The Plant Cell, Vol. 6, 1655–1663, November 1994

Elements of the Maize *A1* Promoter Required for Transactivation by the Anthocyanin *B/C1* or Phlobaphene *P* Regulatory Genes

Jutta A. Tuerck[1] and Michael E. Fromm[2]

Monsanto Company, Plant Sciences, 700 Chesterfield Parkway, St. Louis, Missouri 63198

The extensive genetic and molecular characterization of the flavonoid pathway's structural and regulatory genes has provided some of the most detailed knowledge of gene interactions in plants. In maize flavonoid biosynthesis, the *A1* gene is independently regulated in the anthocyanin and phlobaphene pathways. Anthocyanin production requires the expression of the *C1* or *Pl* and *R* or *B* regulatory genes, whereas phlobaphene production requires only the *P* regulatory gene. By deletion analysis of the *A1* promoter, we show that the sequences between −123 and −88 are critical for activation by anthocyanin and phlobaphene regulatory genes. Linker-scanner mutations indicated that the −123 to −100 region is more important for transactivation by the P protein. The −98 to −88 region is more important for B/C1 transactivation and shows a strong homology with the region of the *Bz1* anthocyanin structural gene promoter shown to be activated by B/C1 and not by P. We identified a 14-bp consensus sequence that is also present in the promoters of three other genes in the anthocyanin pathway, and we propose a model for how the flavonoid regulatory proteins interact with the promoters of the structural genes.

INTRODUCTION

The extensive genetic and molecular characterization of flavonoid pigment production in maize, *Antirrhinum*, and petunia has provided some of the most detailed knowledge of gene interactions in plants. Most of the structural and regulatory genes have been isolated, and extensive genetic data exist on their developmental and environmental regulation, as well as their often complex allelic interactions and expression patterns (for review, see Dooner et al., 1991). In maize, the flavonoid pathway produces both the anthocyanins, which are located in most parts of the plant, and the phlobaphenes, which are limited primarily to the seed pericarp, cob, and floral structures. The two pigment pathways share the initial biosynthetic steps encoded by the genes *C2* (chalcone synthase; Wienand et al., 1986), *CHI* (chalcone isomerase; Grotewold and Peterson, 1994), and *A1* (NADPH-dependent flavonol reductase; Schwarz-Sommer et al., 1987). However, subsequent structural genes in the anthocyanin pathway, such as *A2* (flavonol dehydroxylase; Menssen et al., 1990), *Bronze-1* (*Bz1*, UDP glucose:flavonol-3-O-glucosyl transferase; Furtek et al., 1988), and *Bronze-2* (*Bz2;* Nash et al., 1990), are not required in the phlobaphene pathway (Coe et al., 1988).

Four regulatory genes in two gene families control the expression of the anthocyanin biosynthetic structural genes. A member of each of the *B/R* and *C1/Pl* gene families must be expressed for activation of the anthocyanin pathway. The *B* and *R* genes show extensive allelic diversity (Styles, 1970) and can substitute for each other genetically or when expressed from a constitutive cauliflower mosaic virus (CaMV) 35S promoter (Ludwig et al., 1989; Goff et al., 1990). The B/R proteins show sequence homology to the basic-helix-loop-helix (bHLH) DNA binding/dimerization domain found in MYC and MYOD proteins (Chandler et al., 1989; Ludwig et al., 1989). The *C1* and *Pl* genes show less genetic diversity and encode proteins with sequence homology with the DNA binding regions of the mammalian MYB protooncogene family of proteins (Paz-Ares et al., 1987; Cone et al., 1993).

Genetic studies have found *P* to be the only regulatory gene in the phlobaphene biosynthetic pathway of maize (Styles and Ceska, 1977). The *P* gene regulates the accumulation of the transcripts of the *C2*, *CHI*, and *A1* genes in the pericarp but does not activate the other anthocyanin structural genes such as *Bz1* (Grotewold et al., 1994). The P protein has homology with the DNA binding regions of MYB proteins, and this N-terminal region of the predicted P protein is 80% identical to the C1 protein (Grotewold et al., 1991). P protein produced in *Escherichia coli* binds to the consensus site ACCTACCA in vitro, and this sequence appears to play a role in the inducibility of the *A1* promoter in vivo (Grotewold et al., 1994).

Previously, a deletion analysis of the *Bz1* promoter defined a *cis*-acting region that responds to regulation by B and C1

[1] Current address: Department of Genetics, University of Cambridge, Downing Street, Cambridge, CB2 3EH, UK.
[2] To whom correspondence should be addressed.

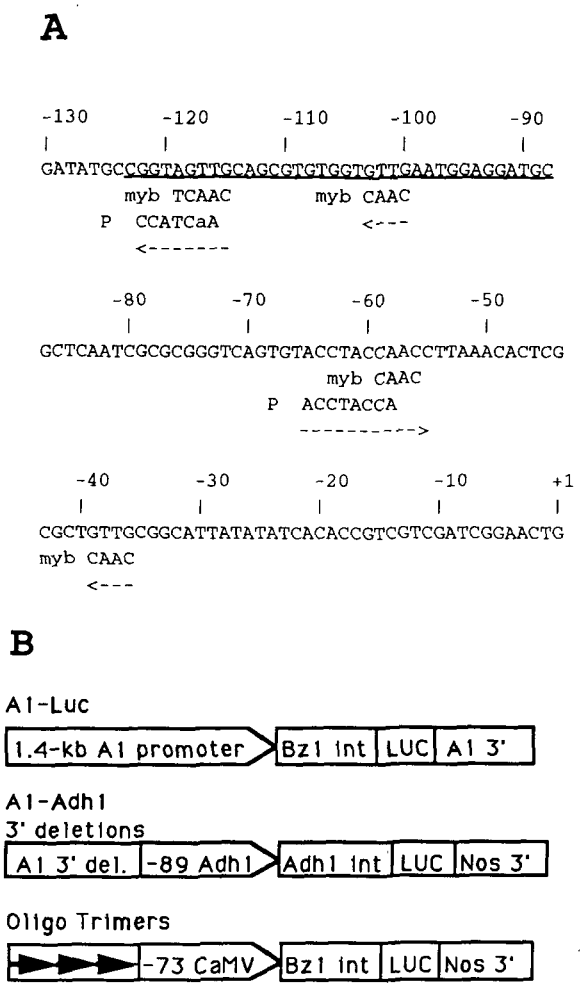

Figure 1. *A1* Promoter Sequence and Structure of Genes Containing Regions of the *A1* Promoter.

