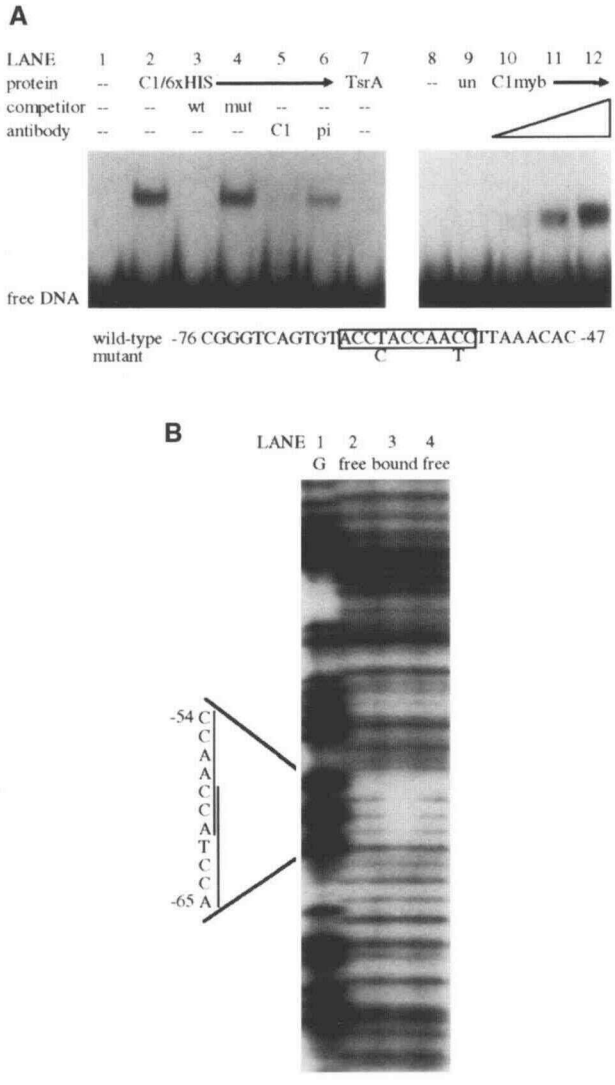

Figure 1. Sequence-Specific DNA Binding by C1/6×HIS to a Site in the −76 to −47 *a1* Promoter.

(A) Sequence-specific binding by C1/6×HIS to the −76 to −47 region of the *a1* promoter. A gel retardation assay was performed with extracts from bacteria expressing C1/6×HIS (lanes 2 to 6) and a radiolabeled −76 to −47 region of the *a1* sequence (shown as wild type below). Dashes indicate the absence of a given component. Lane 1 contains the free DNA; lane 7 shows the DNA binding activity in extracts from bacteria expressing TsrA (Ames and Parkinson, 1994), which serves as a control for bacterial DNA binding proteins induced by protein overexpression. Competitions were performed using the wild-type sequence (wt; lane 3) or a mutant derivative (mut; lane 4), with two base pair changes (indicated below the wild-type sequence) in the boxed P binding site (Grotewold et al., 1994). Reactions were preincubated with the C1 antibody (C1; lane 5) or preimmune serum (pi; lane 6). In a different gel (lanes 8 to 12), extracts from bacteria expressing the C1 Myb domain (C1myb) for 1 hr (lane 10), 2 hr (lane 11), and 3 hr (lane 12) were used in gel retardation assays with the radiolabeled −76 to −47 *a1* promoter sequence. Lane 8, free DNA; lane 9, extract from uninduced (un) cells.

(B) DNase I footprint of C1/6×HIS bound to a site at −65 to −54 in the *a1* promoter. DNase I footprinting was performed using purified C1/6×HIS and a radiolabeled −138 to −1 region of the *a1* promoter. Lane 1 contains the Maxam and Gilbert G sequencing reaction (G); lanes 2 and 4 contain free DNA (free); lane 3 contains C1/6×HIS bound DNA. The sequence shown is to the strand opposite the one labeled. The lines at left denote the two overlapping P binding sites previously reported by Grotewold et al. (1994).

ylation interference assays indicated that C1/6×HIS was prevented from binding by methylation of the six guanine bases in this region (cytosines on the strand shown in Figure 1B; data not shown). Thus, like P (Grotewold et al., 1994), C1 binds to the −65 to −54 site in the *a1* promoter.

C1 and P Bind to Two Functionally Important Sites in the *a1* Promoter

Sequence-specific DNA binding by C1 and P to the −76 to −47 region of the *a1* promoter provides an explanation for the ability of this region of *a1* to confer C1 plus R and P inducibility to a heterologous minimal promoter (Grotewold et al., 1994). However, a region of the *a1* promoter from −140 to −79 can also confer inducibility by C1 plus B and by P (Tuerck and Fromm, 1994). Activation of the −140 to −79 region of the *a1* promoter could be mediated by C1 and P directly binding to this sequence or through the binding of other proteins to this region that recruit C1 or P to the DNA.

To explore further their DNA binding characteristics, C1 and P proteins were purified from crude bacterial extracts. Unfortunately, full-length P is not expressed as well as C1 in *E. coli* and is highly susceptible to proteolysis (data not shown). We overcame this problem by using a fusion protein with the Myb DNA binding domain of P fused to the C terminus of C1/6×HIS (Pmyb/C1/6×HIS), which was expressed to the same extent as C1/6×HIS in bacteria (data not shown). Several lines of evidence suggest that Pmyb/C1 functions similarly to P. Like P, Pmyb/C1 activates *a1* independently of R when expressed in maize cells (Grotewold et al., 1994). Furthermore, the P Myb domain is sufficient for DNA binding (Williams and Grotewold, 1997). We were able to purify C1/6×HIS and Pmyb/C1/6×HIS from bacterial extracts to the same extent (~80% pure; data not shown), facilitating comparative analyses of the DNA binding activities mediated by the C1 and P Myb domains in the context of the C1/6×HIS C terminus.

To test the hypothesis that C1 and P directly activate the −140 to −79 region of the *a1* promoter, we examined whether C1/6×HIS could bind to a distal (−133 to −83) *a1* sequence in gel retardation assays. As shown in Figure 2A, C1/6×HIS bound to the distal *a1* sequence (lane 2), and C1/6×HIS binding was diminished by the addition of the C1 antibody (lane 4) but not by preimmune serum (lane 3). Specific binding to the distal *a1* sequence was competed by an excess of unlabeled distal DNA (lane 6) and even more effectively by

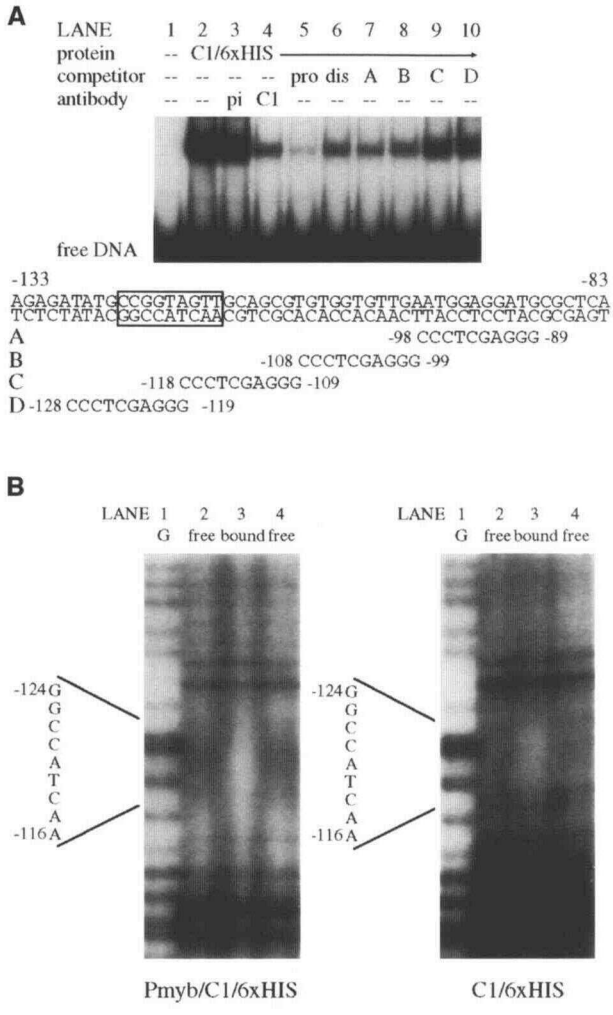

Figure 2. C1/6×HIS and Pmyb/C1/6×HIS DNA Binding to a Site in the −133 to −83 Region of the *a1* Promoter.

(A) Sequence-specific binding by C1/6×HIS to the −133 to −83 region of the *a1* promoter. Gel retardation assays were done with equal amounts of the purified C1/6×HIS protein (lanes 2 to 10) and the radiolabeled wild-type −133 to −83 *a1* sequence. Dashes indicate the absence of a given component. Lane 1 contains the free DNA. Reactions were preincubated with preimmune serum (pi; lane 3), C1 antibody (C1; lane 4), or unlabeled competitor DNA (lanes 5 to 10) before the addition of radiolabeled DNA. The competitor DNA used in lanes 5 and 6 are the proximal (pro; −76 to −47) and distal (dis; −133 to −83) *a1* sequences, respectively. The sequences changed in the mutant *a1* promoter derivatives used as competitor DNA (A to D; lanes 7 to 10) and corresponding to the top strand of the wild-type sequence are shown, along with their coordinates. The boxed region is bound by C1/6×HIS and Pmyb/C1/6×HIS in DNase I footprinting assays **(B)**.

