The EMBO Journal vol.14 no.8 pp.1773–1784, 1995

Dual DNA binding specificity of a petal epidermis-specific MYB transcription factor (MYB.Ph3) from *Petunia hybrida*

Roberto Solano[1], Concepción Nieto, Julio Avila[2], Luis Cañas[3], Isabel Diaz[4] and Javier Paz-Ares[1,5]

Departamento de Biología Vegetal, Centro de Investigaciones Biológicas-CSIC, Velazquez 144, 28006-Madrid and [4]Departamento de Biotecnologia, Escuela Técnica Superior de Ingenieros Agrónomos-UPM, Ciudad Universitaria s/n, 28040-Madrid, Spain
[1]Present address: Departamento de Genética Molecular de Plantas, Centro Nacional de Biotecnología-CSIC, Campus de Cantoblanco, 28049-Madrid, Spain
[2]Present address: Departamento de Biología Vegetal, UDI Fisiología Vegetal, Universidad de la Laguna, Astrofísico Fco. Sánchez, 38206 Sta. Cruz de Tenerife, Spain
[3]Present address: Instituto de Biología Molecular y Celular de Plantas, UPV-CSIC Camino de Vera, 14, 46022-Valencia, Spain

[5]Corresponding author

R.Solano and C.Nieto contributed equally to this work

Communicated by F.García-Olmedo

The MYB.Ph3 protein recognized two DNA sequences that resemble the two known types of MYB DNA binding site: consensus I (MBSI), aaaAaaC(G/C)-GTTA, and consensus II (MBSII), aaaAGTTAGTTA. Optimal MBSI was recognized by animal c-MYB and not by Am305 from *Antirrhinum*, whereas MBSII showed the reverse behaviour. Different constraints on MYB.Ph3 binding to the two classes of sequences were demonstrated. DNA binding studies with mutated MBSI and MBSII and hydroxyl radical footprinting analysis, pointed to the N-terminal MYB repeat (R2) as the most involved in determining the dual DNA binding specificity of MYB.Ph3 and supported the idea that binding to MBSI and MBSII does not involve alternative orientations of the two repeats of MYB.Ph3. Minimal promoters containing either MBSI and MBSII were activated to the same extent by MYB.Ph3 in yeast, indicating that both types of binding site can be functionally equivalent. MYB.Ph3 binding sites are present in the promoter of flavonoid biosynthetic genes, such as the *Petunia chsJ* gene, which was transcriptionally activated by MYB.Ph3 in tobacco protoplasts. MYB.Ph3 was immunolocalized in the epidermal cell layer of petals, where flavonoid biosynthetic genes are actively expressed. This strongly suggests a role for MYB.Ph3 in the regulation of flavonoid biosynthesis.
Key words: DNA binding/flavonoids/MYB/*Petunia*/transcription factor

Introduction

Proteins related to the product of the avian myeloblastosis oncogen (v-*myb*) and its cellular counterpart (c-*myb*), MYB proteins, have been identified in different eukaryotes, ranging from yeast to humans (for reviews, see Lüscher and Eiseman, 1990; Graf, 1992). They share a conserved DNA binding (MYB) domain and are known to act as transcriptional regulators (Grotewold *et al.*, 1994; Sablowski *et al.*, 1994, and references therein). The numbers of MYB proteins which have been identified in different species have been extremely variable; thus, only one MYB protein has so far been described in *Drosophila* (Katzen *et al.*, 1985), while at least 20–40 have been detected in *Petunia hybrida* (Avila *et al.*, 1993). Quite different functions have been demonstrated for different members of this protein family, e.g. chicken c-MYB is involved in the control of hematopoiesis, yeast BAS1 controls histidine biosynthesis and maize C1 regulates the synthesis of the plant-specific phenylpropanoid pigments, anthocyanins (Cone *et al.*, 1986; Graf, 1992; Paz-Ares *et al.*, 1987; Tice-Baldwin *et al.*, 1989).

The above diversity of function does not appear to hold for highly similar MYB proteins from a given species and its relatives. In fact, different vertebrate c-MYB proteins regulate hematopoiesis in their respective organisms, and another vertebrate MYB protein, chicken B-MYB can interact with at least some targets of its c-MYB counterpart (Foos *et al.*, 1992; Graf, 1992). In plants, six MYB proteins [C1, Pl, P, Zm1 and Zm38 from maize and MYB305 (Am305) from *Antirrhinum*] have been implicated in the regulation of phenylpropanoid biosynthetic genes (Cone *et al.*, 1986, 1993; Franken *et al.*, 1994; Grotewold *et al.*, 1991; Paz-Ares *et al.*, 1987; Sablowski *et al.*, 1994). Phenylpropanoids are a class of metabolites originating from phenylalanine, which includes among others flavonoids, coumarins and cinnamyl alcohols, the latter being precursors of lignin and other cell wall phenolics (Hahlbrock and Scheel, 1989). In spite of their chemical diversity, the synthesis of these compounds does include common enzymatic steps. There are two other plant MYB proteins with known function, the Gl1 protein from *Arabidopsis*, which is involved in trichome differentiation, and the *Antirrhinum* MIXTA protein, which controls the shape of epidermal petal cells (Oppenheimer *et al.*, 1991; Noda *et al.*, 1994). Target genes for these proteins have not been molecularly identified although, in the case of MIXTA, its effect on cell shape correlates with changes in the cell wall, a structure which contains phenylpropanoid derivatives (Noda *et al.*, 1994).

The MYB domain most often contains three imperfect, 51/52 residue repeats (R1, R2 and R3), but in plants it contains only two (R2 and R3), or exceptionally one (Baranowskij *et al.*, 1994). A helix–turn–helix motif, characteristic of many prokaryotic and eukaryotic DNA binding proteins, can be predicted within each MYB repeat, which is consistent with the ability of the domain to bind DNA (Frampton *et al.*, 1991). The role of each repeat in DNA binding is, however, different. R1 of animal

MYB proteins has a limited participation, as it can be deleted without a significant effect on the sequence specificity of DNA binding properties of the resulting protein, although its absence decreases the stability of the protein–DNA complex (Howe *et al.*, 1990; Tanikawa *et al.*,1993). In c-MYB, the different roles of R2 and R3 in sequence-specific DNA binding is indicated by the different sequence requirements of these two repeats and by the fact that while R3 can form the helix–turn–helix structure in solution, as indicated by NMR analysis, R2 appears to depend on sequence-specific DNA binding to adopt/stabilize this conformation (Ogata *et al.*, 1992; Jamin *et al.*, 1993; Myrset *et al.*, 1993; Sarai *et al.*, 1993; Tanikawa *et al.*, 1993). An additional difference between R2 and R3 repeats is the differential effect of mutations at equivalent positions in these repeats (Frampton *et al.*, 1991).

Similarities between MYB proteins are mostly confined to the MYB domain, and amino acid residues corresponding to putative recognition helices are most highly conserved, which suggests a similarity of their binding sites (Avila *et al.*, 1993). Accordingly, all known animal MYB proteins, and at least two plant MYB proteins, ATMYB1 and ATMYB2 from *Arabidopsis*, bind to sequences containing the core CNGTT(A/G) (Lüscher and Eiseman, 1990; Urao *et al.*, 1993). In contrast, however, two other plant proteins, Am305 from *Antirrhinum* and P from maize, which are closely related to ATMYB2, recognize a different class of sequences: G(G/T)T(T/A)GGT(T/G) or GGT(T/A)GGT(T/G) (Grotewold *et al.*, 1994; Sablowski *et al.*, 1994).

Here we report on the DNA binding properties of protein MYB.Ph3, which recognizes two types of sequence: consensus MBSI, aaaAaaC(C/G)GTTA, and consensus MBSII, aaaAGTTAGTTA, respectively resembling the binding sites of animal c-MYB and of plant Am305 and P MYB proteins, and differentially bound by c-MYB and Am305 proteins. In addition, each class of binding site is shown to have different sequence and spacing constraints. We suggest that the R2 repeat is the most involved in determining the dual DNA binding specificity, and that binding of MYB.Ph3 to MBSI and MBSII does not involve alternative orientations of its two MYB repeats. We also show that MYB.Ph3 is able to activate transcription in yeast, from promoters containing either of the binding sites, and in plant protoplasts, from the promoter of the *chsJ* gene, which contains MYB.Ph3 binding sites. Finally, we show that MYB.Ph3 is synthesized in epidermal cells of *Petunia* petals, where phenylpropanoid (flavonoid) biosynthetic genes, such as *chsJ*, are also expressed. These results strongly suggest the involvement of MYB.Ph3 in the regulation of flavonoid biosynthetic genes.

Results

Production of full-length and mutant MYB proteins

To study the DNA binding properties of MYB.Ph3, appropriate parts of *myb.Ph3* cDNA, or of its mutant version, *myb.Ph3*(d→e), were placed under the control of the T7 promoter from plasmid pET3b (Studier *et al.*, 1990). A full size and several amino-terminal and carboxy-terminal deletion derivatives, as well as a D→E mutant