(A) Sequence of the *A1* promoter from −130 to +1. Numbering is relative to the *A1* start of transcription. The experimentally defined region important for *A1* promoter activity is underlined (−123 to −88). Potential imperfect *myb*-like (myb) sites and a *P*-like (P) site that have homology to the consensus sequences for MYB binding (C/TAACG/TG) and P binding (ACCTACCA; at −60) are indicated under the *A1* sequence. Leftward arrows indicate that the *myb*-like sites or *P*-like site are located on the complementary DNA strand and in such cases show the complementary sequence in a 3′ to 5′ orientation. The lowercase letter in the leftward P site indicates that it does not match the P consensus sequence.

(B) Structure of *Luc* Reporter Genes Fused to *A1* Promoter Elements. A1-Luc contains the full 1.4-kb *A1* promoter that was used for constructing 5′ deletions of the *A1* promoter. The A1-Adh1 3′ deletions series contains 3′ deletions of the *A1* promoter attached to the minimal −89 *Adh1* core promoter. The trimer series contains oligonucleotide oligotrimers of portions of the *A1* promoter in their normal orientation upstream of the minimal −73 CaMV 35S promoter. −73 CaMV, cauliflower mosaic virus 35S −73 5′ end of the minimal promoter; Bz1 int., *Bz1* intron; Adh1 int., *Adh1* intron 1; Nos 3′, nopaline synthase 3′ polyadenylation region; A1 3′, *A1* polyadenylation and 3′ region.

(Roth et al., 1991). This 32-bp region contains sequences with homologies to the consensus DNA binding sites of the bHLH and MYB proteins (Sen and Baltimore, 1986; Biedenkapp et al., 1988). Thus, the homology of the B and C1 DNA binding domains with bHLH and MYB proteins, respectively, and the presence of *bHLH* (CANNTG) and *myb* (C/TAACG/TG) consensus DNA binding sites in the *Bz1* promoter appeared to explain how these regulatory proteins interacted with the *Bz1* promoter. However, homologies to this area of the *Bz1* promoter were not readily apparent in the promoters of other coordinately regulated anthocyanin genes, such as the *A1* or *Bz2* promoters.

In this study, we sought to define the regions of the *A1* promoter responding to B and C1 in the anthocyanin pathway and responding to P in the phlobaphene pathway. Plasmid constructs containing various deletion or substitution derivatives and small oligonucleotide trimers of the *A1* promoter were used in a transient expression assay system to define the critical regions required for B/C1 or P inducibility. Two regions containing *myb*-like sites were found to be important. The 5′ site was more important for P induction and contained a *P*-like site, whereas both sites were important for B/C1 induction of the *A1* promoter. Adjacent to these two *myb*-like sites was an area critical for B/C1 transactivation. When aligned with the critical region of the *Bz1* promoter, a 15-bp region of the *A1* promoter was identical at 11 nucleotides. A 14-nucleotide consensus sequence of this region can also be found in the *C2*, *A2*, and *Bz2* anthocyanin structural gene promoters. These results, together with earlier results on the functional domains of the regulatory proteins, are used as the basis of a model for how the flavonoid regulatory proteins interact with the promoters of the structural genes.

RESULTS

Expression of 5′ Deleted *A1* Promoter Constructs

The sequence of the *A1* promoter from −130 to +1 contains several possible *myb*-like sites and two *P*-like sites (Figure 1A). The P protein has been demonstrated to bind to the P site at −60 (Grotewold et al., 1994). To define the 5′ boundary of the region important for *A1* promoter activity, a series of 5′ deletions of the 1.4-kb *A1* promoter (Schwarz-Sommer et al., 1987; Klein et al., 1989) was constructed. Each construct derived from A1-Luc (Figure 1B) contained a 5′ deleted *A1* promoter fragment (5′ end points from −1400 to +10), the first intron from the *Bz1* gene to increase the level of expression (Callis et al., 1987), the coding region from the firefly luciferase (*Luc*) gene (Ow et al., 1987), and the 3′ polyadenylation region from *A1*.

LUC activity from the constructs was measured in transient gene expression assays following the introduction of the plasmids into Black Mexican Sweet (BMS) maize suspension cells by microprojectile bombardment. The BMS cells require the expression of both B and C1 to transactivate the *A1* or *Bz1*

promoters and to synthesize anthocyanin pigments. The *A1* promoter can also be transactivated by P alone (Grotewold et al., 1994). Thus, expression from the A1-Luc constructs was transactivated by cobombardment with either B and C1 plasmids or a P plasmid. A plasmid containing the chloramphenicol acetyltransferase (*CAT*) gene driven by the maize *alcohol dehydrogenase 1* (*Adh1*) promoter (pAdh-CAT; Callis et al., 1987) served as an internal control for the efficiency of the bombardment and extraction (Klein et al., 1989). Expression from the A1-Luc plasmids was recorded as a ratio of LUC-to-CAT activity.