(B) DNase I footprints of Pmyb/C1/6×HIS and C1/6×HIS bound to a site at −116 to −124 in the distal *a1* promoter. DNase I footprinting was performed using purified Pmyb/C1/6×HIS or C1/6×HIS and a radiolabeled −219 to +7 sequence of the *a1* promoter with a mutated −65 to −54 binding site. Lanes 1 contain the Maxam and Gilbert G sequencing reaction (G); lanes 2 and 4 contain free DNA (free); lanes 3 contain Pmyb/C1/6×HIS or C1/6×HIS bound DNA (bound) as indicated. The sequence shown is to the strand opposite the one labeled.

equivalent amounts of the unlabeled proximal (−76 to −47) *a1* DNA (lane 5).

To identify where C1/6×HIS was binding within the distal *a1* promoter, DNA sequences with specific mutations were used to compete with radiolabeled wild-type distal DNA for binding by C1/6×HIS (Figure 2A). These mutations had been shown previously to decrease activation in vivo (Tuerck and Fromm, 1994). DNA fragments with mutations A and B were as effective competitors as the wild-type −133 to −83 region of *a1* DNA (Figure 2A, lanes 7 and 8), suggesting that C1/6×HIS bound to fragments with these mutated sequences with the same affinity with which it bound to the wild type. In contrast, DNA fragments with mutations C or D were less effective as competitors compared with the wild-type distal fragment (Figure 2A, lanes 9 and 10), suggesting that C1/6×HIS does not bind well to the mutated sequences in the C and D variants.

Similar results were obtained in experiments involving direct binding of C1/6×HIS to radiolabeled wild-type and mutant DNA fragments. C1/6×HIS bound ~30% of the wild-type fragment and ~30% of the fragments with the A and B mutations; in contrast, only 5% of DNA sequences bearing the C and D mutations were bound by an equivalent amount of C1/6×HIS (data not shown). These data are consistent with the competition results, and together they suggest that C1/6×HIS binds to the *a1* promoter in the −128 to −109 region spanned by the C and D mutations. The binding and competition experiments were repeated with Pmyb/C1/6×HIS, with results similar to those obtained with C1/6×HIS (data not shown), indicating that the P Myb domain also mediates binding to the distal *a1* sequence.

DNase I footprinting of the wild-type *a1* promoter failed to reveal C1/6×HIS (data not shown) or P (Grotewold et al., 1994) binding to distal *a1* sequences. This result, combined with the more effective competition exhibited by the proximal *a1* fragment relative to the distal sequence (Figure 2A, lanes 5 and 6), suggests that C1 and P bind with higher affinity to the overlapping $ACC^{T}/_{A}ACC$ sequence at −65 to −54 in the *a1* promoter. To identify more precisely the C1 and P binding sites within the distal *a1* sequence, a −219 to +7 *a1* sequence with a mutated −65 to −54 proximal binding site was used in DNase I footprinting experiments. As seen in Figure 2B, when the proximal site is mutated, C1/6×HIS and Pmyb/C1/6×HIS exhibit binding to a site at −116 to −124 (AACTACCGG), in the opposite orientation relative to the −65 to −54 binding site. This sequence is within the region covered by the C and D mutations in the distal *a1* sequence, as predicted from the results of the gel retardation experiments.

Our identification of C1 and P binding sites in each of the regions sufficient for activation by C1 plus B and by P (Grotewold et al., 1994; Tuerck and Fromm, 1994) suggests that C1 and P bind to both sites in *a1* independently and that a single site suffices to assemble a functional transcription complex. Previous studies have shown that mutations within either binding site in the context of the full-length *a1* promoter had detectable but minor effects (Grotewold et al., 1994; Tuerck and Fromm, 1994). We now hypothesize that this was because C1 or P could still bind to the other site and activate transcription. A prediction of this hypothesis is that mutating both binding sites should have a more severe effect on activation by C1 plus B and by P.

To test this hypothesis, we assayed C1 plus B and P activation of the wild-type *a1* promoter, promoter derivatives with mutations in each of the C1 and P binding sites, or a derivative with both sites mutated. Activation of wild-type and mutant *a1* promoters was tested in cultured maize cells transiently transformed by microprojectile bombardment; the constructs used are depicted in Figure 3A. As seen in Figure 3B, changes in either of the C1 and P binding sites moderately reduced C1 plus B and P activation, in agreement with previous studies (Grotewold et al., 1994; Tuerck and Fromm, 1994). Mutation of both binding sites had a more severe effect on activation by C1 plus B and by P (Figure 3B). Mutation of the proximal or distal binding sites decreased activation by P similarly and was 18 and 16% of the levels observed with the wild-type *a1* promoter, respectively. When both binding sites were mutated, activation by P was only threefold over background levels, suggesting that P binding sites mediate the major portion of *a1* activation by P in vivo. The addition of B did not alter activation by P (data not shown). For activation by C1 plus B, mutations of the proximal and distal C1 binding sites had similarly mild effects and were 35 and 53% of wild-type *a1* promoter activity, respectively. C1 plus B activation of the *a1* promoter double mutant was further decreased to 17% of wild-type *a1* levels, yet it remained 12-fold over background. For activation by both C1 plus B and by P, the phenotype of the double-site mutant was additive relative to the single-site mutations, suggesting that C1 and P bind independently to the two sites. A summary of our results, combined with previously published data, is presented in Figure 3C.

Although the identified C1 binding sites are important for activation, other sequences within the *a1* promoter may contribute as well. When mutated, the −98 to −89 region caused a major decrease in C1 plus B activation of *a1* (Figure 3C, mutant A; Tuerck and Fromm, 1994). Nevertheless, C1 binding to the −133 to −83 *a1* promoter fragment in the absence of the distal binding site was extremely weak, and mutations of the −98 to −89 region did not affect C1 binding to the distal *a1* site (Figure 2A). These results suggest that other proteins important for activation by C1 plus B may be binding in the −98 to −89 region. If true, this may account for the residual level of C1 plus B activation we observed when both C1 binding sites in *a1* were mutated. A logical candidate for a protein binding to the −98 to −89 region is B or R. However, when using an in vitro–translated B protein or B protein preparations from *E. coli* or Sf9 insect cells, we did not observe DNA binding to any functionally important anthocyanin promoter sequence (data not shown). Potentially, B may require an HLH dimerization partner to bind to DNA.

C1 DNA Binding Activity Is Important for Activation of the *a1* Promoter

If C1 DNA binding activity is important for transcriptional activation of the anthocyanin promoters, then a C1 mutant that fails to bind to DNA in vitro should be unable to activate in vivo. The C1 Myb domain is sufficient for DNA binding (Figure 1A) and for interacting with B (Goff et al., 1992), providing two independent functional assays. This enables the identification of C1 mutants specifically defective in DNA binding or the interaction with B. Mutations that result in misfolded proteins would be predicted to affect both activities.

A candidate C1 DNA binding mutant was noted in previous studies. A C1 mutant in which the aspartate (D) at position 101 was changed to glutamate (E) was isolated from a naturally occurring dominant inhibitor allele of *c1* called *C1-I* (Paz-Ares et al., 1990; Goff et al., 1991). D-101 is conserved in mouse Myb. Based on the available structure of the mouse Myb repeats bound to DNA (Ogata et al., 1994), D-101 is at the beginning of the putative DNA recognition helix in the second Myb repeat of C1. Previous experiments had shown that the C1:D101E mutant was unable to activate the *bronze1* anthocyanin promoter in maize cells (Goff et al., 1991). Nevertheless, C1:D101E was able to interact with B as well as wild-type C1 in maize two-hybrid assays, leading to the prediction that C1:D101E was a DNA binding mutant (Goff et al., 1992).

To test this prediction, wild-type and mutant C1 proteins were assayed for in vitro DNA binding to the proximal *a1* sequence. As seen in Figure 4A, DNA binding by C1:D101E/6×HIS (lane 5) was extremely weak, with a faint band apparent only in very long exposures (data not shown). In contrast, a 1:10 dilution of wild-type C1/6×HIS (lane 4) generated a visible band in this experiment. Thus, C1:D101E/6×HIS is defective in DNA binding. A DNA binding defect is also caused by the analogous D117E mutation in MYB.Ph3, a plant Myb domain protein from petunia (Solano et al., 1995).