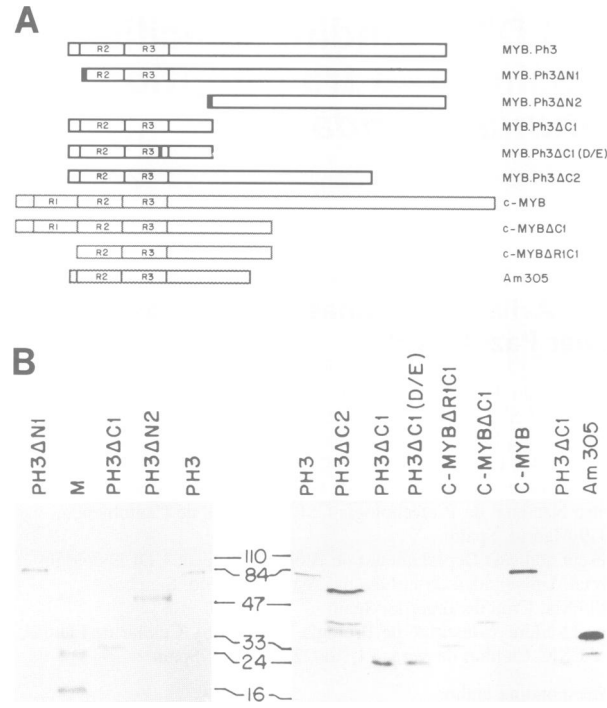

Fig. 1. MYB proteins and mutant derivatives expressed in *E.coli* and/or *in vitro*. (**A**) Schematic representation of the different proteins expressed. MYB.Ph3, full size MYB.Ph3 protein (Avila *et al.*, 1993); MYB.Ph3ΔN1 and MYB.Ph3ΔN2, amino-terminally truncated derivatives, lacking 32 and 180 amino acid residues, respectively; MYB.Ph3ΔC1 and MYB.Ph3ΔC2, carboxy-terminally truncated derivatives, lacking 337 and 107 amino acid residues, respectively; MYB.Ph3ΔC1(D→E), derivative of MYB.Ph3ΔC1, with a single mutation within the MYB domain (asp117→glu; Avila *et al.*, 1993); c-MYB, full size murine c-MYB protein (Howe *et al.*, 1990); c-MYBΔC1, carboxy-terminal truncation derivative, lacking 339 amino acid residues; c-MYBΔR1C1, amino-terminally truncated derivative of c-MYBΔC1, lacking the first MYB repeat R1; Am305, full size *Antirrhinum* MYB protein Am305 (Jacskon *et al.*, 1991). MYB repeats present in the different proteins are indicated (R1, R2 and R3). (**B**) SDS–PAGE analysis of the proteins expressed in *E.coli* (left) and *in vitro* (right). The *E.coli* synthesized proteins were detected by immunoblotting with MYB.Ph3 antibodies. The *in vitro* synthesized proteins were labelled with [^{35}S]methionine and detected by fluorography. The presence of several faster migrating bands, in some of the lanes, is attributed to premature termination of translation and/or proteolytic cleavage products. The numbers indicate the apparent molecular mass, in kDa, of pre-stained markers (M)

affecting a highly conserved D residue in the R3 repeat (D117, Avila *et al.*, 1993), were synthesized in *Escherichia coli* and/or *in vitro*, as summarized in Figure 1. As expected from earlier experience, apparent molecular masses of these proteins were larger than those deduced from their amino acid sequences (Avila *et al.*, 1993). Other proteins used in these studies were the c-MYB and two of its derivatives from mice and Am305 from *Antirrhinum* (Howe *et al.*, 1990; Jackson *et al.*, 1991; Figure 1). These proteins, except c-MYBΔR1C1, were synthesized by *in vitro* transcription and translation of previously described constructs (Howe *et al.*, 1990; Sablowski *et al.*, 1994). To synthesize c-MYBΔR1C1, a derivative of the c-MYBΔC1 construct was prepared by removal of the DNA region encoding the first repeat, using PCR.

Initial DNA binding studies were performed with proteins MYB.Ph3ΔC1 and MYB.Ph3ΔN2, because the

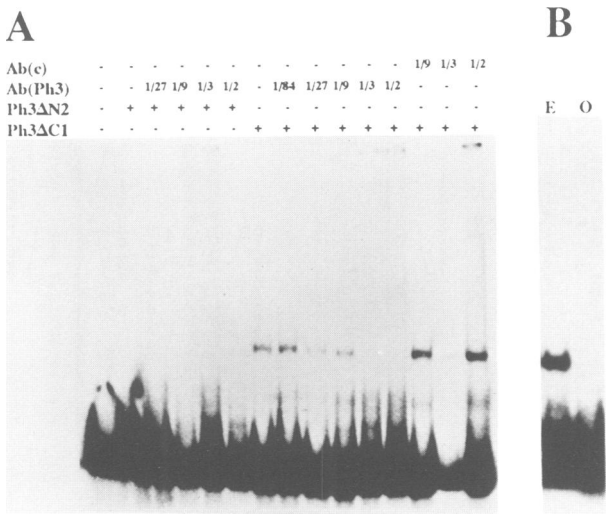

Fig. 2. DNA binding by MYB.Ph3ΔC1, as shown by EMSA. (**A**) Protein extracts of *E.coli* synthesizing MYB.Ph3ΔC1 or MYB.Ph3ΔN1 (1 μg) were incubated with a pool of random oligomers in the presence or absence of MYB.Ph3 antibodies, Ab(Ph3), or non-specific antibodies, Ab(C). Free and bound DNA were resolved by non-denaturing PAGE. Numbers indicate the dilution of the antibodies. Other symbols are: +, present and −, absent. (**B**) Binding of MYB.Ph3ΔC1 to the mixture of oligonucleotides after none (O) or after five rounds of binding site selection (E), using 0.1 μg of *E.coli* extracts.

E.coli-synthesized, full-length MYB.Ph3 protein was found to aggregate in inclusion bodies. This problem was subsequently avoided by using the *in vitro* synthesized protein.

Selection of MYB.Ph3 binding sites

To identify MYB.Ph3 binding sites, a random site selection strategy based on the electrophoretic mobility shift assay was used (EMSA; Blackwell and Weintraub, 1990). Extracts containing either MYB.Ph3ΔC1 or MYB.Ph3ΔN2 were incubated with a mixture of oligomers with a random 18mer core flanked by two constant regions. The DNA−protein complexes were resolved from the free DNA by EMSA, as shown in Figure 2A. A retarded band specific to extracts containing MYB.Ph3ΔC1 was clearly visible. The ability of MYB.Ph3 antibodies to specifically compete this protein−DNA complex demonstrated that MYB.Ph3ΔC1 was involved in its formation. DNA present in the complex with MYB.Ph3 was amplified by PCR and used in a subsequent round of selection. This process was repeated three more times. The only difference between rounds was that the protein concentration was progressively reduced, to increase stringency of selection. The final degree of enrichment was clearly significant, as evaluated by EMSA (Figure 2B). The selected oligonucleotides were cloned in pUC19, and inserts from independent clones were sequenced. An inspection of these sequences revealed the presence of two types of motifs, one with a CNGTTA core similar to that of animal MYB proteins (Lüscher and Eiseman, 1990) and another with the core GTTAGTTA, which resembles that of *Antirrhinum* Am305 and maize P proteins (Figure 3; Grotewold *et al.*, 1994; Sablowski *et al.*, 1994). Two consensus binding sites were derived from the sequences of inserts with single-core motifs: type I (MBSI), aaaAaaC-

Single CNGTTA

atccCCCTGAAAAACG**GTTA**CAcgat
atccG**TAAC**G**TTA**CGcgat
atccAAATTAAAAATT**CCGTTA**cgat
atccCATACTTTG**TAACCG**CCTcgat
atcc**TAACG**GTTTTAATTTCTTcgat
atcc**TAACCG**TTTCCTGTACTTcgat
atcc**TAACG**GTTTTTTCCATTCcgat
atccTTAACGTAAAAA**CCGTTA**cgat
atccATAAA**CCGTTA**CATCCACcgat
atccCCTCTTTTA**TAACCG**TTCcgat
atcc**TAACA**GATTTGAACCGATcgat
atccAATCGGAAAAAA**CGGTTA**cgat
atccCCA**TAACCG**TTTTCTCACcgat
atcc**TAACG**GTCTATCTTCATAcgat
atccTAAGCATAAAAA**CCGTTA**cgat
atccTTG**TAACC**GTTTTTTTCGcgat
atccTTATCGTAGAAA**CCGTTA**cgat
atccGTG**TAACC**GTTTTATTGCcgat
atccATCTGTGTTTC**TAACGG**Ccgat
atccGATAAA**CGGTTA**TGGAATcgat
atccTTAAAA**CGGTTA**CGCCGAcgat
atccCACG**TAACC**GTTTTTATAcgat
atccAAACTGAAA**CGGTTA**CGGcgat
atccCTAATACT**ATAACGG**CCAcgat
atcc**TAACA**GAATCCACAGTGGcgat
atcc**TAACG**GTTTCCATATGGAcgat
atccAAAAATT**CAGTTA**GTGGCcgat
atccA**TAACC**GTTTTTTTGTTAcgat
atcc**TAACC**GTCTTATACAGTTcgat
atccGAGT**TAACG**GTTTTATTAcgat
atccTC**TAACG**ACTTTTAGCTAcgat
atccG**TAACCG**CCTTAAGGAACcgat
atccCCCAACATAAAA**CGGTTA**cgat
atcc**TAACC**GTTTTATACCGGAcgat
atccGGGCTTCTCTAACC**GTTA**cgat
atccTATT**ATAACTG**CCTCATGcgat
atccG**TAACCG**TTTATTCGTAAcgat
atccCATACAGAAAAA**CCGTTA**cgat
atcc**TAACC**GTTTTTTTATCACTcgat
atccTACCTGAGTAG**ACGGTTA**cgat
atccCCTAACAAA**TAACCG**AATcgat
atccCAAAGGAG**CCGTTA**CGATcgat
atccGTTTAAA**CCGTTA**CATCTcgat
atccCATCTAAGAAAA**CGGTTA**cgat
atcc**TAACTG**CCATCTATGCATcgat
atcc**TAACC**GTTTTTGTCCGCAcgat
atccGCCATTAAAAAG**CGGTTA**cgat

Single GTTAGTTA

atcc**TAACTAAC**TCCTAATCGAcgat
atccTCCTGAAATG**TAACTACC**cgat
atcc**TAACTAAC**TGCCTCAACCcgat
atccTC**CAACAACC**TATTCTAAcgat
atccATCCCAGAAA**GTTAGTTA**cgat
atccTATAATTA**GTTTGTTA**TAcgat
atccCCCAGTATAA**GTTAGTTA**cgat
atcc**TAACTAAC**TCATTCTTACcgat
atccCG**TAACTAAC**TTTCTGTAcgat
atccATTAGATTTA**GTTAGTTA**cgat
atccGCCAATCCG**TAACTAAC**Tcgat
atcc**TAACTAAC**TTTTTCCAGAcgat
atcc**TAACAAAC**TCCCCTACACcgat
atccT**CAACTAAC**TTTTTGAGAcgat
atccTCGGATAAGA**GTTAGTTA**cgat
atcc**TAACTAAC**TTTTCTGCTCcgat
atcc**TAACTAAC**TCCTTGTTACcgat
atccG**TAACTAAC**TAGACAAACcgat
atccGTG**TAACTAAC**TTTGTCCcgat
atcc**TAACTAAC**TTCATCAACGcgat
atccCTACTCAAAA**GTTAGTTA**cgat
atccC**CAACAACC**ATGACCTAAcgat
atcc**TAACTAAC**TCCCTCAGCAcgat
atcc**CAACTAAC**TCCCTTAACAcgat
atccCC**TAACTAAC**TCTTAGGCcgat
atccC**TAACTAAC**TTTACACCCcgat