The activity of the 5′ promoter deletions relative to the activity of the full-length 1.4-kb *A1* promoter is exhibited in Figure 2. Relative activities from the constructs with 5′ deletion end points between −600 and −123 were essentially equal to the full-length promoter. For the constructs with end points at −98 and −91, relative activities of 5 and 8% were observed. The relative activities of constructs with end points at −78, −34, and +10 were 2% or less. None of the 5′ deletions or the intact *A1* promoter was transactivated by either B or C1 alone (data not shown) or expressed at significant levels in the absence of both B and C1 (Figure 2). These results suggest that the regions essential for the regulation of *A1* by B/C1 and by P lie within 123 bp of the start of transcription.

The binding of *Escherichia coli*–expressed C1 protein to the 3′ end of the *A1* gene has been reported previously (Wienand et al., 1990). To exclude the possibility that A1-Luc expression was influenced by sequences from the 3′ end of *A1*, this region of A1-Luc was replaced with a fragment containing the 3′ polyadenylation region from the nopaline synthase gene. When cobombarded with B/C1 plasmids, LUC activity from this construct was equivalent to that from the plasmid with the *A1* 3′ end (data not shown). This result indicates that the 3′ end of the *A1* gene does not contain an enhancer necessary for promoter activity.

Expression of 3′ Deleted A1 Promoter Constructs

To delineate the 3′ boundary of the regions responding to B/C1 and P regulation, it was necessary to distinguish between minimal core promoter sequences and B/C1- or P-responsive sequences. This was achieved by fusing 3′ deletions of the *A1* promoter to a minimal *Adh1* promoter (5′ end located at *Adh1* nucleotide position −89; Figure 1B). Attaching the B/C1- and P-responsive regions of the *A1* promoter upstream of the *Adh1* core promoter should confer B/C1 and P inducibility to the *Adh1* promoter because the *Adh1* core promoter was shown previously to be inducible by B/C1 when fused to the *Bz1* promoter (Roth et al., 1991).

The 3′ *A1* promoter deletion constructs contained a fixed 5′ end point at −1400 with 3′ deletion end points at −30, −55, −88, −108, −118, and −128. The *A1* 3′ deletion–*Adh1* promoter constructs were cobombarded into BMS cells, along with B and C1 or P expression vectors and pAdh-CAT. The LUC activity levels were measured relative to those of the *Adh1* core promoter (Figure 3). In the absence of B/C1 or P, all of the 3′ deletion constructs expressed at the same level as the *Adh1* core promoter. This indicates that there is no enhancement or repression of the A1-Adh1 core promoter fusions in the absence of B/C1 or P. However, in the presence of B/C1 or P, the *A1* promoter fragments with 3′ boundaries at −30, −55, and −88 resulted in a five- to 10-fold increase in LUC activity of the *Adh1* core promoter. The absolute amounts of LUC produced from these constructs were as high or higher than those produced from the intact *A1* promoter in the A1-Luc construct. This indicates that the *A1* regions of the fusion constructs are providing a strong enhancement of the minimal *Adh1* core promoter.

Only a twofold increase in expression was observed for the A1-Adh1 constructs with 3′ end points at −108 and −118. Both the −128 end point and the minimal *Adh1* core promoter alone showed no induction with either B/C1 or P. These data define the critical 3′ boundary of the B/C1- and P-responsive sequences of the *A1* promoter to be at −88. For comparison, the inducibility was measured from the previously reported construct containing *Bz1* sequences (−515 to −45) upstream of the *Adh1* minimal core promoter (Roth et al., 1991). The *Bz1* sequences increased the *Adh1* minimal promoter activity approximately sevenfold when cobombarded with B and C1 (Figure 3). The coexpression of P did not increase the Bz1-

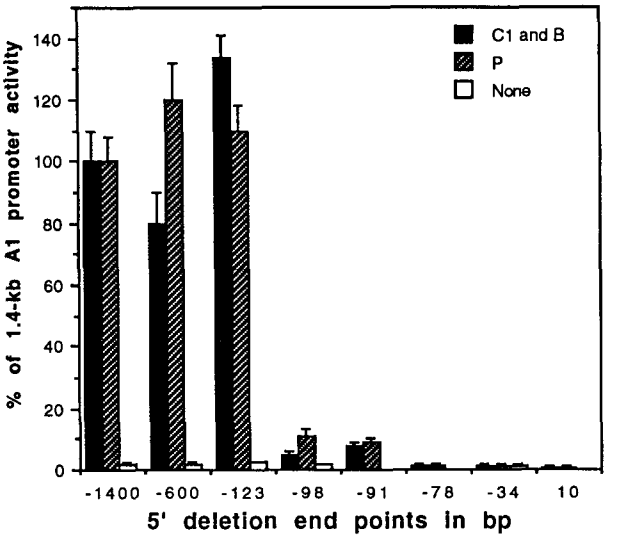

Figure 2. Transient Expression of 5′ Deletion Mutants of the 1.4-kb *A1* Promoter.

Constructs with 5′ promoter deletions of A1-Luc were bombarded into BMS-EG5 cell suspensions along with plasmids expressing the B/C1 or P cDNAs and the Adh-CAT reference plasmid. The bars represent the relative LUC activity versus the amount from the intact 1.4-kb promoter in A1-Luc with B/C1 (black columns), P (striped columns), or no transacting factor present (white columns; values not determined for the −91, −78, and +10 deletions). At least five independent bombardments were performed for each construct. Error bars represent the standard error of the mean.