Even though C1 and C1:D101E interacted similarly with B in maize two-hybrid assays (Goff et al., 1992), we are unable to quantify expression of the two proteins in this system. Because C1 and C1:D101E are expressed similarly in yeast (data not shown), we used yeast two-hybrid assays to confirm that wild-type and mutant C1 proteins interact similarly with B. The wild-type C1/GAL4 activation domain and the B/GAL4 DNA binding domain fusion proteins generated a robust level of β-galactosidase activity when coexpressed in yeast cells bearing a *GAL1* promoter–*lacZ* reporter gene. As

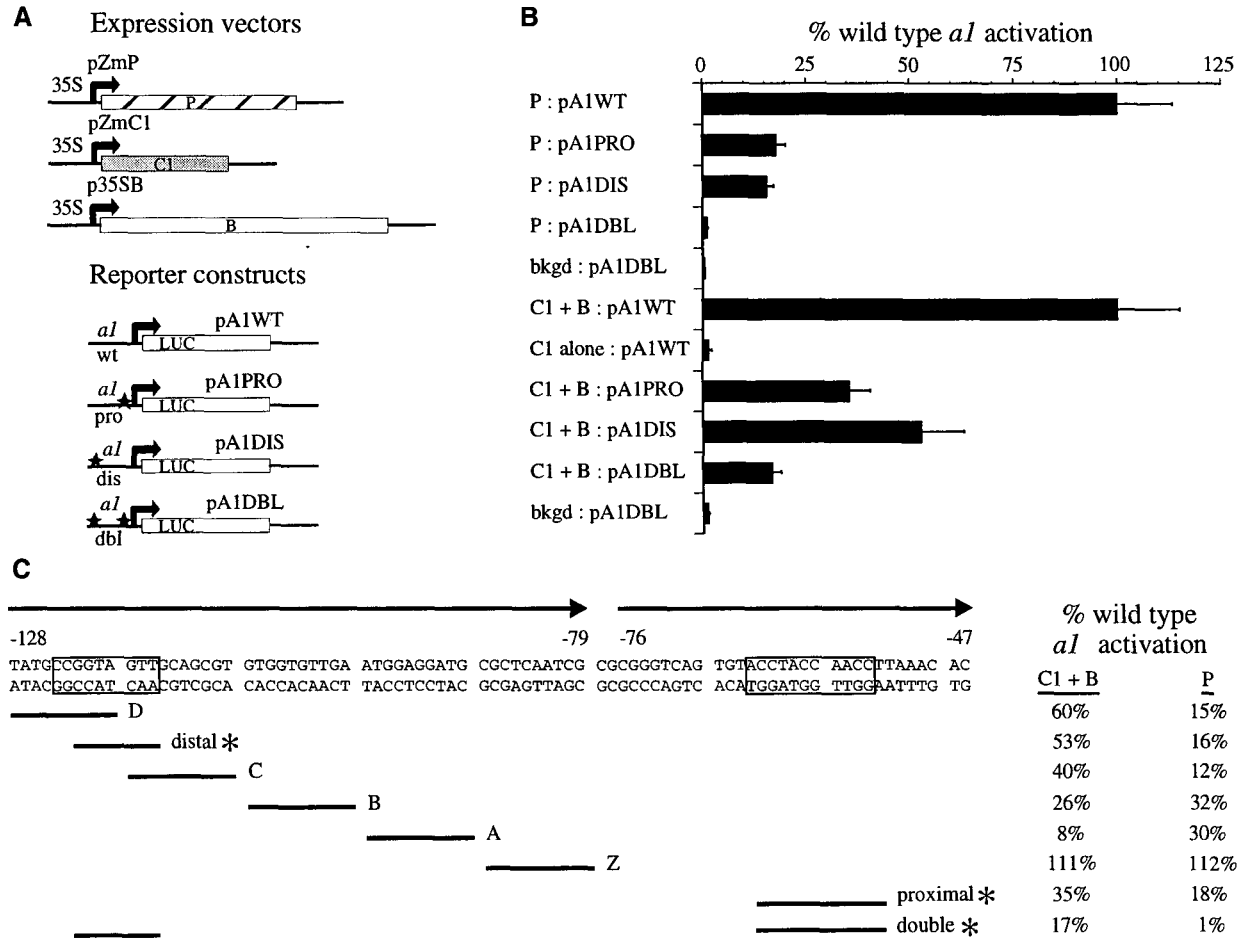

Figure 3. Activation of Wild-Type and Mutant *a1* Promoters by P and by C1 plus B.

(A) Constructs used in maize transient transformation assays. Plasmids designed to express the activator proteins (pZmP, pZmC1, and p35SB) are shown. Reporter constructs with the wild-type (wt) −219 to +7 region of the *a1* promoter (pA1WT) or mutant derivatives fused to a firefly luciferase (LUC) gene are shown. The proximal mutant (pA1PRO) has base pair changes in the C1 and P binding site at −65 to −54 (pro). The distal mutant (pA1DIS) changes the C1 and Pmyb/C1 binding site at −116 to −124 (dis), and the double (dbl) mutant *a1* promoter reporter (pA1DBL) has both the proximal and distal binding site mutations.

(B) Activation by P and by C1 plus B of wild-type and mutant *a1* promoters. Cultured maize cells were transiently transformed with expression vectors, reporter constructs, and a constitutively expressed β-glucuronidase reporter gene, which is included as a transformation and extraction control. The ratio of luciferase to β-glucuronidase activity in extracts from the transiently transformed cells was determined and then expressed as the fold induction observed in the presence of the activators relative to the expression observed in their absence. Fold induction for the mutant promoter constructs was normalized to the fold induction by P or by C1 plus B of pA1WT, which was set at 100%. The observed induction of pA1WT is robust: an average of 272-fold for P and 90-fold for C1 plus B. The amount of background (bkgd) activation was similarly low for all three promoter constructs. Each histogram represents the average of 10 to 12 samples from two separate experiments (except C1 alone, $n = 6$ in one experiment; C1 plus B, $n \approx 16$ in three experiments). Error bars represent the standard error.

(C) Summary of *a1* promoter studies. The large arrows denote the two regions of the *a1* promoter that can independently confer inducibility by C1 plus B/R and inducibility by P when fused to a heterologous minimal promoter (Grotewold et al., 1994; Tuerck and Fromm, 1994). C1 and P binding sites are boxed. Linker scanning mutations are depicted by lines underneath the altered sequences; those denoted by an asterisk represent data from **(B)**, whereas others had been published previously (Tuerck and Fromm, 1994). The proximal linker scanner had been previously tested for activation by C1 plus R and by P, with results similar to ours (Grotewold et al., 1994).

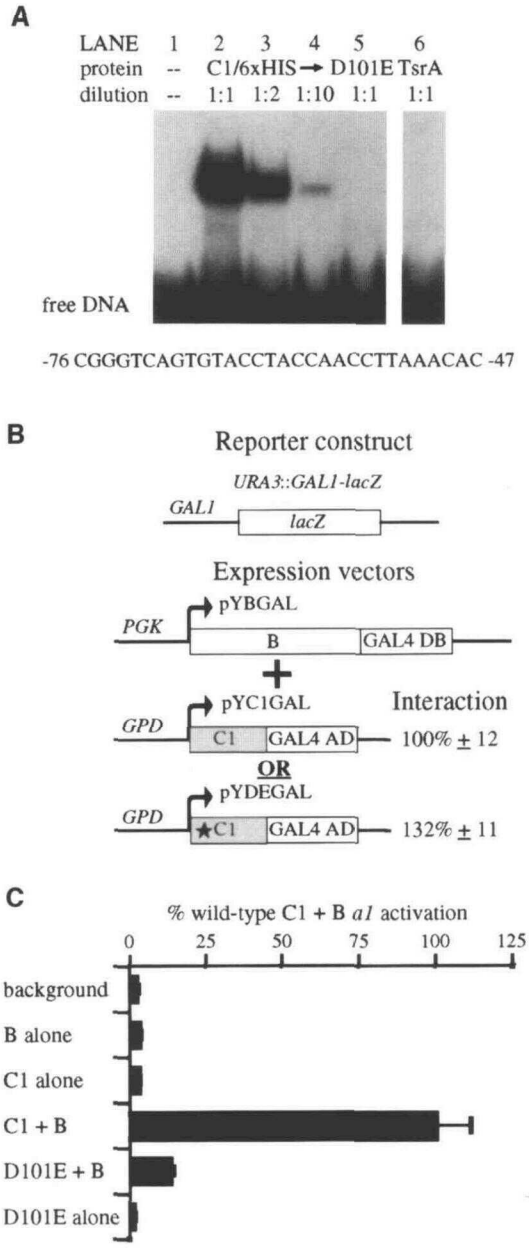

Figure 4. Functional Assays of the C1:D101E Mutant.

(A) C1:D101E mutant DNA binding activity in gel retardation assays. Gel retardation assays were done with crude extracts from bacteria expressing wild-type C1/6×HIS or the mutant C1:D101E/6×HIS proteins and the radiolabeled −76 to −47 region of the *a1* promoter. Dashes indicate the absence of a given component. Wild-type C1/6×HIS (lane 2) and dilutions (lanes 3 and 4) were used for comparisons to DNA binding by C1:D101E/6×HIS (D101E; lane 5). Lane 1 contains the free DNA; lane 6 contains the TsrA negative control from the same gel.

(B) C1:D101E mutant activity in yeast two-hybrid protein–protein interaction assays with B. The cDNAs encoding the C1 (pYC1GAL) or C1:D101E (pYDEGAL) proteins fused to sequences coding for the GAL4 activation domain in a yeast expression vector were cotransformed into yeast along with a plasmid designed to express a B/GAL4 DNA binding domain fusion protein (pYBGAL; Goff et al., 1992). *PGK* and *GPD* indicate the yeast 3-phosphoglycerate kinase and glyceraldehyde-3-phosphate dehydrogenase promoters, respectively; GAL4 DB and GAL4 AD indicate the yeast GAL4 protein DNA binding and activation domains, respectively. The interaction of C1 or C1:D101E with B was monitored by activation of a *GAL1–lacZ* reporter present in the yeast strain, as measured by β-galactosidase activity in the transformed cells. The background β-galactosidase activity generated by the empty vector plus pYBGAL averaged <1 Miller unit. pYC1GAL plus pYBGAL generated an average of 225 Miller units of β-galactosidase activity in these experiments. Data from pYDEGAL plus pYBGAL were normalized to the β-galactosidase activity generated by pYC1GAL plus pYBGAL and represent the average of 10 to 15 independent transformants from two separate experiments, ±SE.