Complex

atcc**TAACTGACAGTTACAAGTcgat**
atcc**TTAACTGTCGGTTATGACcgat**
atcc**TTAACTAACTAACTAAC**Gcgat
atccCACTACTA**CAGTTAGTTA**cgat
atcc**TAACGGCTAATTCAGTTA**cgat

Summary of selected sequences

MBS-I

POSITION	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5
A	22	21	27	40	33	33	-	3	-	-	-	46	1
G	4	7	7	2	8	8	-	26	46	-	-	-	2
T	11	12	6	3	4	5	-	2	-	46	46	-	7
C	3	1	1	-	-	-	46	15	-	-	-	-	13
Ph3-I	**a**	**a**	**a**	**A**	**a/D**	**a/D**	**C**	**G/C**	**G**	**T**	**T**	**A**	
c-MYB					**D**	**D**	**C**	**A/C**	**G**	**T**	**T**	**A/G**	

MBS-II

POSITION	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5
A	12	12	11	25	-	-	-	23	-	-	-	23	1
G	6	7	8	-	27	3	-	-	27	-	-	4	4
T	6	4	4	1	-	24	27	4	-	27	27	-	1
C	1	2	2	-	-	-	-	-	-	-	-	-	5
Ph3-II	**a**	**a**	**a**	**A**	**G**	**T**	**T**	**A**	**G**	**T**	**T**	**A**	
Am305					**G**	**G/T**	**T**	**T/A**	**G**	**G**	**T**	**T/G**	

Fig. 3. Sequences of selected MYB.Ph3ΔC1 binding sites. Seventy eight site-selected oligonucleotides were sequenced after cloning in a plasmid vector, revealing the existence of two types of binding sites: CNGTTA and GTTAGTTA, respectively. Single site oligonucleotides were classified according to the type of site they contained, and those oligonucleotides containing two or more binding sites were classified as 'complex'. Bases corresponding to binding sites are highlighted by bold letters. Upper case indicates sequences derived from random core sequences; lower case indicates sequences within the constant flanking regions. Enlarged consensus sequences were determined from quantification of bases present at each position in oligonucleotides containing a single type of binding site; only bases derived from random core sequences were considered. The consensus sequences of each type of MYB.Ph3ΔC1 binding site, MBSI and MBSII, are compared with the most similar consensus sequences of binding sites of other MYB proteins, animal c-MYB and *Antirrhinum* Am305, respectively (Weston, 1992; Sablowski *et al.*, 1994). D = A, G or T.

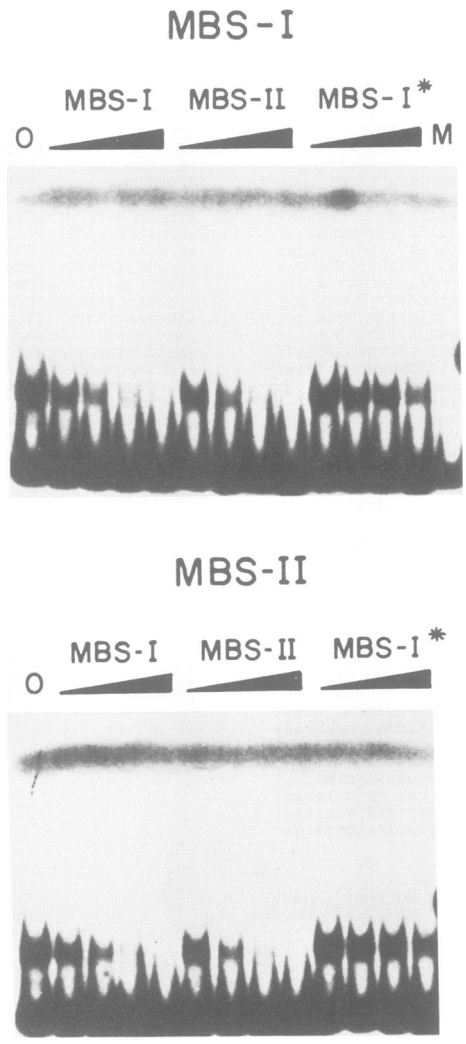

Fig. 4. Relative binding affinities of MYB.Ph3ΔC1 for MBSI (AAACGTTA) and MBSII (GTTAGTTA). Binding reactions were performed with rabbit reticulocyte lysate-translated MYB.Ph3ΔC1 (2 μl), the MBSI (top) and MBSII (bottom) probes (1 ng), and increasing amounts (10, 30, 90 and 270 ng) of either MBSI, MBSII or mutant MBSI (MBSI*, AAACGATCA) competitor. O indicates that no competitor was used; M, indicates that MBSI* was used as probe, without competitor. All reactions were performed in the presence of poly(dI-dC) (500 ng) and 150 ng of denatured calf thymus DNA.

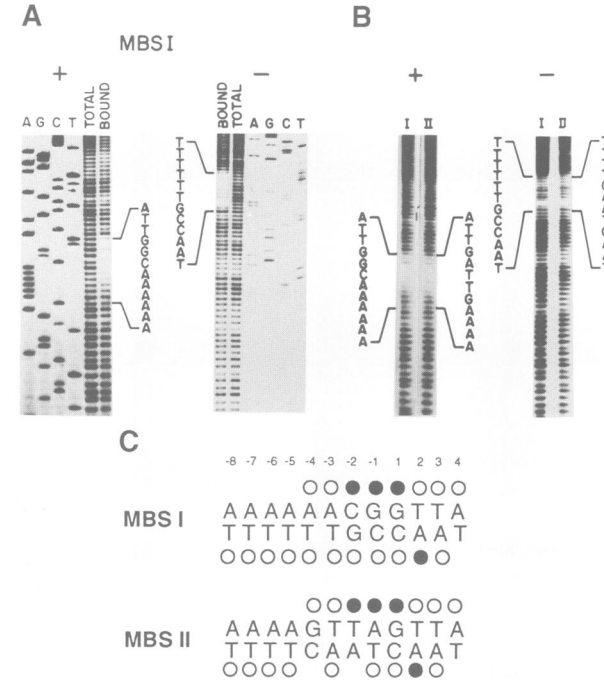

Fig. 5. Missing nucleoside, hydroxyl radical footprinting analysis of binding of protein MYB.Ph3ΔC1 to its binding sites, MBSI and MBSII. (**A**) Hydroxyl radical-treated, MBSI-containing DNA was incubated with extract from bacteria expressing MYB.Ph3ΔC1. Bound DNA was resolved from free DNA in non-denaturing PAGE, eluted from the gel and compared with original DNA (TOTAL) in denaturing PAGE. The analysis was performed on DNAs in which the + and − strands, respectively, were terminally labelled. Sequencing reactions (A, G, C and T) made by the method of Sanger *et al.* (1977) were run as controls.The region corresponding to MBSI is indicated. (**B**) Denaturing PAGE analysis of the hydroxyl radical-treated, bound MBSI DNA (I) and MBSII DNA (II) after incubation with rabbit reticulocyte lysate-translated MYB.Ph3ΔC1 protein. (**C**) Summary of missing nucleoside data. Nucleosides whose absence interferes with binding by MYB.Ph3ΔC1 are indicated: ●, strong; ○, partial.

(C/G)GTTA, and type II (MBSII), aaaAGTTAGTTA (Figure 3). These binding sites respectively represent (imperfect) inverted and direct repeated versions of a GTTA motif preceded by a 5′, A-rich flanking sequence.

Similar binding affinity of MYB.Ph3 to the two binding sites

The validity of the derived consensus sequences and the relative affinity of MYB.Ph3ΔC1 binding to MBSI and MBSII, were evaluated in competition experiments (Figure 4). Complexes between *in vitro* synthesized MYB.Ph3ΔC1 and ^{32}P-labelled oligonucleotides containing either MBSI or MBSII were more efficiently competed by an excess of cold oligonucleotide containing either of these sequences than by that containing a mutant of MBSI (AAACGATCA), indicating the involvement of MBSI and MBSII in the binding. The binding behaviour of MBSI and MBSII oligonucleotides in the competition experiments was very similar, which indicated that the affinity of MYB.Ph3 for the two binding sites was also very similar.

Similar hydroxyl radical footprints on the two binding sites

To obtain further evidence of the implication of the defined consensus binding sites, and to evaluate the relative importance of each nucleoside of MBSI and MBSII in their interaction with MYB.Ph3, hydroxyl radical footprinting analysis (the missing nucleoside technique; Hayes and Tullius, 1989) was performed. Hydroxyl radical-treated DNA, containing MBSI or MBSII, was incubated with extracts containing MYB.Ph3ΔC1 protein synthesized in *E.coli* and/or *in vitro*. Free and bound DNAs were separated by non-denaturing PAGE and the bound fraction was eluted and analysed in denaturing PAGE. As shown in Figure 5, several of the bands corresponding to MBSI or MBSII had decreased intensity, demonstrating the relevance of these sequences in mediating binding to MYB.Ph3. Moreover, the footprints of MYB.Ph3ΔC1 on MBSI and on MBSII were very similar (Figure 5B), which, as is discussed below, supported the idea that binding to MBSI and MBSII does not involve alternative orientations of MYB.Ph3 repeats. Footprints correspond-