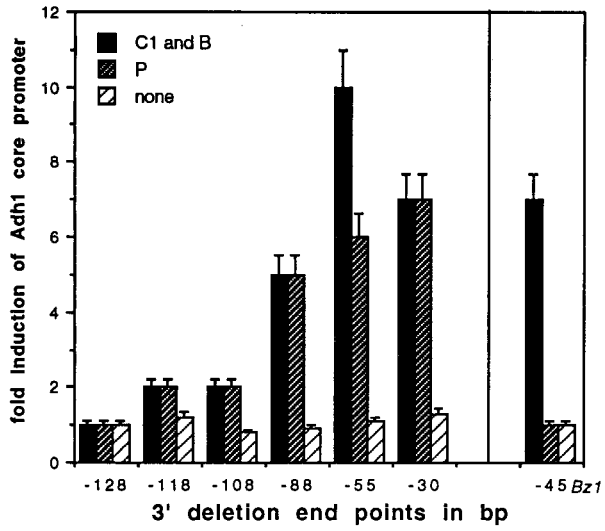

Figure 3. Enhancer Assay of 3′ Deletions of the *A1* Promoter Fused 5′ of the *Adh1* Core Promoter.

The indicated *A1* promoter 3′ deletion end points were fused to the −89 *Adh1* minimal promoter and bombarded into BMS-EG5 cell suspensions. The relative enhancement of activity was measured when cell suspensions were bombarded alone or with B/C1 or P. Shown for comparison is the value for the fusion of the 3′ end of the *Bz1* promoter at −45 (−45 *Bz1*) to the −89 *Adh1* minimal promoter. These values were compared to the expression from the −89 *Adh1* minimal promoter. The bars represent the activities with B/C1 (black columns), P (gray striped columns), or no transacting factor present (white-striped columns). Error bars represent the standard error of the mean.

Adh1 minimal promoter expression (Figure 3). This was expected because the *Bz1* promoter is not induced by P (Grotewold et al., 1994).

Site-Directed Mutations of the *A1* Promoter

The results presented above indicate that the sequences responsible for the transactivation by B/C1 or P are located between −88 and −123 of the *A1* promoter. To further define specific sequences necessary for regulation by B/C1 and P, linker-scanner mutations were constructed in the −128 to −59 region. Each mutation replaced 10 nucleotides of the *A1* promoter with 10 nucleotides that included an XhoI site (Figure 4). Only a 10-bp region of the entire 1.4-kb *A1* promoter sequence is changed in each of these constructs. The spacing of mutated sequences relative to the start of transcription was not changed. The mutant promoter plasmids were introduced into maize suspension cells with either B and C1 plasmids, or with the P plasmid, or without transacting factor plasmids. Figure 4 shows the activities of the mutant constructs from −123 to −79, relative to the unaltered 1.4-kb promoter in A1-Luc. The activities of some linker-scanner mutations were also measured in embryos and aleurones (Figure 4).

In BMS cells, the linker-scanner mutations with changes spanning −88/79, −78/69 (data not shown), and −68/59 (data not shown) resulted in promoter activities between 98 and 117%. This indicates that the −88 to −59 region contains no essential sequences for *A1* promoter activity. However, any of the four site-directed mutations in the −128/−89 region reduced

	-128	-118	-108	-98	-88	-79
A1	TATGCCGGTA	GTTGCAGCGT	GTGGTGTTGA	ATGGAGGATG	CGCTCAATCG	
LS oligo	ccctCgaGgg	ccctCgagGg	ccctcGagGg	ccctcGaggG	CcCTCgAggG	
	-128/119	-118/109	-108/99	-98/89	-88/79	
Factors						
BMS						
B and C1	60 ± 6	40 ± 3	26 ± 3	8 ± 2	111 ± 10	
P	15 ± 3	12 ± 2	32 ± 4	30 ± 3	112 ± 12	
None	2 ± 1	1.7 ± 1	1.8 ± 1	2 ± 1.5	2.2 ± 1	
embryos						
C1, R^{scm2}		16 ± 2	28 ± 3	4 ± 2	80 ± 7	
aleurones						
C1, R		21 ± 3	30 ± 4	3 ± 1	70 ± 8	

Figure 4. Linker-Scanner Mutations of the *A1* Promoter in the −128 to −78 Region.

XhoI linker-scanner mutations were made over 10-base regions in the 1.4-kb *A1* promoter of A1-Luc and bombarded into either BMS cells, embryos, or aleurone cells. *B/C1*, *P*, or no transacting factor constructs were cobombarded in BMS cells. Endogenous genes supplied C1 and R factors in the aleurone (*a1 C1 R*) or embryo (*bz1 C1 R^{scm2}*) tissues. Nucleotides that were changed in the linker-scanner mutants (LS oligo) relative to the *A1* promoter sequences are shown in lowercase letters, and conserved bases are shown in uppercase letters. Potential *myb*-like consensus DNA binding sites (CAAC is on the opposite strand) are underlined. The level of LUC activity relative to that of the intact promoter in A1-Luc is shown below the corresponding mutated regions.

both B/C1- and P-mediated transactivation to some extent. The overall expression from the promoter constructs mutated in this region was never higher than 60% for B/C1-mediated transactivation and 32% for P-mediated transactivation. The changes in the −128/119 and −118/109 region reduced P transactivation most severely to 15 and 12% of normal. The mutations in the −98/89 region reduced B/C1 transactivation most severely to 3 to 8% of normal (Figure 4). These results, in addition to the 5′ deletion at −123, indicate that the sequences between −123 and −109 are most important for P transactivation, whereas sequences between −98 and −89 are most important for B/C1 transactivation.