(C) Activation of the *a1* promoter by the C1:D101E mutant. Activation by C1 plus B and C1:D101E plus B was compared in transiently transformed cultured maize cells. Transformations included the full-length (1.4-kb) *a1* promoter fused to a firefly luciferase reporter gene (pA1Luc; Klein et al., 1989) and a constitutively expressed β-glucuronidase reporter gene, which serves as a transformation and extraction control. Data were calculated as the ratio of luciferase-to-β-glucuronidase activity in extracts from the transiently transformed cells and normalized to wild-type C1 plus B activity, which was set at 100%. Activation by C1 plus B averaged 42-fold over the background observed in the absence of C1 and B. Histograms represent the average of 18 samples from three separate experiments (except C1 alone and B alone, $n = 12$ from two experiments; C1:D101E alone, $n = 6$ from one experiment). Error bars represent the standard error.

seen in Figure 4B, C1:D101E interacted with B at least as well as with wild-type C1 in yeast two-hybrid assays. Hence, the D101E mutation only affects the DNA binding activity of C1 and not its ability to interact with B, strongly suggesting that the DNA binding defect is not due to a general misfolding of the mutant protein.

To assess the importance of C1 DNA binding activity for *a1* activation, the ability of C1:D101E to activate the *a1* promoter was assayed using transient transformation of cultured maize cells. As seen in Figure 4C, C1:D101E activated the wild-type *a1* promoter at only 11% of wild-type C1 levels in the presence of B, similar to the reduction in activation observed when both C1 binding sites in *a1* were mutated (Figure 3B). Decreased activation by C1:D101E, a mutant specifically affected in DNA binding, strongly argues that C1 DNA binding activity is important in activating the *a1* promoter.

Quantitative Comparison of C1 and P DNA Binding Activities

We have shown that C1 and P bind to the same *a1* sequences in vitro. If both proteins bound to their sites with

the same affinity in vivo, then one might expect both of them to activate transcription, because both proteins have activation domains (Goff et al., 1991; E. Grotewold, M.B. Sainz, and V.L. Chandler, unpublished results). However, C1 is completely dependent on B for activation, despite the absence of a transcriptional activation domain in B (Goff et al., 1992; Sainz et al., 1997). Thus, one simple model is that C1 requires B to bind to DNA with high affinity. A prediction of this model is that C1 alone would have a significantly lower DNA binding affinity than would P in vitro.

To test this prediction, we compared the relative DNA binding affinities of the C1 and P Myb domains for the high-affinity site in the proximal *a1* promoter. Proteins with a 10-histidine residue tag fused to the Myb domains of C1 (C1myb) and of P (Pmyb) were expressed in and purified from *E. coli* and used in gel retardation assays with the −76 to −47 *a1* sequence. Free and bound bands were quantified, and the data were subjected to Scatchard analysis (Scatchard, 1949); representative graphs are depicted in Figures 5A and 5B. When we used the average of six different experiments that yielded curves with root mean square (RMS) values ≥0.8, Pmyb exhibited an equilibrium dissociation constant (K_d) of 52 nM ± 4 SE for the high-affinity *a1* site, which is comparable to the 28 nM ± 3 value recently reported by Williams and Grotewold (1997). In contrast, C1myb had a K_d of 330 nM ± 50 SE, an approximately sixfold lower affinity for the proximal *a1* sequence relative to Pmyb. These K_d values are within the range previously reported for mouse Myb domain binding to its high-affinity binding site under experimental conditions similar to ours (Ramsay et al., 1992). Although binding constants have not been determined for nearly full-length proteins, similar affinity differences have also been observed for these proteins, because Pmyb/C1/6×HIS binds to the high-affinity *a1* site with approximately three- to fourfold greater affinity than does C1/6×HIS (data not shown).

To address whether C1 and P exhibited differences in affinity for other functionally important sequences, we determined the DNA binding affinity of the C1 and P Myb domains for the low-affinity site in the distal *a1* promoter.

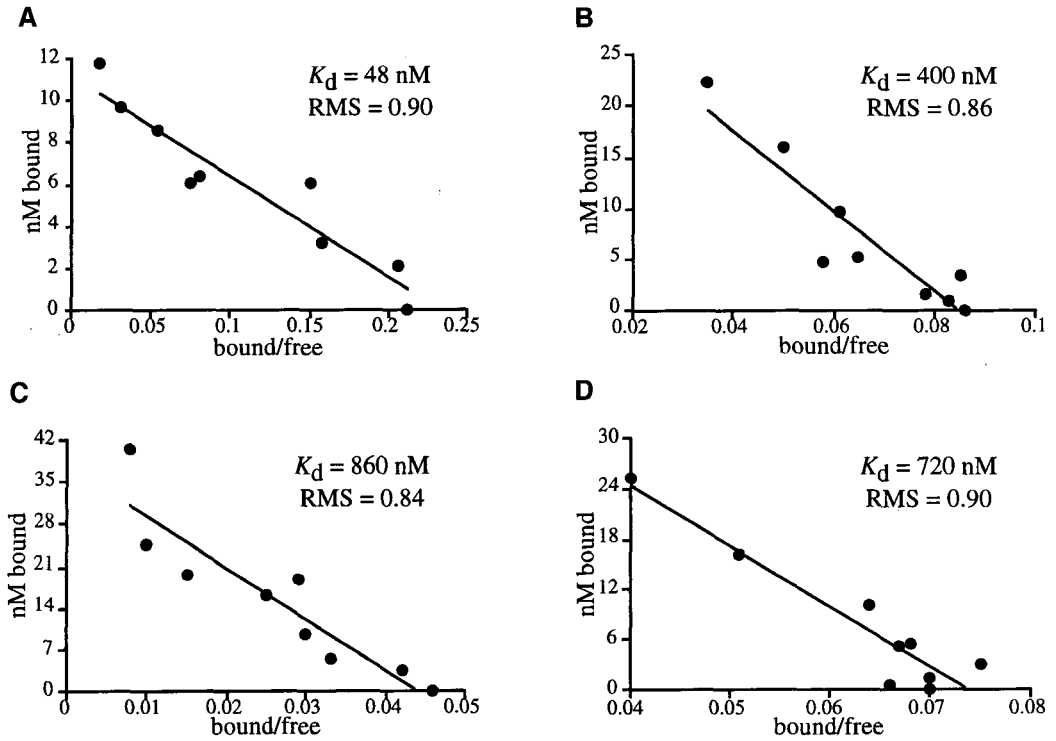

Figure 5. Scatchard Analysis of C1myb and Pmyb Binding to the High- and Low-Affinity Binding Sites on the *a1* Promoter.

Representative curves are shown, with RMS and equilibrium K_d values close to the mean values from the six experiments done for each protein and each binding site, as indicated.

(A) Pmyb binding to the high-affinity *a1* binding site.

(B) C1myb binding to the high-affinity *a1* binding site.

(C) Pmyb binding to the low-affinity *a1* binding site.

(D) C1myb binding to the low-affinity *a1* binding site.

Representative Scatchard plots are shown in Figures 5C and 5D. When we used the average of six different experiments that yielded curves with RMS values ≥0.8, C1myb and Pmyb exhibited equivalent dissociation constants for the distal *a1* sequence of 780 nM ± 70 and 860 nM ± 150 SE, respectively.

To determine what DNA sequence C1 most prefers to bind, we performed polymerase chain reaction (PCR)–based site selection, which involved cycles of C1 binding to a random pool of annealed oligonucleotides, immunoprecipitation of the protein–DNA complexes, and PCR amplification of the selected sequences (Pollock and Treisman, 1990). In contrast to P, which selected primarily sequences with an $ACC^{T}/_{A}ACC$ motif (Grotewold et al., 1994), C1 selected 26 diverse sequences from the same random pool of oligonucleotides used for the P experiment. Only some of these fragments had sites similar to the C1 binding sites in the *a1* promoter. The binding of C1 to the various site-selected sequences was quantified relative to the proximal (−76 to −47) *a1* sequence, using the same amounts of radiolabeled fragments and C1 protein in gel retardation assays. No detectable C1 binding to nine of the selected sequences was observed; the remaining 17 sequences are ranked in order of affinity for C1 in Figure 6. None of the selected sequences bound C1 with greater affinity than did the proximal *a1* sequence, suggesting that C1, like P, binds with highest affinity to the proximal *a1* binding site.

Although we have not directly identified where C1 is binding within the selected sequences, many of them have sites similar to an $A^{C}/_{A}C^{T}/_{A}A^{C}/_{A}C$ motif present in the *a1* promoter C1 binding sites. This motif, originally identified in plant flavonoid biosynthetic gene promoters (Lois et al., 1989; Loake et al., 1992), is the consensus binding site for plant Myb domain proteins from diverse species (Grotewold et al., 1994; Sablowski et al., 1994; Solano et al., 1995). None of the sequences selected by C1 had animal Myb consensus binding sites ($^{T}/_{C}AAC^{T}/_{G}G$; Biedenkapp et al., 1988), and the three selected sequences with the highest affinity for C1/6×HIS possess sites with no mismatches relative to the $A^{C}/_{A}C^{T}/_{A}A^{C}/_{A}C$ motif, suggesting that C1/6×HIS prefers to bind to sequences similar to the plant Myb domain protein consensus binding site. C1/6×HIS binds to the distal *a1* binding site with 42% of the affinity with which it binds to the high-affinity binding site in the −74 to −47 region of the *a1* promoter (Figure 5). Thus, although binding constants have not been determined, the four site-selected sequences with the highest affinity for C1/6×HIS (Figure 6) are bound by the protein with an affinity comparable (within twofold) to that of the functionally important −133 to −83 distal *a1* promoter sequence.