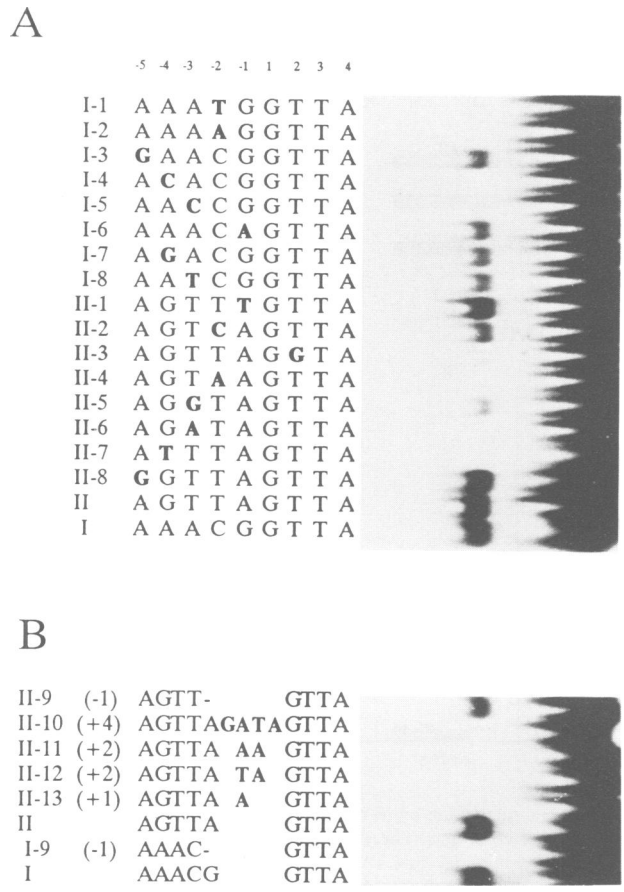

Fig. 6. Effect on binding of mutations in optimal MYB.Ph3ΔC1 binding sites. (**A**) Replacement mutations; (**B**) insertion/deletion mutations. The type of mutated oligonucleotide is indicated: I, MBSI and II, MBSII. The sequence of the relevant part of the top strand of each oligonucleotide is shown, with mutated (or inserted) bases highlighted in bold. Numbers within parenthesis in (B) indicate the number of bases inserted (+) or deleted (−). All reactions were performed with approximately equimolar amounts of protein and labelled DNA with the same specific activity (see Materials and methods).

ing to proteins synthesized in *E.coli* and *in vitro* were also similar, showing that the DNA binding properties of MYB.Ph3ΔC1 were not affected by the method of preparing the protein.

Different sequence and spacing constraints for the two sites

The effects of different mutations on binding of MBSI and MBSII were evaluated by EMSA. As shown in Figure 6A, MYB.Ph3ΔC1 protein bound better to oligonucleotides containing optimal MBSI or MBSII sequences, as judged by the binding site selection experiments, than to any of the mutated oligonucleotides, including MBSI/MBSII chimeras, indicating that the binding site selection strategy, including the PCR steps, did not introduce any bias on the selected sequences. This fully supports the conclusion that protein MYB.Ph3 does bind two types of sequences. As expected from the binding site selection experiments, MBSI had more relaxed sequence requirements than MBSII. However, a change of T to G at position −3 on MBSII (GGTAGTTA) still allowed the resulting mutant to bind to MYB.Ph3. In contrast, a change

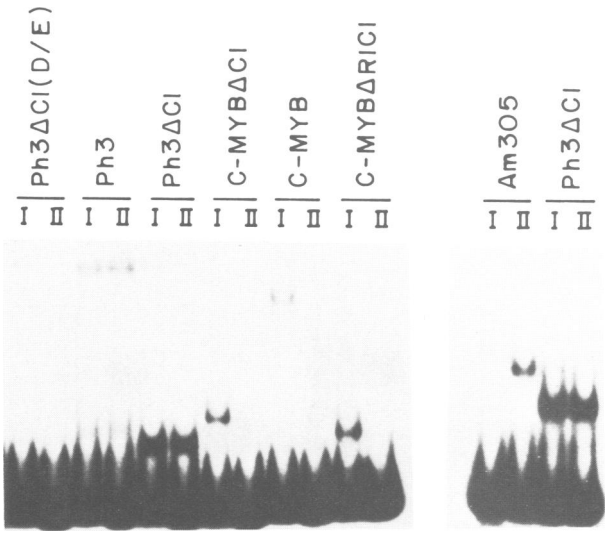

Fig. 7. Binding of different MYB proteins and mutant derivatives to optimal MYB.Ph3 binding sites. All reactions contained approximately equimolar amounts of protein and of DNA. Nomenclature for proteins as in Figure 1.

of A to C on MBSI (AACCGGTTA) greatly affected binding of the protein to the resulting oligonucleotide. These changes in MBSI and MBSII would be equivalent if binding to these sites involved alternatively oriented MYB repeats.

The effect of changes of the spacing between the two GTTA repeats in MBSII was also examined (Figure 6B). All of the insertions analysed (+1, +2 and +4) greatly reduced binding by MYB.Ph3. In contrast, deletion of one base (−1 spacing, GTTGTTA) did not greatly affect binding by the protein. The equivalent deletion in MBSI (AAACGTTA) resulted in negligible binding by MYB.Ph3.

Differential binding of other MYB proteins to MBSI and MBSII

The DNA binding experiments described above were performed with a carboxy-terminally truncated derivative of MYB.Ph3. To evaluate whether full size MYB.Ph3 protein could also bind to MBSI and MBSII, EMSA with *in vitro* synthesized proteins was performed. In addition, binding of these two sequences by murine c-MYB protein and *Antirrhinum* Am305 proteins was examined. To compare relative binding affinities, all reactions contained equimolar amounts of DNA and of proteins. As shown in Figure 7, full size MYB.Ph3 could bind equally well to MBSI and MBSII. Animal c-MYB and its deletion derivatives could bind with high affinity to MBSI and only very weakly to MBSII, and the opposite occurred with Am305 protein.

Although the relative binding affinities of full size MYB.Ph3 and of MYB.Ph3ΔC1 to MBSI and MBSII are the same, the overall affinity of MYB.Ph3 for both targets was lower than that of the truncated protein (Figure 7). To evaluate this difference further, DNA binding studies were performed with several mutants of optimal MBSI and MBSII, and in all cases the relative binding of MYB.Ph3 to these mutants was the same as that of MYB.Ph3ΔC1 (results not shown).

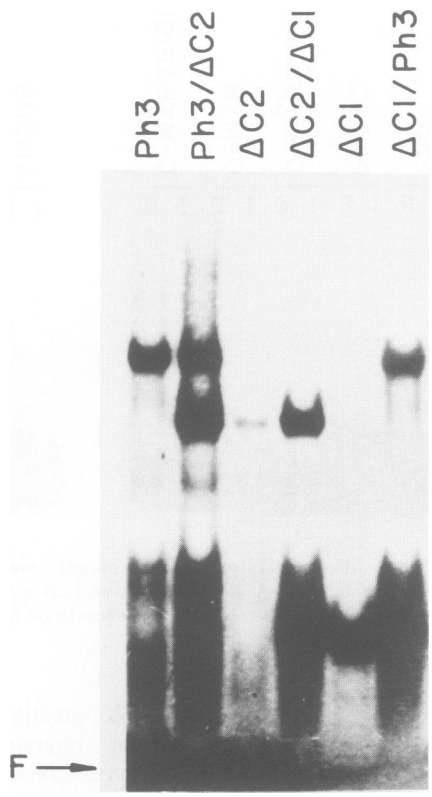

Fig. 8. EMSA of binding to MBSII by co-translated, full size and/or carboxy-terminally truncated MYB.Ph3 proteins. Full size (Ph3) and two deletions of the protein, MYB.Ph3ΔC1 (ΔC1) and MYB.Ph3ΔC2 (ΔC2), all containing an intact MYB domain, were generated by *in vitro* translation, alone or co-translated in pairwise combinations, and bound to target DNA. Two μl of lysate were used in each DNA binding assay. The position of free DNA is indicated (F). Additional, faster migrating complexes, in tracks containing *in vitro* translated MYB.Ph3 and/or MYB.Ph3ΔC2 proteins, are attributed to the presence in these protein preparations of degradation (or premature termination) products (see Figure 1).

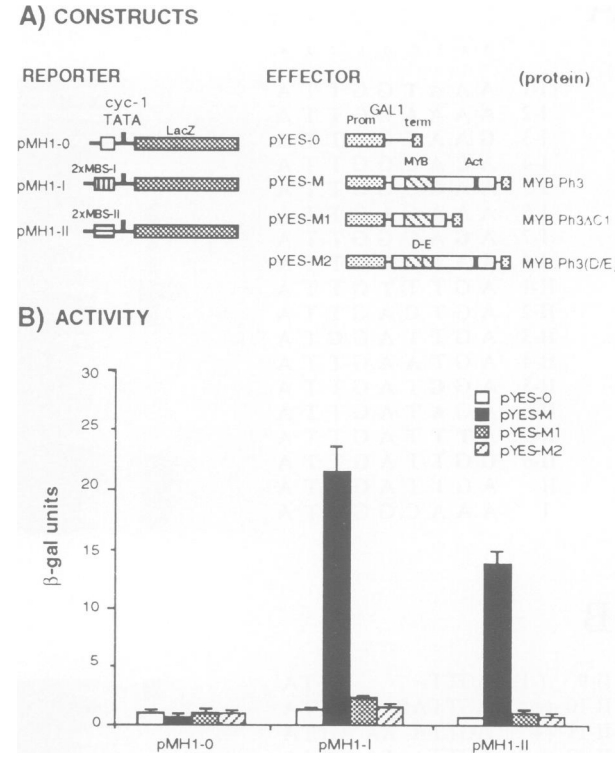

Fig. 9. Transactivation by MYB.Ph3 in yeast. (**A**) MYB.Ph3 reporter and effector constructs. The yeast integration plasmid pMH1-O, containing a minimal *cyc1* TATA box in front of the *lacZ* reporter gene, was used to prepare reporter constructs pMH-I and pMH-II by insertion of two copies in tandem of MBSI and MBSII, respectively, in the upstream region of the promoter (see Materials and methods) The yeast replicating plasmid pYES2 (pYES-O), containing the galactose-inducible *GAL1* promoter and terminator, was used as a source of effector constructs expressing proteins MYB.Ph3, MYB.Ph3ΔC1 and MYB.Ph3ΔC1(D→E). MYB, MYB DNA binding domain; Act, predicted activator domain. (**B**) β-Gal activity of yeast bearing different combinations of reporter and effector constructs. Each value is the mean of three independent transformants.

Also shown in Figure 7 is that a mutant of MYB.Ph3ΔC1, in which an aspartate residue within the MYB domain was changed to glutamate (D117; Avila *et al.*, 1993), did not bind to any of the MYB.Ph3 binding sites. This aspartate residue is conserved in all MYB proteins from plants and animals described so far (Avila *et al.*, 1993), but it is changed to a glutamate residue in the product of the dominant mutant of the *C1* gene, the colour inhibitor *C1-I* allele (Paz-Ares *et al.*, 1990, see Discussion).