Four of the linker-scanner mutations were also assayed in aleurones and embryos expressing the *R* and *C1* genes (Figure 4). A self-colored allele of *R* (*R*^{scm2}) is needed to obtain anthocyanin expression in the embryo because the standard *R* allele is not expressed in embryos. This experiment measures the effect of the promoter mutations in a cell expressing the R and C1 proteins from the endogenous genes at normal levels. The embryo and aleurone results gave lower but generally similar ratios of mutant to wild-type promoter expression as was observed in BMS cells (Figure 4). One major difference was that the −118/109 site was more important relative to the −108/99 site in aleurones and embryos than in BMS cells for *B/C1* transactivation. These results generally confirm the importance and magnitude of these sequence changes on promoter transactivation in the presence of endogenous levels of R and C1 proteins.

Oligonucleotide Trimer Fusions to the CaMV 35S Core Promoter

To further confirm regions of the *A1* promoter that are sufficient to mediate transactivation by B/C1 and P, small oligonucleotides were tested for their ability to induce a minimal CaMV 35S promoter. These oligonucleotide trimers were inserted as direct repeats (in their normal orientation) upstream of a CaMV 35S minimal promoter with a 5′ end at −73, as shown in Figure 1B. The first oligonucleotide spans nucleotides −140 to −79 (pA1TRI-140/79) and contains the regions determined by linker-scanning to be necessary for both B/C1 and P inducibility. The second oligonucleotide includes nucleotides −134 to −100 (pA1TRI-134/100) and lacks the regions thought to be most important for *B/C1* inducibility (−98 to −88). A similar construct (pBz1TRI-78/47), containing the *Bz1* oligonucleotide trimer, previously shown to be sufficient for B/C1 transactivation (Roth et al., 1991), was assayed for comparison.

Figure 5 shows that the pA1TRI-134/100 oligonucleotide trimer enhanced CaMV 35S core promoter expression ~15-fold relative to the core promoter alone when cobombarded with either B/C1 or P expression plasmids. The longer oligonucleotide trimer construct pA1TRI-140/79 increased expression from the minimal promoter 200-fold when bombarded with B/C1 and 50-fold when bombarded with P. Expression from the *Bz1*

Trimer	Fold induction of CaMV 35S core promoter		
	B/C1	P	none
pBz1TRI-78/47	300 ± 11	1 ± 0.6	1 ± 0.5
pA1TRI-140/79	200 ± 12	50 ± 5	1 ± 0.3
pA1TRI-134/100	15 ± 2	15 ± 3	1 ± 0.4
pSG2-2	1 ± 0.5	1 ± 0.5	1

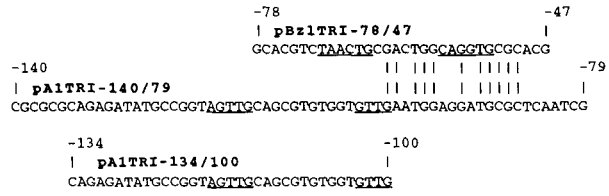

Figure 5. Enhancer Assay of Small *A1* Oligonucleotide Trimers.

The monomer sequences of the *A1* and *Bz1* oligonucleotide regions tested as trimers upstream of the minimal CaMV 35S construct (−73 5′ end) are aligned by the homology area shown by the vertical lines. Each oligonucleotide trimer clone was bombarded at least five times into BMS-EG5 cell suspensions, with or without transactivator genes. All values are shown as the fold induction of LUC activity relative to the minimal −73 CaMV 35S pSG2-2 construct in the absence of B/C1 or P factors. Potential *myb*-like sequences in the *A1* oligonucleotide are underlined and are on the opposite strand relative to the *myb*-like site underlined in the *Bz1* sequence (C/TAACG/TG). Also underlined is the *bHLH*-like site in *Bz1* (CAGGTG).

oligonucleotide trimer construct pBz1TRI-78/47 was increased 300-fold by cobombardment with B/C1 but was unaffected by P.

Both *A1* trimer constructs were tested in bombardment experiments with either C1 or B alone. Introduction of either regulatory gene alone led to only a two- to fivefold enhancement of LUC activity relative to bombardments without the activator plasmids (data not shown). These oligonucleotide trimer experiments demonstrated that the −140 to −79 region of the *A1* promoter appears to contain the most important regions needed for B/C1 and P transactivation. The sequences most important for P-mediated activation are in the −100 to −134 region, with increased stimulation when the −79 to −99 region was included. The −134 to −100 region conferred some B/C1 inducibility, but the inducibility was 12-fold stronger when the −99 to −79 region was also included. The ability of both *A1* trimers to stimulate expression to different levels suggests the presence of more than a single area of contact of the regulatory proteins in this region of the promoter.

DISCUSSION

Mutational analysis of the *A1* promoter has established that the −123 to −88 region is critical for the activation of the *A1* gene by the anthocyanin and phlobaphene regulatory genes. An oligonucleotide trimer from −140 to −79 confers efficient B/C1- or P-dependent activation to a minimal CaMV 35S promoter. The location of this critical region between −123 and −88 is consistent with the varying levels of anthocyanin

biosynthesis observed in revertants of the *a1-m2* mutant allele carrying a defective Suppressor-mutator (*dSpm*) element at −99 (Masson et al., 1987). Presumably, the imprecise excisions of the *dSpm* element account for the reduced and variable expression observed in the revertants. The *A1* promoter does not appear to be negatively regulated because in the absence of B/C1 or P all of the mutant promoters are expressed at the same level as the intact *A1* promoter, which is 50-fold less than the induced levels.

The −123 to −88 region of the *A1* promoter contains two *myb*-like consensus DNA binding sites but does not contain *bHLH*-like consensus DNA binding sites. Previously, the presence of *myb*-like and *bHLH*-like consensus DNA binding sites in the critical region of the *Bz1* promoter was consistent with a model of B and C1 interactions with the promoter based on their homologies to the DNA binding regions of the bHLH and MYB proteins. This model is still likely for the *A1* promoter's interactions with C1, PI, or P at the *myb*-like consensus sites spanned by the site-directed mutations at −118/−109 and −108/−99 and the P binding site at −65/−55 (Grotewold et al., 1994). Mutations at any of these sites, particularly the upstream sites, affect both P and B/C1 induction of the *A1* promoter.