DISCUSSION

Sequence-specific DNA binding by C1/6×HIS to two functionally important regions of the *a1* promoter provides direct evidence that C1 directly activates the transcription of anthocyanin biosynthetic genes. The importance of C1 DNA binding activity in the activation of *a1* is demonstrated by several lines of evidence. First, C1 binding sites have been identified in each of the *a1* regions that independently confer C1 plus B inducibility to a heterologous promoter. Second, the significance of C1 DNA binding activity to C1 plus B–mediated activation is indicated by the large reduction in activation observed when both C1 binding sites in *a1* were mutated and also by the reduced activation exhibited by the C1:D101E mutant, which is specifically affected in DNA binding activity. Given that C1 binds to the *a1* promoter and interacts with the absolutely corequired B protein, B is also likely to be directly involved in activation of the *a1* gene.

The diversity of sequences identified by site selection suggests that C1 has a broader DNA binding specificity than does P. Although the selected sequences with the highest relative affinities for C1 show similarity to the plant Myb domain protein consensus binding site ($A^{C}/_{A}C^{T}/_{A}A^{C}/_{A}C$), several sequences selected by C1 have two mismatches relative to this consensus motif. An example is clone 4, to which C1/6×HIS binds with 25% of the affinity of the high-affinity *a1* binding site (Figure 6). The recognition of diverse sequences by C1 makes it difficult to identify potential C1 binding sites in anthocyanin promoters by sequence scanning. In a previous study, mutation of an animal Myb consensus binding site in the *bronze1* promoter decreased C1 plus B activation to 10% of wild-type levels and led to the hypothesis that C1 binds to this site (Roth et al., 1991). C1 exhibits sequence-specific DNA binding to the *bronze1* anthocyanin promoter (M. Sainz and V. Chandler, unpublished results), and the hypothesis that C1 binds to the animal Myb consensus site is currently being tested. Nevertheless, it appears that C1 binds with high affinity to another site in the *bronze1* promoter (M. Sainz and V. Chandler, unpublished results).

Several lines of evidence suggest that C1 does not require B or R to bind to anthocyanin promoter sequences in vivo. First, C1 exhibits sequence-specific DNA binding activity to functionally important sites in the *a1* promoter in vitro in the absence of B. Second and most importantly, the C1 and P Myb domains have similar affinities for the distal *a1* promoter in vitro, but only C1 requires B to activate this promoter fragment in vivo (Tuerck and Fromm, 1994). Hence, differences in DNA binding affinity do not correlate with differential requirements for B by C1 and P in activating the distal *a1* promoter sequence. Third, although C1myb exhibits a sixfold lower affinity than does Pmyb for the high-affinity *a1* binding site, this difference is not sufficient to explain why there is no activation by C1 alone. If B or R were required solely to boost C1 DNA binding affinity, residual activation caused by weak DNA binding and activation by C1 in the absence of B or R would be expected. No such residual activation by C1 alone has been observed with any of the anthocyanin promoters studied to date (Klein et al., 1989; Goff et al., 1990; Roth et al., 1991; Bodeau and Walbot, 1992; Grotewold et al., 1994; Tuerck and Fromm, 1994; M. Lesnick

CLONE	SEQUENCE	% -76 TO -47 *a1* BINDING	# SITES WITH MISMATCHES = 0	1	2
-76 to -47 *a1*	CGGGTCAGTGT**ACCTACCAACC**TTAAACAC	100%	2	1	
1	ccccgggtATGGATGAAAAAGTTTCACCTAACTTtggatcc	79%	1		
2	ccccgggtACAGCCTAGGGTAACATACCTACCCAtggatcc	35%	1	2	1
3	ggatccaATCAACTACCTACCAGCTTCCCTCGTacccgggg	32%	2	5	
4	ccccgggtGATGATCATGCTTCCTAAACACGTGGtggatcc	25%			4
5	ggatccaTTATTTTATGCTCATGTATCCTGACTacccgggg	13%		1	2
6	ggatccaACCGACCTGATGTTTCTGCTCCAACCacccgggg	12%		5	5
7	ggatccaGCGACACGTAATGTCTTTCTTCAACCacccgggg	12%		2	4
8	ggatccaAGACTGATCCAACGACCAGAAATGAGacccgggg	8%		1	3
9	ccccgggtTGTGTTAAGTATATCTACCACCCGAAtggatcc	5%		3	3
10	ggatccaTTATTTAGCACTACGGGAGTGCTTCAacccgggg	5%			4
11	ggatccaAAAATGTCTGCCTGCCAATCCATCCGacccgggg	5%			5
12	ggatccaTTGTCCACGAACCGATACACCTTTTCacccgggg	4%		1	1
13	ggatccaTCTAGAACTCTAACCACAAAGGTGTacccgggg	4%			5
14	ggatccaCAGAGAGCTATTGTGGTATACTAGCAacccgggg	4%		1	2
15	ggatccaTTTCGCTTACCATCCGCTCGACTGAAacccgggg	4%		1	1
16	ggatcccaTGATACCGCCTTCCCCAACCGGTCAacccgggg	3%		1	3
17	ccccgggtTCCCTTAGCCTATTATCTGTAAAAGAtggatcc	3%			

Figure 6. PCR Site Selection by C1.

Selected sequences are ranked in order of C1/6×HIS binding affinity. The fraction of bound radiolabeled DNA was determined and normalized to the bound fraction of the −76 to −47 *a1* sequence (set = 100%) in the same gel. The high-affinity C1 binding site in the −76 to −47 *a1* sequence is in boldface; lowercase letters represent linker (nonrandom) sequences. Sites within these sequences with the fewest mismatches (zero or one) relative to an $A^C/_AC^T/_AA^C/_AC$ motif are underlined. One site in clone 9 is on the opposite strand. The number of sites with zero, one, or two mismatches relative to the $A^C/_AC^T/_AA^C/_AC$ motif are indicated. Clone 17 only had sites with three or more mismatches.

and V. Chandler, unpublished results). The arguments detailed above suggest that lower DNA binding affinity is insufficient to explain why C1 requires B. Nevertheless, B may act to increase C1 DNA binding specificity, and possibly affinity, for functional sites in anthocyanin promoters. It is also possible that P requires additional uncharacterized proteins to increase its DNA binding affinity in vivo.

One hypothesis that explains why C1 requires B/R to activate is that the interaction between the C1 Myb domain and B/R alters the activation potential of C1. We speculate that the C1 Myb domain/B interaction may function at least in part to relieve an inhibitory masking of the C1 activation domain by the C1 Myb domain. Studies to date suggest that the C1 C-terminal activation domain is the only one present in the C1/B complex. An extensive analysis of B did not reveal a transcriptional activation domain in B (Goff et al., 1992). Furthermore, mutations in the C1 activation domain cause equivalent decreases in activation when assayed as fusions of the C1 activation domain to a heterologous DNA binding domain in the absence of B and when assayed in the context of otherwise native C1 activating in concert with B (Sainz et al., 1997). A C1 Myb domain/GAL4 activation domain fusion protein remains dependent on B for activating the *bronze1* and *a1* promoters (Goff et al., 1991; M. Sainz and V. Chandler, unpublished results), suggesting that if the model of C1 Myb domain inhibition of activation domain

function is correct, the C1 Myb domain interacts with and masks the GAL4 activation domain. The C1 activation domain is not masked by all Myb domains, because it can function independently of R (Grotewold et al., 1994) or B (M. Sainz and V. Chandler, unpublished results) when fused to the P Myb domain. Fusions of the C1 and P Myb domains to other transcriptional activation domains may further address this issue. Alternatively, B or R may be required for efficient nuclear localization of C1 but not of P. Nuclear localization sequences have been identified in R (Shieh et al., 1993), and similar analyses of C1 and P are in progress.

Regulation of the *a1* promoter by C1 plus B/R and by P illustrates fundamental principles of eukaryotic gene expression. The modular structure of the *a1* promoter is demonstrated by the presence of two binding sites of roughly equal functional importance for activation by C1 plus B and by P in vivo. The importance of combinatorial interactions of transcription factors in the activation of *a1* is demonstrated by the selective requirement for B or R by C1 but not by P. Future experiments with chimeric proteins and other anthocyanin promoters should further elucidate the differential regulation of maize flavonoid biosynthetic genes by distinct sets of transcription factors.

METHODS

Bacterial Expression and Purification of Proteins

For expression of C1/6×HIS and Pmyb/C1/6×HIS, an NdeI site was engineered by site-directed mutagenesis (Ausubel et al., 1987) at the first ATG codons of the cDNAs encoding C1 (Cone et al., 1986) and P (Grotewold et al., 1994). The sequences encoding the Pmyb/C1 protein (P amino acids 1 to 118 plus C1 amino acids 126 to 253) were fused as previously described (Grotewold et al., 1994). C1 or C1:D101E (amino acids 1 to 253) and Pmyb/C1 were cloned as NdeI-XhoI fragments into pET-25b (Novagen, Madison, WI), generating C-terminal translational fusions to the herpes simplex virus epitope and six histidine residues encoded by the vector. Full-length C1 (amino acids 1 to 273) was cloned as an NdeI-BamHI fragment into pBH500 (Hoopes et al., 1992), enabling expression to be driven by the T7 promoter. Sequences encoding the C1 and P Myb domains (amino acids 1 to 119) were amplified using polymerase chain reaction (PCR) primers with appropriate restriction sites and then cloned as XhoI-BamHI fragments into pET-19b (Novagen), generating N-terminal translational fusions to the 10 histidine residues encoded by the vector.