MYB.Ph3 binds as a monomer to its binding sites

To investigate whether MYB.Ph3 interacts with its binding sites in the monomeric form, like its animal counterparts (Howe *et al.*, 1990), the approach of Hope and Struhl (1987) was followed: the full size and two carboxy-terminally truncated MYB.Ph3 derivatives were synthesized *in vitro*, alone or in pairwise combinations. The resulting translation or co-translation products were incubated with the oligomer containing MBSI or MBSII, and free and complexed DNA were resolved by non-denaturing PAGE. As exemplified in Figure 8 for MBSII, no band shift of intermediate mobility was obtained for any of the co-translated mixtures with either of the binding sites, indicating, at least, that the *in vitro* synthesized deletion derivatives of MYB.Ph3 bound to DNA as a monomer. The same is likely to be true for the full size protein, since otherwise a much lower mobility of the MYB.Ph3–DNA complex would have been observed.

MYB.Ph3 can activate transcription in yeast

To gain insight on whether MYB.Ph3 can activate transcription from promoters containing its target sequences, the first experiments were performed in yeast. The *myb.Ph3* coding region or its mutants were placed under the control of the galactose-inducible *GAL1* promoter and the effects of their expression were monitored, using a *lacZ* reporter gene fused to a minimal *cyc1* promoter with (or without) two copies in tandem of either MBSI or MBSII. As shown in Figure 9, the presence of MYB.Ph3 resulted in a binding site-dependent enhancement of LacZ activity. The promoter containing MBSI drove a somewhat higher MYB.Ph3-dependent LacZ activity than the MBSII-containing promoter, but the activity of the MBSI promoter in the absence of MYB.Ph3 was also higher. Deletion of the carboxy-terminal part of MYB.Ph3, putatively containing the activator domain (Avila *et al.*, 1993), practically eliminated its activation properties. The mutant

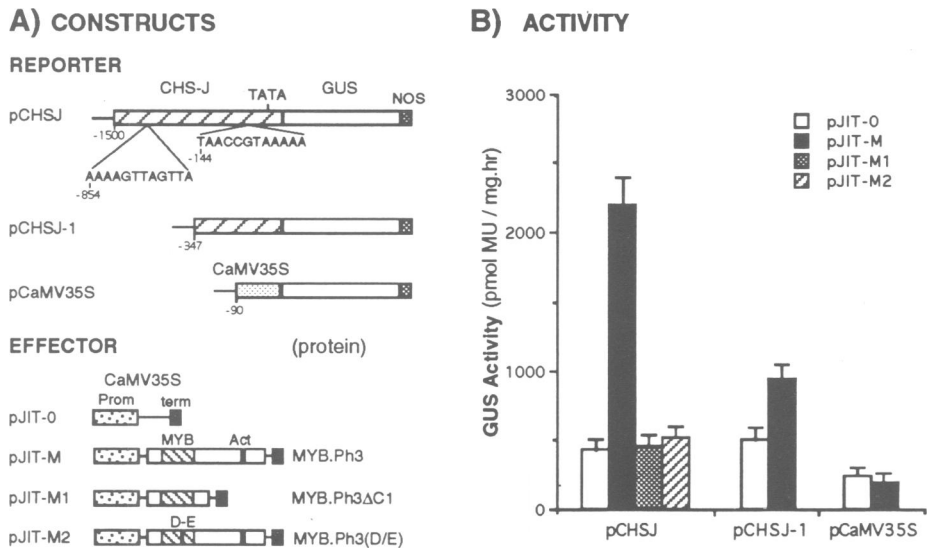

Fig. 10. MYB.Ph3-mediated transactivation from the *chsJ* promoter in tobacco protoplasts. (**A**) MYB.Ph3 reporter and effector constructs. The reporter constructs were in plasmid pUC19, and contained the following fragments fused to the *GUS* coding region: pCHSJ, a 1578 bp fragment from the *chsJ* gene (from nt −1500 to +78) present in construct VIP165 (Koes *et al.*, 1990); pCHSJ1, a 425 bp fragment from the same promoter (from nt −347 to +78, see Materials and methods); pCaMV35S, a fragment of the 35S promoter of CaMV (from nt−90; Lam *et al.*, 1989). MYB.Ph3 binding sites present in these construct are indicated. The pJIT60 plasmid containing a CaMV 35S-derived promoter and the *nos* terminator was used as the source of effector constructs expressing proteins MYB.Ph3, MYB.Ph3ΔC1 and MYB.Ph3(D→E) Other symbols as in Figure 9. (**B**) GUS activity (pmol MU/mg.h) of tobacco protoplasts transfected with different combinations of reporter and effector constructs. Each value is the mean of four independent assays.

Table I. Sequences related to MBSI and MBSII found at the upstream region of chalcone synthase (chs) genes from different species

Gene	Species	Sequence	Position	Reference
CHSJ	*Petunia hybrida*	AAAAGTTAGTTA	−854	Koes *et al.*, 1989
		TTTTTACGGTTA	−144[a]	
CHS1	*Pisum sativum*	AAAAGTTAGTTA	−600	An *et al.*, 1993
CHSG	*Sinapis alba*	ATTAGTTAGTTA	−711	Batschauer *et al.*, 1991
		AGGAGTTAGTTG	−609[a]	
		ACGAGTTAGTTG	−261	
CHS2	*Petroselium crispum*	TAATGTTAGTTA	−510[a]	Herrman *et al.*, 1988
		TCGTTTCCGTTA	−387[a]	
		AAGTAGCAGTTA	−159	
CHS	*Arabidopsis thaliana*	AGTCGTTAGTTA	−560	Feinbaum and Ausubel, 1988
		GTCAGTTTGTTA	−40	
CHS	*Antirrhinum majus*	ATTGGGTGGTTA	−622	Sommer and Saedler 1986
CHS	*Hordeum vulgare*	GGGAGTTAGTTG	−218[a]	Rohde *et al.*, 1991
CHSC2	*Zea mays*	GGAAAACCGTTA	−473[a]	Franken *et al.*, 1991
		CGGGGTTAGTTG	−172[a]	

Only one *chs* gene from each species is shown. The position is given for the most 5′-upstream nucleotide with respect to the start of transcription.
[a]The sequence is found in the reverse orientation.

MYB.Ph3 (D117→E), unable to bind to the MYB.Ph3 site, could not stimulate transcription.

MYB.Ph3 activates transcription from the promoter of a chalcone synthase gene (chsJ) in tobacco protoplasts

To evaluate whether MYB.Ph3 can activate transcription in plant cells, transient assays were performed in tobacco protoplasts. The *Petunia* chalcone synthase gene *chsJ* was chosen as a source of (target) promoter sequences, since it had been found that this promoter has MYB.Ph3 binding sites (Figure 10 and Table I; see also Discussion). The *chsJ*-derived reporter constructs contained the *GUS* coding region fused to a promoter fragment of the *chsJ* gene, starting from nucleotide −1500, which includes two MYB.Ph3 binding sites, or to a deletion derivative starting from nt −347 which lacked the distal MYB.Ph3 site (Figure 10A). In addition, a reporter construct containing the cauliflower mosaic virus (CaMV) 35S promoter from nt −90, which has no MYB.Ph3 sites, was used as a control. The constitutive expression of the coding region of *myb.Ph3* and of its mutant derivatives *myb.Ph3(d→e)* and *myb.Ph3Δc1* was driven by a CaMV 35S-derived promoter. Results of the transient expression assays are shown in Figure 10B. MYB.Ph3 could induce a 5-fold increase of *GUS* activity when the construct with the largest *chsJ* promoter fragment was used, whereas mutant proteins MYB.Ph3ΔC1 and MYB.Ph3(D→E) did not.

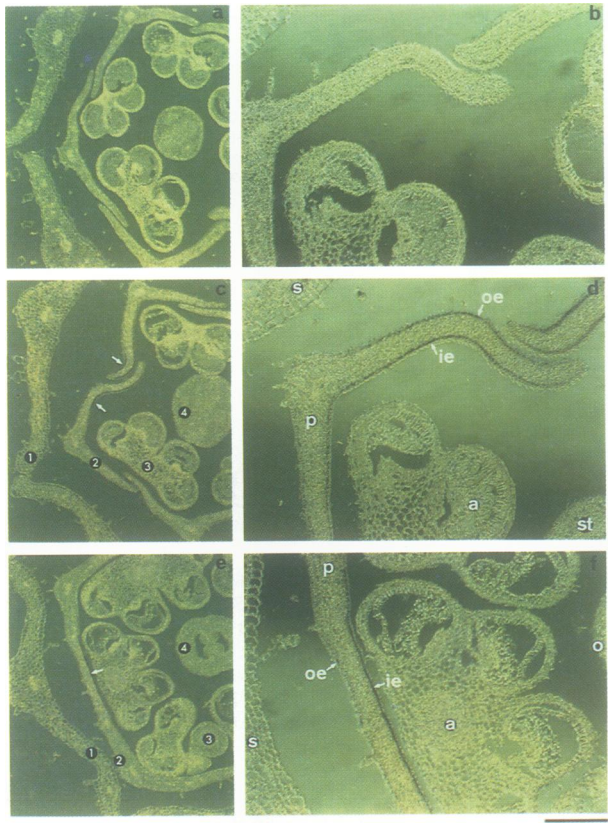

Fig. 11. Immunolocalization of the MYB.Ph3 protein in petal limbs and tubes of *P.hybrida*. The protein is only detected in the epidermis. (**a**–**d**) Cross-sections of developing *Petunia* flowers made at the level of petal limbs, treated with pre-immune serum (a and b) and with MYB.Ph3 monospecific antibodies (c and d). (**e** and **f**) Cross-section made at the level of petal tubes, treated with MYB.Ph3 monospecific antibodies. Sections in (b), (d) and (f) are the same as (a), (c) and (e), respectively, shown at a higher magnification (×2.5) Numbers indicate the four floral whorls (1–4) Other symbols: s, sepals; p, petals [limbs in (d), tubes in (f)]; oe, outer epidermis; ie, inner epidermis; a, anthers; st, pistil (style); o, ovary. Scale bar equals 800 μm in (a), (c) and (e) and 200 μm in (b), (d) and (f)

This MYB.Ph3-mediated enhancement of activity was lower but significant (2-fold) when the *chsJ*−reporter construct lacking the distal MYB.Ph3 site was used, and not significant for the CaMV-based construct. These results indicate that MYB.Ph3 can activate transcription from specific promoters in plant cells and are consistent with a requirement for a MYB.Ph3 binding site for activation. Further evidence for this would require a more detailed analysis. It should be noted that the observed 5-fold stimulation of transcription from the *chsJ* promoter is a minimum estimate of the intrinsic transcriptional activation potential of MYB.Ph3, since other MYB proteins, presumably present in tobacco protoplasts, might affect the background expression of the *chsJ* construct and/or compete with MYB.Ph3 for binding sites (see for instance Sablowski *et al.*, 1994).