The P protein was observed to bind in vitro only to the −55 to −65 region and not to the region between −123 to −88 (Grotewold et al., 1994). However, a site-directed *A1* promoter mutation that abolished P binding in vitro still retained 40 and 75% of wild-type *A1* promoter activity in the presence of P or R/C1, respectively (Grotewold et al., 1994). An indication that the −55 to −65 site is not sufficient for *A1* promoter induction is that 5′ deletion mutants lose 90% of activity at nucleotide −98 and virtually all activity at nucleotide −78. The −55 to −65 site is not required because the −88 3′ *A1* promoter deletion–*Adh1* core promoter construct removes this region and retains 70% of the P- or B/C1-dependent enhancer function. Similarly, the region between −65 and −55 is not present in the inducible oligonucleotide trimer constructs (this study) but does confer some P or B/C1 inducibility as an oligonucleotide trimer (Grotewold et al., 1994). Thus, whereas P binds to the region between −55 and −65 in vitro, it is clear that it must also act through the more important −123 to −88 region in vivo. The lack of P binding to the −123 to −88 region in vitro may be a result of the absence of other required factors or protein modifications not present in the P protein produced in *E. coli*. It is worth noting that a *P*-like site (Grotewold et al., 1994) of the sequence AACTACC at nucleotides −123 to −116 overlaps with the *myb*-like site that we have identified at nucleotides −119 to −115 (Figure 1A).

The critical −123 to −88 region of the *A1* promoter does not contain a *bHLH*-like consensus DNA binding site (CANNTG). This result is consistent with the recent observation that the bHLH region of the B protein is not essential for B transactivation of anthocyanin structural gene promoters. A deleted B protein missing the bHLH domain was still able to transactivate the *Bz1* promoter to 50% (Goff et al., 1992) and the *A1* promoter to 70% (J.A. Tuerck, unpublished results) of the level observed with intact B protein. Thus, the original proposal, that the B/R proteins bind to *bHLH*-like consensus sites in anthocyanin promoters through their bHLH domain, is not valid. This is consistent with the lack of well-defined *bHLH*-like consensus DNA binding sites in the other anthocyanin promoters. However, these *bHLH*-like sites and domains may play a role in some of the fine tuning of the individual genes like *Bz1* under some of the diverse light, stress, and tissue-specific regulation that occurs in anthocyanin production (Taylor and Briggs, 1990).

Although the *bHLH*-like consensus DNA binding site is not conserved, alignment of this region of the *Bz1* promoter and the critical region of the *A1* promoter reveals an 11-bp of 15-bp match (Figure 6). This area (−100 to −86) in the *A1* promoter produced the most severe reduction in *A1* expression when changed by a linker-scanning mutation (3 to 8% residual activity with B/C1). Linker-scanner mutations in the *Bz1* homologous region (−64 to −50) reduced expression to 1 to 3% of normal (Roth et al., 1991). Thus, the sequences most important for B/C1 inducibility show a significant homology in the two promoters. However, the nonidentical nucleotides create a *bHLH*-like consensus binding site in the *Bz1* promoter that is not present in the *A1* promoter (Figure 6). The severe effect of the linker-scanning mutation in the 10-bp region of the *Bz1* promoter that contains the bHLH consensus binding site (−53 to −58; Roth et al., 1991) is now attributed to the mutations in the nucleotides conserved between the *A1* and *Bz1* promoters and not to the concurrent changes in the bHLH consensus binding site.

Inspection of the *A2*, *Bz2*, and *C2* promoters revealed that the sequences required for B/C1 inducibility of the *A1* and *Bz1* promoters have homology with sequences in these promoters (Figure 6). The five anthocyanin promoters match a core

```
                                    myb          bHLH
                                 C/TAACG/TG      CANNTG
Bz1            -78 GCACGTCTAACTGCGACTGGCAGGTGCGC -50
                                    || |||  |  |||||
A1 -123 CGGTAGTTGCAGCGTGTGGTGTTGAATGGAGGATGCGC -86

                                    35535454 43433
consensus                           CGACTGGCNGGTGC

Bz1                          -65    CGACTGGCaGGTGC
A1                          -101    tGAaTGGagGaTGC
Bz2                          -66    CGACTcGCcaGTcg
A2                           -88    CGAaTGGCaGGcag
C2                          -262    tGACTGGCtGcTGC
```

Figure 6. Consensus Sequence for Anthocyanin Promoters.

The consensus sequence was derived from the regions of the five anthocyanin promoters shown. The regions were identified by homology to the aligned regions (vertical lines) of the *A1* and *Bz1* promoters that were found to be critical for *B/C1* inducibility in vivo. The nucleotide numbers are relative to the start of transcription. Nucleotides matching the consensus sequence are represented as uppercase letters. The *myb*-like and *bHLH*-like consensus DNA binding sequences are shown above the *Bz1* oligonucleotide. Potential *myb*-like sequences in the *A1* oligonucleotide are underlined and are on the opposite strand relative to the *myb*-like site in the *Bz1* sequence.

consensus sequence (CGACTGGCNGGTGC) in at least nine of the 14 nucleotides. Four of the positions in the consensus sequence are conserved, and four more match four of the five promoter sequences. The consensus sequence is located within 100 bp of the start of transcription in four of the five promoters, with the exception of the *C2* promoter, in which the consensus sequence starts at −263. In addition, the *A1*, *Bz1*, *Bz2*, and *A2* promoters have a *myb*-like consensus sequence located within 32 nucleotides upstream of the location of their 14-bp consensus sequence (at −98 in *Bz2* and at −105 in *A2*).