Escherichia coli BL21(DE3) was cotransformed with the constructs described above and a plasmid expressing the *E. coli dnaY* gene, which encodes a rare arginine tRNA (Brinkmann et al., 1989) and is required for overexpression of C1 proteins. Cultures with an OD_{600} of 0.8 to 1.0 were induced by adding isopropyl β-D-thiogalactopyranoside to 0.4 mM and then harvested after 1 hr (C1, C1/6×HIS, and Pmyb/C1/6×HIS) or 3 hr (C1myb or Pmyb) at 30°C. Denatured crude protein extracts of C1, C1/6×HIS, and Pmyb/C1/6×HIS were prepared essentially as described for Myb (Garcia et al., 1991), except that cells were lysed by sonication on ice. For protein purification, crude extracts were in lysis buffer minus EDTA (10 mM Tris-HCl, pH 7.8, 50 mM NaCl, 6 M urea, 1 mM phenylmethylsulfonyl fluoride [PMSF], 1 mM leupeptin, 0.3 μM aprotinin, 1 μM leupeptin, and 1 μM pepstatin). For C1myb and Pmyb, cell pellets were resuspended in HP buffer (50 mM Na_2HPO_4/NaH_2PO_4, pH 8.0, 300 μM NaCl, 6 M urea, 10 mM EDTA, 1 μM PMSF, 0.3 mM aprotinin, 1 μM leupeptin, and 1 μM pepstatin) before sonication.

All proteins were purified from denatured crude extracts by Ni^+ affinity chromatography (Abate et al., 1990) by using Ni^+–nitriloacetic acid agarose (Qiagen, Chatsworth, CA) or HIS-bind resin (Novagen) and the conditions for purification recommended by Novagen. Proteins were eluted in elution buffer (0.5 M NaCl, 400 mM imidazole, 6 M urea, 20 mM Tris-HCl, pH 7.9). C1/6×HIS and Pmyb/C1/6×HIS proteins were concentrated using a Centricon-30 (Amicon, Beverly, MA) per the manufacturer's instructions and stored at −20°C. Denatured crude extracts or purified protein preparations of C1, C1/6×HIS, or Pmyb/C1/6×HIS used for gel retardation assays were renatured by dilution, essentially as described by Garcia et al. (1991). The C1/6×HIS or Pmyb/C1/6×HIS preparations used for DNase I footprinting were renatured by dialysis at 4°C against renaturing buffer (50 mM NaCl, 10 mM Tris-HCl, pH 7.8, 1 mM EDTA, 20% glycerol, 1 mM PMSF, 0.3 μM aprotinin, 1 μM leupeptin, and 1 μM pepstatin). For C1myb and Pmyb, eluted proteins were renatured by dialysis at 4°C against WB buffer (50 mM Na_2HPO_4/NaH_2PO_4, pH 8.0, 300 mM NaCl, 1% Tween 20, 10% glycerol, 5 mM 2-mercaptoethanol, 10 mM EDTA, 1 mM PMSF, 0.3 μM aprotinin, 1 μM leupeptin, and 1 μM pepstatin). All renatured proteins were stored in small aliquots at −70°C. Protein concentrations were determined by Lowry assays (Lowry et al., 1951). Expression and the extent of purification for all proteins were assayed on Coomassie Brilliant Blue R 250–stained SDS–polyacrylamide gels by using standard techniques (Ausubel et al., 1987). Wild-type C1/6×HIS and C1:D101E/6×HIS mutant proteins were expressed to the same extent (data not shown).

Gel Retardation Assays

Synthetic oligonucleotides of the appropriate *a1* promoter region were end-labeled using T4 polynucleotide kinase (New England Biolabs, Beverly, MA) and a molar excess of γ-^{32}P-ATP (>7000Ci/mM; ICN Biomedicals, Costa Mesa, CA) for 1 hr at 37°C. An equimolar amount of the complementary oligonucleotide and NaCl to 50 mM NaCl were added to the kinase reaction, and the oligonucleotides were annealed by cooling from 90°C to room temperature over ~2 hr and purified on native polyacrylamide gels.

Gel retardation assays were as described by Garcia et al. (1991), except that binding reactions included 100 μg/mL poly(dI-dC) (Pharmacia Biotechnology), 4 μg/mL sheared salmon sperm DNA, ~20,000 to 60,000 cpm (~10 to 120 fmol) of labeled DNA, and ~5 to 50 μg of C1, C1/6×HIS, or Pmyb/C1/6×HIS protein. For quantitative binding curves, 25-μL reactions were used, the salmon sperm DNA was omitted, and reactions included 0.05 to 2 μg of C1myb or Pmyb protein in WB buffer. Appropriate reactions were preincubated with antisera, purified antibody, or unlabeled DNA for 20 to 30 min at 4°C, followed by the addition of labeled DNA. The amount of unlabeled DNA varied depending on the experiment (225 pmol per reaction for crude bacterial protein extracts, 450 pmol per reaction using purified C1/6×HIS and Pmyb/C1/6×HIS, and 0.25 to 128 pmol per reaction for quantitative binding analysis). Competition with unrelated competitor DNA had little or no effect on C1/6×HIS DNA binding under our conditions (data not shown).

Binding was for 20 to 30 min at 4°C, and binding reactions were loaded on 1.5-mm-thick 5% polyacrylamide gels (80:1 acrylamide–

bisacrylamide) in 0.25 × TBE buffer (1 × TBE is 89 mM Tris base, 89 mM H_3BO_3, and 2 mM EDTA) and run at 20 V/cm for 70 min (C1, C1/6×HIS, and Pmyb/C1/6×HIS) or 90 min (C1myb and Pmyb) at 4°C. Gels were prerun for at least 2 hr. Free and bound DNA complexes were visualized by autoradiography of dried gels and were quantified using an AMBIS (San Diego, CA) or PhosphorImager (Molecular Dynamics, Sunnyvale, CA) radioanalytic scanner, according to the manufacturer's instructions.

DNase I Footprinting

DNA was radioactively labeled using the Klenow fragment of DNA polymerase I (Boehringer Mannheim) to fill in a restriction site with only G and C residues (BssHII) or A and T residues (EcoRI) by using a twofold molar excess of each of the two appropriate radioactive deoxynucleotide triphosphates (Du Pont–New England Nuclear) relative to sites at which they could potentially be incorporated. DNase I protection assays were performed essentially as described by Andrews et al. (1987). Binding reactions were performed as they were for the gel retardation assays, except that ~8 to 80 μg of partially purified C1/6×HIS or Pmyb/C1/6×HIS protein was used in a 45-μL volume, and binding reactions included 2.5 mM $MgCl_2$, 0.5 mM $CaCl_2$, and ~100,000 cpm of radiolabeled DNA. Immediately before loading, 5 μL of a fresh dilution of RNase-free DNase I (Boehringer Mannheim) in DNase I buffer (50% glycerol, 50 mM NaCl, 50 mM Tris-HCl, pH 7.2, 10 mM $MgSO_4$, 100 μg/mL BSA, and 1 mM DTT) was added, and reactions were incubated for 30 sec at 4°C and then stopped by the addition of 5 μL of 0.5 M EDTA, pH 8.0. Gels were run as described for the gel retardation assays, wrapped, and exposed to film, and the free and bound fragments were cut out, electroeluted, and ethanol precipitated. Free and bound DNA was resuspended at equivalent cpm/μL, run on a 6% sequencing gel, dried, and exposed to film.

PCR Site Selection

The pool of DNA used has a central 26-bp random sequence flanked by 25-bp linkers and was constructed by annealing a primer to the constant linker sequence of single-stranded oligonucleotides, extending with the Klenow fragment and purifying the double-stranded products on native polyacrylamide gels, as previously described (Grotewold et al., 1994). Five rounds of selection were performed. Double-stranded oligonucleotides (0.4 ng in rounds 1 and 2 or 0.2 ng in rounds 3 to 5) were used in 50-μL binding reactions, as was described for the gel retardation assays.

In rounds 1 and 2, purified C1/6×HIS protein was used. Protein–DNA complexes were immunoprecipitated by the addition of 1 μg of herpes simplex virus monoclonal antibody (Novagen) and 5 μL of a 50% slurry of protein A–Sepharose (Pharmacia). PCR was performed using primers to the constant linker sequences as follows: 5 min at 94°C, then 1 min at 94°C, 1 min at 50°C, and 30 sec at 72°C for 15 to 20 cycles, followed by a final 20-min extension at 72°C. PCR products to be used in subsequent rounds were gel purified on native polyacrylamide gels.

For rounds 3 through 5, crude extracts from bacteria expressing full-length C1 protein were used in the binding reactions. Gel retardation assays were performed, and the shifted complexes were eluted, ethanol precipitated, and amplified by PCR as before. After the final round, the PCR products were digested with EcoRI and BamHI and cloned into pBluescript KS+ (Stratagene, La Jolla, CA), and 26 clones were sequenced (Ausubel et al., 1987). A constant amount of purified C1/6×HIS protein was used in gel retardation assays with radiolabeled selected DNA sequences, as described above. Binding was quantified as the percentage of DNA bound, using a PhosphorImager per the manufacturer's instructions. Sequences with similarity to $A^{C}/_{A}C^{T}/_{A}A^{C}/_{A}C$ or $^{T}/_{C}AAC^{T}/_{G}G$ motifs were found using the FIND program in the Genetics Computer Group (Madison, WI) sequence analysis package (Devereux et al., 1984).