Cell type specificity of MYB.Ph3

To investigate the developmental specificity of MYB.Ph3, cross-sections of developing *Petunia* flowers were investigated by *in situ* immunolocalization, using purified rabbit antibodies against MYB.Ph3. The protein was only detected in petals, as shown in Figure 11, specifically in the inner and outer epidermis of petal limbs, with a stronger signal at the petal edges, while in petal tubes the protein was mostly located in the inner epidermis. This pattern of MYB.Ph3 synthesis essentially matches that of anthocyanin accumulation, except that these pigments are also synthesized in connective tissue and its surrounding cells (Koes *et al.*, 1990), where MYB.Ph3 was not detected.

Discussion

Dual DNA binding specificity of MYB.Ph3

Binding site selection experiments presented here show that a carboxy-terminally truncated MYB.Ph3 protein, MYB.Ph3ΔC1, recognizes two different types of binding sites: MBSI, aaaAaaC(C/G)GTTA, and MBSII, aaaAGTTAGTTA. Both hydroxyl radical footprinting experiments and band-shift assays confirmed the validity of these sequences, and showed the different sequence/spacing constraints of the two types of site on binding by MYB.Ph3ΔC1. Optimal MBSI and MBSII were bound by MYB.Ph3ΔC1 with equal affinity and the same was true for the full size MYB.Ph3 protein, whereas other MYB proteins showed preferential binding to one or the other. Thus, animal c-MYB bound with high affinity to optimal MBSI but not to optimal MBSII, while the opposite occurred with *Antirrhinum* Am305 protein, which only bound with high affinity to MBSII. Hence, binding to each type of site is differently sensitive to changes in their nucleotide sequence as well as to changes in the amino acid sequence of the MYB domain. Whether other MYB proteins, including c-MYB and Am305, can also bind to two types of sequences remains unanswered, as, for instance, the inability of Am305 to bind efficiently to optimal MBSI does not preclude its binding with high affinity to as yet unknown variants of this type of sequence. In any case, the existence of at least two types of sequences whose binding by MYB proteins is differentially affected by mutations in the MYB domain, could have greatly contributed to the generation of complex MYB regulatory networks, and help to justify the presence of ~100 MYB proteins in the small genome of *Arabidopsis* (I.Romero and J.Paz-Ares, unpublished data).

The overall affinity of the full size protein for either of the binding sites was lower than that of its truncated derivative (Figure 7). The reason for the decreased affinity of MYB.Ph3 remains to be established. However, the fact that all mutations of MBSI and MBSII tested had a similar relative effect on binding by MYB.Ph3 and MYB.Ph3ΔC1 strongly suggests that both proteins have practically identical DNA binding specificities. A similar situation has been reported for animal c-MYB, in which the DNA binding affinity of carboxy-terminally truncated versions is higher than that of full size c-MYB (Ramsay *et al.*, 1992; see also Figure 7). In any case, as shown here, MYB.Ph3 could activate transcription from promoters containing the deduced MYB.Ph3 binding sites both in yeast and in tobacco protoplasts.

The two DNA binding sites of MYB.Ph3, MBSI and MBSII, represent (imperfect) inverted and direct repeats of the GTTA motif, respectively, flanked at their 5′ ends by an A-rich region. However, binding of MYB.Ph3 to these sites does not seem to involve alternative orientations of the two MYB repeats (R2 and R3) of MYB.Ph3,

contrary to the case of other proteins interacting with variably oriented bipartite DNA-binding sites, such as members of the glucocorticoid receptor and POU protein families from animals and the Ara C protein from *E.coli* (Carra and Schleif, 1993; Kurokawa *et al.*, 1993; Li *et al.*, 1993). Thus, the hydroxyl radical footprints of MYB.Ph3 binding on MBSI and MBSII were very similar and did not present specular images in half of the site, as would be expected if MYB.Ph3 repeats adopted two different relative orientations when binding to MBSI and MBSII. In addition, a mutant of MBSII (R.Solano, C.Nieto and J.Paz-Ares, submitted) in which T at position −3 was substituted for a G (oligonucleotide II-5 in Figure 7) was still bound by MYB.Ph3. In contrast, a mutation in MBSI at position −2 (C for A, oligonucleotide I-5 in Figure 7), the 'equivalent' mutation if one of the MYB repeats had the opposite orientation, was greatly impaired in its binding by MYB.Ph3. Moreover, we have shown that MYB.Ph3 induces similar distortions upon binding to MBSI and MBSII (R.Solano, C.Nieto and J.Paz-Ares, submitted). Furthermore, model building studies with MYB domains only allow the tandem orientation for the two MYB repeats (Framptom *et al.*, 1991), an organization experimentally proven for the MYB domain of animal c-MYB (Ogata *et al.*, 1994).

The solution structure of the complex of a MBSI sequence with the minimum DNA binding domain of animal c-MYB indicates that repeats R2 and R3 respectively interact with the 5′ and 3′ parts of the site, although their regions of interaction are partially overlapping (Ogata *et al.*, 1994). Assuming a similar disposition for the equivalent repeats of c-MYB and MYB.Ph3 in binding to MBSI, the same would be true for the case of MBSII, as binding of MYB.Ph3 to both sites results in a similar hydroxyl radical footprint. This would point to repeat R2 as the most involved in determining the dual DNA binding specificity of MYB.Ph3, since the differences between the two types of sites are concentrated in their 5′ part. In agreement with this is the presumed structural flexibility of R2 repeats (Jamin *et al.*, 1993; Myrset *et al.*, 1993; Ording *et al.*, 1994). The specific structural flexibility properties of this repeat in a particular MYB protein would determine, in part, the different sequence constraints in the recognition of MBSI and MBSII, similarly to what has been proposed for bZIP proteins (Kim *et al.*, 1993) .

Functionality of MYB.Ph3: the effect of a single conservative mutation within the MYB domain

Transcription from promoters containing MBSI and/or MBSII binding sites was shown to be activated by MYB.Ph3 in yeast and in plant protoplasts. A mutation within the second repeat of MYB.Ph3 (D117→E; Avila *et al.*, 1993), which abolished binding of the mutant protein to wild-type binding sites, also impaired its capacity to activate transcription in both the plant and yeast system. In addition, removal of the carboxy-terminal part of the protein, where major activator domain(s) have been predicted (Avila *et al.*, 1993), greatly affected transcriptional activation capacity. Thus, MYB.Ph3 has the general properties of transcriptional activators (Ptashne, 1988). However, these results do not preclude the presence in MYB.Ph3 of negative regulatory domains and, under certain conditions, this protein could repress rather than activate transcription, as happens with animal c-MYB (Sakura *et al.*, 1989).

The observed effect on DNA binding, and concomitantly on transcriptional activation properties, of the D→E mutation in MYB.Ph3, could also apply to other MYB proteins, as this residue is highly conserved in plant and animal MYB proteins (Avila *et al.*, 1993). Thus, a similar change in C1 resulted in a mutant protein with a drastically decreased (45-fold) capacity to stimulate transcription from the promoter of the *Bz* gene, a gene under the control of *C1* (Goff *et al.*, 1990, 1991), in agreement with the expected effect of this mutation on MYB DNA binding properties. Recently, Franken *et al.* (1994) have shown that the D→E mutation only results in a modest reduction (2-fold) of transcriptional activation of the *C1*-regulated *A1* gene. This could reflect the fact that transcriptional activation by the C1 protein might not require a highly specific, C1-mediated recognition of the promoter. It has been shown for animal c-MYB that its R2 and R3 repeats interact with each other *in cis* when binding to DNA, and that this interaction involves the conserved aspartate residue which forms a salt bridge with an arginine residue from the R2 repeat, highly conserved in MYB proteins (R64 in MYB.Ph3, Avila *et al.*, 1993; Ogata *et al.*, 1994). In this context, the change of an aspartate residue for a glutamate, which is larger, could affect the R2−R3 interaction, thereby preventing the formation of a stable complex with DNA.

The D→E mutation also occurs in the dominant colour inhibitor mutant (*C1-I*) of the *C1* gene (Paz-Ares *et al.*, 1990), meaning that competition between their proteins for DNA binding sites is unlikely to be the only cause of the dominant negative effect of the *C1-I* allele. Alternatively, these proteins could compete for other proteins, such as the R protein which interacts with C1 protein (Goff *et al.*, 1992), or the C1 and C1-I proteins might interact with each other to produce a transcriptionally inactive complex.

Possible role of MYB.Ph3 in the regulation of flavonoid biosynthesis

Genetic and molecular evidence has shown the involvement of three plant MYB proteins (C1, Pl and P from maize) in the regulation of the synthesis of flavonoids (Cone *et al.*, 1986, 1993; Paz-Ares *et al.*, 1987; Grotewold *et al.*, 1991). In the case of MYB.Ph3, three lines of evidence strongly suggest that it is also involved in the regulation of flavonoid biosynthetic genes, although genetic evidence *in planta* is missing at present. First, inspection of proximal upstream regions of flavonoid biosynthetic genes revealed the presence of one or more MYB.Ph3 binding sites, or related motifs, in several of these genes, particularly in *chs* genes, whose product catalyses the first flavonoid-specific enzymatic step in the phenylpropanoid pathway (Hahlbrock, 1981). As shown in Table I, in all of the plant species examined, there was a member of the *chs* gene family which contained optimal (or close to optimal) MYB.Ph3 binding sites, except in *Antirrhinum*, where only a distantly related motif was found in its only known *chs* gene. It is very unlikely that this finding is due to chance ($P < 10^{-10}$). Second, MYB.Ph3 was shown to activate transcription from the promoter of the *chsJ* gene of *Petunia* in tobacco protoplasts

(Figure 10), in agreement with the presence of the predicted MYB.Ph3 binding sites in this promoter. Third, MYB.Ph3 was found to be located predominantly in epidermal cells of *Petunia* petals (Figure 11), a cell type which is rich in flavonoid pigments and in which flavonoid biosynthetic genes, including *chsJ*, are actively expressed (Koes *et al.*, 1990).