The question of how the anthocyanin and phlobaphene transcription factors interact with the structural gene promoters remains. There are limited data on the in vitro binding of B and C1 (Wienand et al., 1990). P binding has not been observed in vitro at the −123 to −88 region (Grotewold et al., 1994), although our results indicate that this region is essential for *A1* promoter activity in vivo. However, some properties of these regulatory proteins are known from protein domain substitution studies. The C1 acidic region forms a transacting factor that functions independently of B when attached to the GAL4 DNA binding domain (Goff et al., 1991) or the P DNA binding domain (Grotewold et al., 1994). However, the C1 DNA binding domain, which is 80% identical to that of the P protein, still required the B protein for transactivation of the *Bz1* promoter when attached to the GAL4 acidic region (Goff et al., 1991). The dependence of C1 on B is likely to be mediated by a direct interaction because C1 and B have been shown to interact in vivo in protein tethering experiments (Goff et al., 1992). In addition, the B protein did not contain any transactivating regions when fused to the GAL4 DNA binding domain (Goff et al., 1992).

A possible model based on the above promoter and protein domain data is that C1 contains a functional acidic transactivating region but binds *myb*-like consensus sites inefficiently. The B protein contains no transactivating domain but can bind to C1 and to DNA at the 14-bp consensus sequence (CGACTGGCNGGTGC), with a DNA binding domain other than the nonessential bHLH domain. (If B does not bind at the 14-bp consensus site, then a constitutive factor must be binding at the consensus site.) The B/C1 complex binds to the promoter DNA with higher affinity and specificity than C1 alone at *myb*-like sites and optimally with B contacting the anthocyanin 14-bp consensus site. In addition, the oligonucleotide trimer results indicate that B/C1 complex can bind DNA with lower efficiency at *myb*-like sites in the absence of the 14-bp consensus sequence. The dependence of C1 on B allows C1 to recognize *myb*-like sites that are different from the ones to which P binds.

In this model, the P protein, which is 80% identical to C1 in the MYB-homologous N-terminal region, has a higher affinity and specificity than C1 for binding *myb*-like sites (or the overlapping *P*-like consensus sites at −60 and −119; Figure 1A). Thus, the P protein can bind and transactivate promoters in the phlobaphene pathway without interaction with a B protein. The greater specificity of the P protein also allows it to distinguish between phlobaphene *myb*-like binding sites, presumably *P*-like sites such as those in the *A1* promoter, and those unique to the anthocyanin pathway, such as those in the *Bz1* promoter. It is of course possible that both B/C1 and P depend on a yet to be identified factor, which interacts with the conserved 14-bp motif (CGACTGGCNGGTGC), for productive binding and transactivation events. However, this factor must be constitutive because most plant cells produce anthocyanin when B and C1 are introduced into the cells transiently. Similarly, the transactivation of the *A1* promoter in BMS cells requires only the addition of B/C1 or P.

Further studies of these promoters and other anthocyanin and phlobaphene structural gene promoters such as *Bronze-2* (Bodeau and Walbot, 1992) and their binding proteins will be needed to fully resolve these complex interactions. This study defines the critical regions necessary for *A1* promoter activity in vivo, which should help develop accurate in vitro binding systems. These results also provide a basis for determining consensus sites that the B/C1 and P proteins act through to regulate the various structural genes in the anthocyanin and phlobaphene biosynthetic pathways.

METHODS

Plasmids

Plasmids used in this study that have been previously described are as follows: p35S-C1 and p35S-B-Peru, (Goff et al., 1990); pAdh-CAT, (previously described as pAl1CN, Callis et al., 1987); p35S-P cDNA1 (Grotewold et al., 1991); pBz1-Luc, (Klein et al., 1989). A1-Luc (Klein et al., 1989) was modified to contain NcoI, SacII, and HindIII sites 5′ of the *A1* promoter. pSG2-2 contains NcoI, SacII, ApaI, AvrII, NheI, SpeI, XhoI, and HindIII sites upstream of the cauliflower mosaic virus (CaMV) 35S minimal promoter (Roth et al., 1991). pSG2-7 contains NcoI, SacII, ApaI, AvrII, NheI, SpeI, XhoI, and HindIII sites upstream of the −89 *alcohol dehydrogenase 1 (Adh1)* core promoter (Roth et al., 1991).

5′ Deletions of the 1.4-kb *A1* (Klein et al., 1989) promoter fragment were produced by ligating XhoI linkers to blunt-ended Bal31 exonuclease digestion products of A1-Luc DNA. 3′ Deletion products were produced by the polymerase chain reaction with an XhoI site at the deletion boundary. Site-directed mutagenesis (Kunkel et al., 1987) of the *A1* promoter replaced 10 nucleotides of the *A1* promoter with an XhoI site (CCCTCGAGGG) in each mutant. Two oligonucleotide sequences, a 67-mer oligonucleotide (GATCTCGCGCGCAGAGATATGCCGGTAGTTGCAGCGTGTGGTGTTGAATGGAGGATGCGCTCAATCG), containing the *A1* promoter sequences from −140 to −79 and a 41-mer (GATCTCAGAGATATGCCGGTAGTTGCAGCGTGTGGTGTTGG) from −134 to −100, both flanked by BglII and BamHI sites, were trimerized in direct repeats and cloned into the BamHI site of pIC19H+ (Marsh et al., 1984). The XhoI- and SalI-digested trimer DNA was ligated upstream of the −73 CaMV 35S core promoter of pSG2-2, with the sequences in their normal orientation.