Yeast Two-Hybrid Assays and Protein Gel Blot Analyses

The pYBGAL and pYC1GAL vectors, designed to express the B/GAL4 DNA binding domain (*PGK* promoter and terminator, *CEN/ARS*, *LEU2*) and the C1/GAL4 activation domain (*GPD* promoter, 2-μm, *PGK* terminator, *HIS3*) fusion proteins, respectively, have been described previously (Goff et al., 1992). The wild-type sequence encoding C1 amino acids 1 to 144 in pYC1GAL was replaced with a sequence bearing the D101E mutation on a BamHI-AatII fragment, generating pYDEGAL. The presence of the D101E mutation in pYDEGAL was confirmed by sequencing (Ausubel et al., 1987). Yeast strain GGY::171 (*gal4Δ gal80Δ his3 leu2 URA3::Gal1-lacZ*) (Gill and Ptashne, 1987) was transformed using a modified Li^+ transformation protocol (Schiestl and Gietz, 1989), with pYBGAL and pYC1GAL or pYDEGAL. Cells were plated and grown on synthetic defined medium lacking leucine and histidine at 30°C. Independent transformants were transferred to liquid culture and grown to saturation under selection. β-Galactosidase assays and units of activity were calculated as described by Miller (1972). Cultures used for β-galactosidase assays were pooled and used to inoculate fresh selective medium at a 1:10 dilution, grown to saturation, and used to make protein extracts. Yeast protein extracts and protein gel blot analysis were as described previously (Sainz et al., 1997), except that secondary antibodies coupled to horseradish peroxidase (Bio-Rad) were detected using a chemiluminescence kit (Durrant, 1990).

Maize Transient Transformation Assays

A mutation in the distal C1 and P binding site at −116 to −124 in the *a1* promoter, changing the sequence from AACTACCGG to TCTAGAGGG, was generated by PCR site-directed mutagenesis (Ausubel et al., 1987) in plasmids with either the wild-type −219 to +7 *a1* promoter or a derivative with a mutated (ACCTACCAACC to ACCCGATCGTC) proximal C1 and P binding site (Grotewold et al., 1994). The pABR4 plant reporter vector was made by cloning the sequence containing the first intron of the maize *Adh1-S* alcohol dehydrogenase gene, the firefly luciferase coding region, and the *Agrobacterium tumefaciens* nopaline synthase terminator (*nos* 3′ end) from pJD312 (Luehrsen et al., 1992) downstream of the polylinker of the pUC6s vector (Vieira and Messing, 1991). Plasmids with the wild-type sequence (pA1WT), either of the single binding site mutations (pA1PRO and pA1DIS) or both mutations (pA1DBL), were generated by cloning the appropriate *a1* promoter sequences upstream of the maize *Adh1* intron in pABR4.

Vectors designed to express full-length C1, C1:D101E, or P in maize cells were constructed and named pZmC1, pZmDE, and pZmP, respectively. Coding regions of the respective proteins with the 45-nucleotide shortened *c1* 5′ untranslated leader sequence (Goff et al., 1992) were cloned downstream of the cauliflower mosaic virus 35S promoter and *Adh1* intron and upstream of the *nos* 3′ end in pMF6 (Callis et al., 1987). The pZmP vector includes 3′ *p* cDNA un-

translated trailer sequences and the bacteriophage T7 gene 10 terminator upstream of the *nos* 3′ end in pMF6. The *B-Peru* cDNA driven by the 35S cauliflower mosaic virus promoter (p35SB; Goff et al., 1990) and the pA1Luc reporter plasmid featuring the 1.4-kb *a1* promoter fused to the firefly luciferase gene (Klein et al., 1989) have been described previously. The constitutively expressed β-glucuronidase gene plasmid p35SIG (Bodeau and Walbot, 1992) was included as a transformation and extraction control. All plasmid DNA was purified by $CsCl_2$ equilibrium gradient ultracentrifugation and resuspended at ~1 mg/mL. For the −219 to +7 *a1* promoter reporter constructs and for pZmC1, pZmDE, and pZmP, two different plasmid DNA preparations were purified and tested in separate experiments.

Maize transient transformation by microprojectile bombardment was performed as previously described (Sainz et al., 1997). Precipitations onto 1-μm gold microprojectiles included 10 μg each of reporter plasmid and p35SIG control plasmid together with a total of 2 μg of expression vector: 1 μg of p35SB plus 1 μg of pZmC1 or pZmDE, 1 μg each of pZmP and pMF6, or simply 2 μg of pMF6 (empty vector background). Tissue samples consisting of 0.4 mL of packed L6 maize suspension cells (Ciba-Geigy Agricultural Biotechnology Unit, Research Triangle Park, NC) spread on N6 media (Lowe et al., 1985) were bombarded with gold microprojectiles accelerated by a helium gas shock wave. Plates were incubated for ~36 hr in the dark at 28°C. Luciferase was assayed as previously described (Goff et al., 1992). Luciferase activity was expressed as light units detected by using a luminometer (model 3010; Analytical Scientific Instruments, Richmond, CA; or model TD-20e; Turner Designs, Sunnyvale, CA) at room temperature for 10 or 15 sec, respectively. β-Glucuronidase assays (Jefferson, 1987) were performed as previously described (Sainz et al., 1997). Activation was calculated for each sample as the ratio of luciferase to β-glucuronidase activity, minus the respective backgrounds.

ACKNOWLEDGMENTS

We appreciate the helpful comments of Alice Barkan and Marc Lesnick. We thank Devon Brown for construction of the pABR4 plasmid and Steve Goff for the C1:D101E mutant. M.B.S. was a Howard Hughes Medical Institute predoctoral fellow, and the work was supported by National Science Foundation Grants No. MCB-9406233 to E.G. and No. MCB-9304687 to V.L.C.

Received November 14, 1996; accepted February 14, 1997.

REFERENCES

Abate, C., Luk, D., Gentz, R., Rauscher III, F.J., and Curran, T. (1990). Expression and purification of the leucine zipper and DNA-binding domains of Fos and Jun: Both Fos and Jun contact DNA directly. Proc. Natl. Acad. Sci. USA **87,** 1032–1036.

Ames, P., and Parkinson, J.S. (1994). Constitutively signaling fragments of Tsr, the *Escherichia coli* serine chemoreceptor. J. Bacteriol. **176,** 6340–6348.

Andrews, B.J., Beatty, L.G., and Sadowski, P.D. (1987). Isolation of intermediates in the binding of FLP recombinase of the yeast plasmid 2-micron circle to its target sequence. J. Mol. Biol. **193,** 345–358.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K., eds (1987). Current Protocols in Molecular Biology. (New York: Wiley-Interscience).

Biedenkapp, H., Borgmeyer, U., Sippel, A.E., and Klempnauer, K.-H. (1988). Viral *myb* oncogene encodes a sequence-specific DNA-binding activity. Nature **335,** 835–837.

Bodeau, J.P., and Walbot, V. (1992). Regulated transcription of the maize *Bronze-2* promoter in electroporated protoplasts requires the *C1* and *R* gene products. Mol. Gen. Genet. **233,** 379–387.

Brinkmann, U., Mattes, R.E., and Buckel, P. (1989). High-level expression of recombinant genes in *Escherichia coli* is dependent on the availability of the *dnaY* gene product. Gene **85,** 109–114.

Callis, J., Fromm, M., and Walbot, V. (1987). Introns increase gene expression in cultured maize cells. Genes Dev. **1,** 1183–1200.

Chandler, V.L., Radicella, J.P., Robbins, T.P., Chen, J., and Turks, D. (1989). Two regulatory genes of the maize anthocyanin pathway are homologous: Isolation of *B* utilizing *R* genomic sequences. Plant Cell **1,** 1175–1183.

Cone, K.C., and Burr, B. (1989). Molecular and genetic analysis of the light requirement for anthocyanin synthesis in maize. In The Genetics of Flavonoids, D.E. Styles, G.A. Gavazzi, and M.L. Racchi, eds (Milan: Edizioni Unicopli), pp. 143–145.

Cone, K.C., Burr, F.A., and Burr, B. (1986). Molecular analysis of the maize anthocyanin regulatory locus *C1*. Proc. Natl. Acad. Sci. USA **83,** 9631–9635.

Cone, K.C., Cocciolone, S.M., Burr, F.A., and Burr, B. (1993). Maize anthocyanin regulatory gene *pl* is a duplicate of *c1* that functions in the plant. Plant Cell **5,** 1795–1805.

Dellaporta, S.L., Greenblatt, I., Kermicle, J.L., Hicks, J.B., and Wessler, S.R. (1988). Molecular cloning of the maize *R-nj* allele by transposon tagging with *Ac*. In Chromosome Structure and Function: Impact of New Concepts, 18th Stadler Genetics Symposium, J.P. Gustafson and R. Appels, eds (New York: Plenum Press), pp. 263–282.

Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12,** 387–395.

Dooner, H.K., Robbins, T.P., and Jorgensen, R.A. (1991). Genetic and developmental control of anthocyanin biosynthesis. Annu. Rev. Genet. **25,** 173–199.