Data for MYB.Ph3, one of the known plant MYB proteins most distantly related to the C1, Pl and P proteins (Avila *et al.*, 1993), further adds to recent evidence implicating three more MYB proteins (Zm1 and Zm38 from maize and Am305 from *Antirrhinum*) in the regulation of the synthesis of flavonoids and/or related compounds (Franken *et al.*, 1994; Sablowski *et al.*, 1994), a role previously proposed for most known plant MYB proteins (Avila *et al.*, 1993). However, the possibility that MYB.Ph3 or other related MYB proteins might also regulate unrelated pathways (target genes) should not be excluded.

We notice that some *chs* genes contain more than one MYB.Ph3 binding site (Table I), and, in addition, in many *chs* promoters there are other motifs more closely resembling binding sites of Am305 and/or P proteins than of MYB.Ph3 protein (not shown). This probably reflects that, in the generation of novel regulatory circuits, there was duplication and divergence of target sites, in addition to duplication and divergence of regulatory proteins.

Materials and methods

Standard molecular procedures

All methods, except where indicated, were performed as previously described (Avila *et al.*, 1993; Sambrook *et al.*, 1989)

Constructs for E.coli and in vitro synthesis of proteins

MYB.Ph3 protein and mutant derivatives shown in Figure 1 were expressed in *E.coli* and/or *in vitro* after cloning appropriate fragments of the *myb.Ph3* cDNA in the T7 RNA polymerase expression vector pET3b (Studier *et al.*, 1990). For synthesis of amino-terminal truncations, MYB.Ph3ΔN1 and MYB.Ph3ΔN2, the corresponding 5′-truncated fragments of *myb.Ph3* cDNA (from the *Nco*I site at position 393 and the *Eco*RV site at position 838, respectively; Avila *et al.*, 1993) were cloned in the *Bam*HI site of pET3b. These proteins contained 11 extra amino acid residues from the vector sequences. For expression of full size MYB.Ph3 and of its carboxy-terminally truncated versions, an *Nde*I site was engineered, by PCR, in the initiation codon of *myb.Ph3* cDNA. MYB.Ph3ΔC1 and MYB.Ph3ΔC2 coding fragments end at an *Eco*RV site (position 838) and at an *Mae*I site (position 1527), respectively (Avila *et al.*, 1993). The mutant MYB.Ph3ΔC1(D→E), which replaces an aspartate residue (position 117 in MYB.Ph3; Avila *et al.*, 1993) with a glutamate residue in MYB.Ph3ΔC1, was obtained after site-directed mutagenesis of *myb.Ph3* cDNA by PCR as described by Cormack (1992).

Antirrhinum Am305 protein, murine c-MYB and its carboxy-terminal truncated version c-MYBΔC1 (c-MYBΔ45 in Howe *et al.*, 1990) were prepared from previously described T7-based constructs (Howe *et al.*, 1990; Sablowski *et al.*, 1994). To synthesize c-MYBΔR1C1, the coding region 5′ upstream of repeat R2 of the c-*mybΔc1* construct was substituted by an ATG codon by means of PCR. All PCR fragments used in the different constructs were confirmed by sequencing

Protein synthesis in E.coli and in vitro

Bacterial protein synthesis was induced from *E.coli* strain BL21(DE3) harbouring the corresponding pET:*myb.Ph3* construct, essentially as described by Studier *et al.* (1990), except that cultures corresponding to proteins to be used in DNA binding experiments (full size MYB.Ph3, MYB.Ph3ΔC1 and MYB.Ph3ΔN2) were grown at 30°C. Bacterial cells were harvested by centrifugation and were frozen at −80°C. After thawing and resuspension in 1/100 vol of extraction buffer (20 mM Tris−HCl, pH 7.5; 1 mM EDTA; 400 mM NaCl; 5 mM β-ME; 1 mM PMSF), the cells were lysed by passage through a French press (1200 p.s.i.). The lysate was centrifuged at 5000 *g* and the supernatant was further centrifuged (1 h, 40 000 *g*). The resulting supernatant was made 20% glycerol and stored at −80°C, after aliquoting and freezing in liquid N_2. In the case of MYB.Ph3ΔN1, the 5000 *g* pellet containing the inclusion bodies was used as a source of protein for the production of antibodies.

In vitro protein synthesis was performed in the rabbit reticulocyte system (nuclease-treated) in the presence of [^{35}S]methionine, with RNA prepared by *in vitro* transcription of the corresponding constructs with the T7 DNA polymerase, following the manufacturer's instructions (Promega). Prior to the transcription, the constructs were linearized with appropriate restriction enzymes.

Anti-MYB.Ph3 antibodies

Polyclonal anti-MYB.Ph3 antibodies were raised by immunization of rabbits with MYB.Ph3ΔN1 protein. The source of MYB.Ph3ΔN1 was the washed, inclusion body fraction of extracts of induced cells harbouring the corresponding construct, in which this protein was the major component (not shown). From this fraction, the protein was further purified by SDS−PAGE (Laemmli, 1970) and electroeluted. For *in situ* immunolocalization experiments, MYB.Ph3 monospecific antibodies were isolated by incubation with nitrocellulose-immobilized MYB.Ph3ΔN2 protein, as described by Goding (1983). These antibodies only recognized MYB.Ph3, as indicated by Western blot analysis (not shown).

Oligonucleotides

Oligonucleotide mixture O (5′-TCGACTCGAGTCGACATCGN_{18}GG-ATCCTGCAGAATTCGCG-3′), containing 18 random nucleotides flanked by 19 bases of defined sequence, was used as the source of MYB.Ph3 binding sites in the binding site selection procedure. The oligonucleotide mixture was made double-stranded with primer F0 (5′-CGCGAATTCTGCAGGATCC-3′). Primers F0 and R0 (5′-TCGAC-TCGAGTCGACATCG-3′) were used for PCR labelling and amplification during site selection.

MBSI (5′-TGCCGCTCGAGCTTGTAACCGTTTTTTGTCGACTC-GAGCGGT-3′) and MBSII (5′-TGCCGCTCGAGCTTGTAACTAA-CTTTTGTCGACTCGAGCGGT-3′) oligonucleotides were used as representatives of optimal MYB.Ph3 binding sites, MBSI and MBSII respectively. The sequences of other oligonucleotides containing mutated versions of MBSI and MBSII are shown in Figures 2 and 7. The experiment shown in Figure 7 was done with oligonucleotides directly labelled and amplified with primers F1 (5′-ACCGCTCGAGTCGAC-3′) and R1 (5′-TGCCGCTCGAGCTTG-3′). For all other experiments, oligonucleotides were made double-stranded with primer F1 and were cloned in the *Sal*I site of pUC19 following digestion with *Sal*I−*Xho*I. Labelling of cloned oligonucleotides, to be used in gel retardation experiments, was done by PCR with the normal and reverse universal sequencing primers. For the missing nucleoside assay, DNA was terminally labelled on one strand by PCR amplification with a mixture of polynucleotide kinase-labelled, normal and cold reverse universal sequencing primers (or the reciprocal). In all cases, PCR amplification was for 20 cycles and the labelling mixture consisted of 50 μM dNTP (each) and 20 μCi of [α-^{32}P]dCTP or of one ^{32}P-labelled primer. Before using for DNA binding assays, the labelled DNA was purified by PAGE.

DNA binding reactions and electrophoretic mobility shift assay (EMSA)

DNA binding reactions were performed in a buffer containing 10 mM Tris (pH 8), 1 mM EDTA, 100 mM NaCl, 2 mM dithiothreitol and 10% glycerol. Bacterial protein extracts (30–1000 ng) were mixed with (1–10 ng) ^{32}P-labelled DNA, 0.5 μg poly(dI-dC) and BSA (250 μg/μl) and incubated for 30 min on ice. For *in vitro* synthesized proteins, 1 ng of labelled DNA and up to 2 μl of rabbit reticulocyte lysate were used under the same conditions, except that denatured salmon sperm DNA (150 ng) was included and BSA was omitted. In experiments in which equimolar amounts of different proteins were used, the amount of each protein was estimated, after SDS−PAGE of rabbit reticulocyte extracts, by measurement of ^{35}S c.p.m. in the corresponding protein band and correction for Met content. EMSA to separate free and bound DNA was in 6% polyacrylamide gel (40:1 bisacrylamide cast in 0.5× TBE; TBE is 89 mM Tris, 89 mM boric acid and 2 mM EDTA).

Binding site selection

The double-stranded oligonucleotide mixture O was used as a source of MYB.Ph3 binding sites in the binding site selection procedure. The dsDNA (10 ng) was incubated with extracts (1000 ng) from bacteria

expressing either MYB.Ph3ΔC1 (~3–5% of total protein, not shown) or MYB.Ph3ΔN2 (control) and the protein–DNA complexes were separated from the free DNA by EMSA. The resulting gel was dried without any fixation treatment, and the DNA present in the complex with MYB.Ph3ΔC1 was amplified directly from the dried gel as described by Pollock and Treisman (1990). The amplified DNA was used in a second round of selection and the process was repeated three more times. The conditions for these subsequent rounds of selection were the same as in the first round except for the amount of protein used. This was empirically calculated in each round of selection to result in a retarded band corresponding to a 2–4% of the DNA used in the experiment. After five rounds of selection, the amplified DNA was digested with SalI and EcoRI and cloned into pUC19. Individual clones were sequenced.

Missing nucleoside assay

The missing nucleoside assay was as described by Hayes and Tullius (1989), with minor modifications. Single-stranded, terminally labelled DNA fragments containing the MBSI or MBSII oligonucleotide were treated with the hydroxyl radical reagents described by Tullius and Dombroski (1986), except that the concentration of the Fe(II)–EDTA solution and of H_2O_2 was increased (2×). Treated DNAs (10 ng) were incubated with extract containing MYB.Ph3ΔC1 protein synthesized in *E.coli* and/or *in vitro* (10 µl), in a final volume of 75 µl. Bound DNA was separated from free DNA by EMSA, eluted in a solution of 0.5 M ammonium acetate, 0.1 mM EDTA, 10 mM magnesium acetate and 0.1% SDS and analysed in denaturing 8% polyacrylamide gels. Sequencing reactions, shown as control in Figure 5, were carried out by the method of Sanger *et al.* (1977). The correspondence between the fragments generated by this sequencing method and those generated by the hydroxyl radical treatment, was determined by comparison with fragments of control DNA generated by the chemical sequencing method (Maxam and Gilbert, 1980) (not shown).