Plant Material and Cell Bombardment

The Black Mexican Sweet maize suspension line (BMS-EG5, Chourey and Zurawski, 1981) used was able to produce anthocyanin pigments

after bombardment with *B/C1* expression plasmids. For the biolistic bombardment of embryogenic callus or BMS suspension cells, 3 μg of the experimental plasmid, 3 μg of either p35S-C1 and p35S-B-Peru or p35S-P, and 2 μg of pAdh-CAT DNA were precipitated and delivered to the cells as previously described (Klein et al., 1989; Goff et al., 1991). For bombardment of *bz1,C1,R^{scm2}* embryos (Roth et al., 1991) or *a1,C1,R* aleurones (Klein et al., 1989), 6 μg of the experimental plasmid and 2 μg of pAdh-CAT were coprecipitated onto tungsten particles.

Each construct was bombarded in at least five independent bombardments into BMS-EG5 cells, or in at least three independent experiments into embryos or aleurones. One independent experiment comprised three replica bombardments. The mean of those experiments and the standard error of the mean were calculated and are presented in the figures.

Enzyme Assays

LUC was assayed in cell extracts (Callis et al., 1987), and its activity was expressed as the number of light units detected in 10 sec at 25°C. Chloramphenicol acetyltransferase (CAT) activity was determined by the amount of counts per minute of ^{14}C-labeled acetyl coenzyme A converted to ^{14}C-acetoxy-chloramphenicol in 1 hr at 37°C (Sleigh, 1986). The LUC-to-CAT ratio was calculated as previously described (Roth et al., 1991).

ACKNOWLEDGMENTS

We thank Tom Peterson for the P cDNA clone, and Doug Russell, Tim Conner, and Virginia Peschke for critical reading of the manuscript.

Received May 18, 1994; accepted September 21, 1994.

REFERENCES

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 *C1* and *R* gene products. Mol. Gen. Genet. **233,** 379–387.

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.

Chourey, P.S., and Zurawski, D.B. (1981). Callus formation from protoplasts of maize cell culture. Theor. Appl. Genet. **59,** 341–344.

Coe, E.H., Jr., Neuffer, M.G., and Hoisington, D.A. (1988). The genetics of corn. In Corn and Corn Improvements, G.F. Sprague and J. Dudley, eds (Madison, WI: American Agronomy Society), pp. 81–258.

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.

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

Furtek, D., Schiefelbein, J., Johnston, F., and Nelson, O.E. (1988). Sequence comparisons of three wild-type *Bronze-1* alleles from *Zea mays*. Plant Mol. Biol. **11,** 473–481.

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., and Peterson, T. (1994). Isolation and characterization of a maize gene encoding chalcone flavanone isomerase. Mol. Gen. Genet. **242,** 1–8.

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

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

Kunkel, T.A., Roberts, J.D., and Zakour, R.A. (1987). Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154,** 367–382.

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 a transcriptional activator and contains the *bHLH*-homology region. Proc. Natl. Acad. Sci. USA **86,** 7092–7096.

Marsh, J.L., Erfle, M., and Wykes, E.J. (1984). The pIC plasmid and phage vectors with versatile cloning sites for recombinant selection by insertional inactivation. Gene **32,** 481–485.

Masson, P., Surosky, R., Kingsbury, A., and Fedoroff, N.V. (1987). Genetic and molecular analysis of the Spm-dependent *a-m2* alleles of the maize a locus. Genetics **177,** 117–137.

Menssen, A., Höhmann, S., Martin, W., Schnable, P.S., Peterson, P.A., Saedler, H., and Gierl, A. (1990). The *En/Spm* transposable element of *Zea mays* contains splice sites at the termini generating a novel intron from a *dSpm* element in the *A2* gene. EMBO J. **8,** 3051–3057.

Nash, J., Luehrsen, K.R., and Walbot, V. (1990). *Bronze-2* gene of maize: Reconstruction of a wild-type allele and analysis of transcription and splicing. Plant Cell **2,** 1039–1049.

Ow, D.W., Jacobs, J.D., and Howell, S.H. (1987). Functional regions of the cauliflower mosaic virus promoter determined by use of the

firefly luciferase gene as a reporter of promoter activity. Proc. Natl. Acad. Sci. USA **84,** 4870–4874.

Pas-Ares, J., Ghosal, D., Wienand, U., Peterson, P., 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.

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 *bHLH* binding sites. Plant Cell **3,** 317–325.

Schwarz-Sommer, Z., Shepard, N.S., Tacke, E., Gierl, A., Rohde, W., Leclerq, L., Mattes, M., Berndtgen, R., Peterson, 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.

Sen, R., and Baltimore, D. (1986). Multiple factors interact with the immunoglobulin enhancer sequences. Cell **46,** 705–716.

Sleigh, M.J. (1986). A non-chromatographic assay for expression of the chloramphenicol acetyltransferase gene in eucaryotic cells. Anal. Biochem. **156,** 251–256.

Styles, E.D. (1970). Functionally duplicate genes conditioning anthocyanin formation in maize. Can. J. Genet. Cytol. **12,** 397.

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

Taylor, L.P., and Briggs, W.R. (1990). Genetic regulation and photocontrol of anthocyanin accumulation in maize seedlings. Plant Cell **2,** 115–127.

Wienand, U., Weydemann, U., Niesbach-Klösgen, U., Peterson, P.A., and Saedler, H. (1986). Molecular cloning of the *c2* locus of *Zea mays*, the gene coding for chalcone synthase. Mol. Gen. Genet. **203,** 202–207.

Wienand, U., Paz-Ares, J., Scheffler, B., and Saedler, H. (1990). Molecular analysis of gene regulation in the anthocyanin pathway of *Zea mays*. In Plant Gene Transfer, UCLA Symposia on Molecular and Cellular Biology, Vol. 129, C.L. Lamb and R.N. Beachy, eds (New York: Wiley-Liss), pp. 111–124.