Durrant, I. (1990). Light-based detection of biomolecules. Nature **346,** 297–298.

Garcia, A., LaMontagne, K., Reavis, D., Stober-Grasser, U., and Lipsick, J.S. (1991). Determinants of sequence-specific DNA-binding by $p48^{v\text{-}myb}$. Oncogene **6,** 265–273.

Gill, G., and Ptashne, M. (1987). Mutants of GAL4 protein altered in an activation function. Cell **51,** 121–126.

Goff, S.A., Klein, T.M., Roth, B.A., Fromm, M.E., Cone, K.C., Radicella, J.P., and Chandler, V.L. (1990). Transactivation of anthocyanin biosynthetic genes following transfer of *B* regulatory genes into maize tissues. EMBO J. **9,** 2517–2522.

Goff, S.A., Cone, K.C., and Fromm, M.E. (1991). Identification of functional domains in the maize transcriptional activator C1:

Comparison of wild-type and dominant inhibitor proteins. Genes Dev. **5,** 298–309.

Goff, S.A., Cone, K.C., and Chandler, V.L. (1992). Functional analysis of the transcriptional activator encoded by the maize B gene: Evidence for a direct functional interaction between two classes of regulatory proteins. Genes Dev. **6,** 864–875.

Grotewold, E., Athma, P., and Peterson, T. (1991). Alternatively spliced products of the maize *P* gene encode proteins with homology to the DNA-binding domain of *myb*-like transcription factors. Proc. Natl. Acad. Sci. USA **88,** 4587–4591.

Grotewold, E., Drummond, B.J., Bowen, B., and Peterson, T. (1994). The *myb*-homologous *P* gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset. Cell **76,** 543–553.

Hoopes, B.C., LeBlanc, J.F., and Hawley, D.K. (1992). Kinetic analysis of yeast TFIID-TATA box complex formation suggests a multi-step pathway. J. Biol. Chem. **267,** 11539–11547.

Howe, K.M., Reakes, C.F.L., and Watson, R.J. (1990). Characterization of the sequence-specific interaction of mouse *c-myb* protein with DNA. EMBO J. **9,** 161–169.

Jefferson, R.A. (1987). Assaying chimeric genes in plants: The GUS gene fusion system. Plant Mol. Biol. Rep. **5,** 387–405.

Klein, T.M., Roth, B.A., and Fromm, M.E. (1989). Regulation of anthocyanin biosynthetic genes introduced into intact maize tissues by microprojectiles. Proc. Natl. Acad. Sci. USA **86,** 6681–6685.

Loake, G.J., Faktor, O., Lamb, C.J., and Dixon, R.A. (1992). Combination of H-box [CCTACC$(N)_7$CT] and G-box (CACGTG) *cis* elements is necessary for feed-forward stimulation of a chalcone synthase promoter by the phenylpropanoid-pathway intermediate *p*-coumaric acid. Proc. Natl. Acad. Sci. USA **89,** 9230–9234.

Lois, R., Dietrich, A., Hahlbrock, K., and Schulz, W. (1989). A phenylalanine ammonia-lyase gene from parsley: Structure, regulation, and identification of elicitor and light-responsive *cis*-acting elements. EMBO J. **8,** 1641–1648.

Lowe, K.S., Taylor, D.B., Ryan, P.L., and Paterson, K.P. (1985). Plant regeneration via organogenesis and embryogenesis in the maize inbred line B73. Plant Sci. **41,** 125–132.

Lowry, O.H., Rosebrough, N.J., Farr, A.L., and Randall, R.J. (1951). Protein measurement with the Folin phenol reagent. J. Biol. Chem. **193,** 265–275.

Ludwig, S.R., and Wessler, S.R. (1990). Maize *R* gene family: Tissue-specific helix-loop-helix proteins. Cell **62,** 849–851.

Ludwig, S.R., Habera, L.F., Dellaporta, S.L., and Wessler, S.R. (1989). *Lc*, a member of the maize *R* gene family responsible for tissue-specific anthocyanin production, encodes a protein similar to transcriptional activators and contains the *myc*-homology region. Proc. Natl. Acad. Sci. USA **86,** 7092–7096.

Luehrsen, K.R., de Wet, J.R., and Walbot, V. (1992). Transient expression analysis in plants using the firefly luciferase reporter gene. Methods Enzymol. **216,** 397–414.

Miller, J.H. (1972). Assay of β-galactosidase. In Experiments in Molecular Genetics (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory), pp. 352–355.

Ogata, K., Morikawa, S., Nakamura, H., Sekikawa, A., Inoue, T., Kanai, H., Sarai, A., Ishii, S., and Nishimura, Y. (1994). Solution structure of a specific DNA complex of the Myb DNA-binding domain with cooperative recognition helices. Cell **79,** 639–648.

Patterson, G.I., Thorpe, C.J., and Chandler, V.L. (1993). Paramutation, an allelic interaction, is associated with a stable and heritable reduction of transcription of the maize *b* regulatory gene. Genetics **135,** 881–894.

Paz-Ares, J., Ghosal, D., Wienand, U., Peterson, P.A., and Saedler, H. (1987). The regulatory *c1* locus of *Zea mays* encodes a protein with homology to *myb* proto-oncogene products and with structural similarities to transcriptional activators. EMBO J. **6,** 3553–3558.

Paz-Ares, J., Ghosal, D., and Saedler, H. (1990). Molecular analysis of the C1-I allele from *Zea mays*: A dominant mutant of the regulatory C1 locus. EMBO J. **9,** 315–321.

Pollock, R., and Treisman, R. (1990). A sensitive method for the determination of protein–DNA binding specificities. Nucleic Acids Res. **18,** 6197–6204.

Radicella, J.P., Turks, D., and Chandler, V.L. (1991). Cloning and nucleotide sequence of a cDNA encoding *B-Peru*, a regulatory protein of the anthocyanin pathway in maize. Plant Mol. Biol. **17,** 127–130.

Ramsay, R.G., Ishii, S., and Gonda, T.J. (1992). Interaction of the Myb protein with specific DNA binding sites. J. Biol. Chem. **267,** 5656–5662.

Roth, B.A., Goff, S.A., Klein, T.M., and Fromm, M.E. (1991). *C1*- and *R*-dependent expression of the maize *Bz1* gene requires sequences with homology to mammalian *myb* and *myc* binding sites. Plant Cell **3,** 317–325.

Sablowski, R.W.M., Moyano, E., Culianez-Macia, F.A., Schuch, W., Martin, C., and Bevan, M. (1994). A flower-specific Myb protein activates transcription of phenylpropanoid biosynthetic genes. EMBO J. **13,** 128–137.

Sainz, M.B., Goff, S.A., and Chandler, V.L. (1997). Extensive mutagenesis of a transcriptional activation domain identifies single hydrophobic and acidic amino acids important for activation in vivo. Mol. Cell. Biol. **17,** 115–122.

Scatchard, G. (1949). The attractions of proteins for small molecules and ions. Ann. N.Y. Acad. Sci. **51,** 660–672.

Schiestl, R.H., and Gietz, R.D. (1989). High efficiency transformation of intact yeast cells using single stranded nucleic acids as carrier. Curr. Genet. **16,** 339–346.

Schwarz-Sommer, Z., Shepherd, N., Tacke, E., Gierl, A., Rohde, W., Leclerq, L., Mattes, M., Berndtgen, R., Petersen, P.A., and Saedler, H. (1987). Influence of transposable elements on the structure and function of the *A1* gene of *Zea mays*. EMBO J. **6,** 287–294.

Shieh, M.W., Wessler, S.R., and Raikhel, N.V. (1993). Nuclear targeting of the maize R protein requires two nuclear localization sequences. Plant Physiol. **101,** 353–361.

Solano, R., Nieto, C., Avila, J., Canas, L., Diaz, I., and Paz-Ares, J. (1995). Dual DNA binding specificity of a petal epidermis-specific MYB transcription factor (MYB.Ph3) from *Petunia hybrida*. EMBO J. **14,** 1773–1784.

Styles, E.D., and Ceska, O. (1977). The genetic control of flavonoid synthesis in maize. Can. J. Genet. Cytol. **19,** 289–302.

Tuerck, J.A., and Fromm, M.E. (1994). Elements of the maize *A1* promoter required for transactivation by the anthocyanin *B/C1* or phlobaphene *P* regulatory genes. Plant Cell **6,** 1655–1663.

Van der Meer, I.M., Stuitje, A.R., and Mol, J.N.M. (1993). Regulation of general phenylpropanoid and flavonoid gene expression. In

Control of Plant Gene Expression, D.P.S. Verma, ed (Boca Raton, FL: CRC Press), pp. 125–155.

Vieira, J., and Messing, J. (1991). New pUC-derived cloning vectors with different selectable markers and DNA replication origins. Gene **100,** 189–194.

Weintraub, H., Davis, R., Tapscott, S., Thayer, M., Krause, M., Benezra, R., Blackwell, T.K., Turner, D., Rupp, R., Hollenberg, S., Zhuang, Y., and Lassar, A. (1991). The *myoD* gene family: Nodal point during specification of the muscle cell lineage. Science **251,** 761–766.

Williams, C.E., and Grotewold, E. (1997). Differences between plant and animal Myb domains are fundamental for DNA-binding activity and chimeric Myb domains have novel DNA-binding specificities. J. Biol. Chem. **272,** 563–571.