Transcriptional activation assays in yeast

Plasmid vectors for preparing reporter and effector constructs used in this work (see Figure 9) were the integrating plasmid pMH1 (M.Holsworth, manuscript in preparation) and the autonomously replicating plasmid pYES2 (Invitrogen Corp), respectively, as in Sablowski *et al.* (1994). pMH1 was first modified by cloning the polylinker sequence of plasmid pBluescript (Stratagene; the blunt-ended *Sac*I–*Kpn*I fragment) in its unique blunt-ended *Bgl*II site, producing pMH1-O. Two tandem copies of MBS-I and MBSII double-stranded oligonucleotides, digested with *Sal*I and *Xho*I, were first cloned in the *Sal*I site of pUC19 (Vieira and Messing, 1982). The *Sal*I–*Sma*I fragment from these clones containing MBSI and MBSII sequences were cloned in *Sal*I/*Sma*I-digested pMH1-O, giving pMH1-I and pMH1-II, respectively. Transformation of yeast cells (*Saccharomyces cerevisiae* strain JB811: *ura3-52*, *leu2*, *trp1*) with reporter and effector constructs and growth conditions of transformants were essentially as described by Sablowski *et al.* (1994). β-Galactosidase activity was assayed as described by Miller (1972). Three independent transformants were used to determine each value.

Transcriptional activation assays in tobacco protoplasts

Reporter constructs, pCHSJ, pCHSJ1 and pCaMV35S (Figure 10), were in pUC19 and all contained the *GUS* coding region fused to the *nos* terminator. pCHSJ contains the *chsJ–GUS* fusion, VIP 165, described by Koes *et al.* (1990). pCHSJ1 contains a 5′-deletion derivative of the construct in pCHSJ (up to nt −348, Koes *et al.*, 1989), and was prepared by deleting the *Bam*HI–*Hin*dIII fragment from pCHSJ. pCaMV35S contains a fragment of the 35S promoter of CaMV from nt −90 (Lam *et al.*, 1989). Effector constructs, pJIT-O, pJIT-M, pJIT-M1 and pJIT-M2, were prepared by cloning appropriate parts of *myb.Ph3* cDNA or of its mutant *myb.Ph3(d→e)* in the plasmid pJIT60, containing a CaMV 35S-derived promoter and the *nos* terminator (J.F.Guerineau, unpublished).

Leaf protoplasts were isolated from axenic shoot cultures of *Nicotiana tabacum* W38 by established procedures (Power *et al.*, 1989). 5×10^5 protoplasts resuspended in 12% mannitol solution, were transiently transfected using 50 µg of salmon sperm DNA as a carrier, 7 µg of reporter plasmid and 42 µg of effector plasmid, by the electroporation method at 48 Ω, 500 µF, 750 V/cm for 20–30 ms, with an Electro Cell Manipulator 600 (BTX Electroporation system). Protoplasts were plated under the conditions previously described by Piñero *et al.* (1994) for 30 h and *GUS* activity was measured by the fluorometric assay of Jefferson (1987). Protein content was determined using the Bio-Rad kit, with bovine serum albumin as standard.

In situ immunolocalization

Petunia hybrida variety V26 plants were grown under normal greenhouse conditions. Flower buds between 1 and 3 cm long, where *myb.Ph3* expression is maximal (Avila *et al.*, 1993), were harvested and fixed immediately in phosphate-buffered saline (PBS) with 4% paraformaldehyde. After 4 h at 4°C, tissues were washed in PBS (3×10 min), dehydrated in ethanol/*tert*-butanol series and embedded in paraffin (Paraplast Plus). Specimens were mounted on a cryostat Micron HM 500 and sections of 4–6 µm thick were transferred to glass slides coated with 2% Tespa (3-aminopropyltriethoxysilane) in acetone. To immunolocalize MYB.Ph3, the sections were incubated for 2 h with MYB.Ph3 monospecific antibodies and then for 2 h with alkaline phosphatase-conjugated second antibodies (Parets-Soler *et al.*, 1993). Control slides were treated with pre-immune serum. A Nikon Diaphot microscope was used for sample visualization and photography.

Computer analysis

A computer-assisted search of MYB.Ph3 binding sites at the promoter region (1000 bp) of flavonoid biosynthetic genes was done using the 'geneman' program of the DNASTAR software package.

Acknowledgements

We are very grateful to Drs Cathie Martin, Mike Bevan, Enriqueta Moyano and Robert Sablowski for exchanging results prior to publication and for providing us with the *Am305* construct, the yeast strain and yeast vectors used in this study as well as the pJIT60 vector, to Professor Roger Watson for providing us with the animal c-*myb* constructs, and to Drs Ronald Koes and Francesca Quattrocchio for providing the pCHSJ construct. We thank Professor Francisco García-Olmedo and Drs Miguel Angel Peñalva and Cathie Martin for critical reading of the manuscript. The excellent technical assistance of Maria Jesus Benito is gratefully acknowledged. R.S. was recipient of a pre-doctoral fellowship from the Comunidad Autónoma de Madrid, L.C. and J.A. were respectively recipients of a post-doctoral and a pre-doctoral fellowship from the Ministerio de Educación y Ciencia. This research was supported by the European Community (contracts BIOT-CT90-0164A and BIO 2-CT93-0101).

References

An,C., Ichinose,Y., Yamada,T., Tanaka,Y., Shiraishi,T. and Oku,H. (1993) *Plant Mol. Biol.*, **21**, 789–803.

Avila,J., Nieto,C., Cañas,L.A. and Paz-Ares,J. (1993) *Plant J.*, **3**, 553–562.

Baranowskij,N., Frohberg,C., Prat,S. and Willmitzer,L. (1994) *EMBO J.*, **13**, 5383–5392.

Batschauer,A., Ehmann,B. and Schaefer,E. (1991) *Plant Mol. Biol.*, **16**, 175–185.

Blackwell,T.K. and Weintraub,H. (1990) *Science*, **250**, 1104–1110.

Carra,J.H. and Schleif,R.F. (1993) *EMBO J.*, **12**, 35–44.

Cone,K.C., Burr,F.A. and Burr,B. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 9631–9635.

Cone,K.C., Cocciolone,S.M., Burr,F.A. and Burr,B. (1993) *Plant Cell*, **5**, 1795–1805.

Cormack,B. (1992) In Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Smith,J.A., Seidman,J.G. and Struhl,K. (eds), *Current Protocols in Molecular Biology*. Wiley-Interscience, New York, Vol. 1, Unit 8.5.

Feinbaum,R.L. and Ausubel,F.M. (1988) *Mol. Cell. Biol.*, **8**, 1985–1992.

Foos,G., Grimm,S. and Klempnauer,K.H. (1992) *EMBO J.*, **11**, 4619–4629.

Frampton,J., Gibson,T.J., Ness,S.A., Döderlein,G. and Graf,T. (1991) *Protein Engng*, **4**, 891–901.

Franken,P., Schrell,S., Peterson,P.A., Saedler,H. and Wienand,U. (1994) *Plant J.*, **6**, 21–30.

Franken,P., Niesbach-Kloesgen,U., Weydemann,U., Marechal-Drouard, L., Saedler,H. and Wienand,U. (1991) *EMBO J.*, **10**, 2005–2012.

Goding,J.W. (1983) *Monoclonal Antibodies: Principles and Practice*. Academic Press, New York.

Goff,S.A., Klein,T.M., Roth,B.A., Fromm,M.E., Cone,K.C., Radicella,J.P. and Chandler,V.L. (1990) *EMBO J.*, **9**, 2517–2522.

Goff,S.A., Cone,K.C. and Fromm,M.E. (1991) *Genes Dev.*, **5**, 298–309.

Goff,S.A., Cone,K.C. and Chandler,V.L. (1992) *Genes Dev.*, **6**, 864–875.

Graf,T. (1992) *Curr. Opin. Genet. Dev.*, **2**, 249–255.

Grotewold,E., Athma,P. and Peterson,T. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 4587–4591.
Grotewold,E., Drummond,B.J., Bowen,B. and Peterson,T. (1994) *Cell*, **76**, 543–553.
Hahlbrock,K. (1981) In Conn,E.E. (ed.), *The Biochemistry of Plants*. Academic Press Inc., London, Vol. 7, pp.425–453.
Hahlbrock,K. and Scheel,D. (1989) *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **40**, 347–369.
Hayes,J.J. and Tullius,T.D. (1989) *Biochemistry*, **28**, 9521–9527.
Herrman,A., Schulz,W. and Hahlbrock,K. (1988) *Mol. Gen. Genet.*, **212**, 93–98
Hope,I.A. and Struhl,K. (1987) *EMBO J.*, **6**, 2781–2784.
Howe,K.M., Reakes,C.F.L. and Watson,R.J. (1990) *EMBO J.*, **9**, 161–169
Jamin,N., Gabrielsen,O.S., Gilles,N., Lirsac,P.N. and Toma,F. (1993) *Eur. J. Biochem.*, **216**, 147–154.
Jackson,D., Culianez-Macia,F., Prescott,A., Roberts,K. and Martin,C. (1991) *Plant Cell*, **3**, 115–125
Jefferson,R.A. (1987) *Plant Mol. Biol. Rep.*, **5**, 387–405.
Katzen,A.L., Kornberg,T.B. and Bishop,J.M. (1985) *Cell*, **41**, 449–456.
Kim,J., Tzamarias,D., Ellenberger,T., Harrison,S.C. and Struhl,K. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 4513–4517.
Koes,R.E., Spelt,C.E., van Den Elzen,P.J.M. and Mol,J.N.M. (1989) *Gene*, **81**, 245–257.
Koes,R.E., van Blokland,R., Quattrocchio,F., van Tunen,A.J. and Mol,J.N.M. (1990) *Plant Cell*, **2**, 379–392.
Kurokawa,R., Yu,V., Näär,A., Kyakumoto,S., Han,Z., Silverman,S., Rosenfeld,M.G. and Glass,C.K. (1993) *Genes Dev.*, **7**, 1423–1435.
Laemmli,U.K. (1970) *Nature*, **227**, 680–695.
Lam,E., Benfey,P.N., Gilmartin,P.M., Fang,R.-X. and Chua,N.-H. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 7890–7894.
Li,P., He,X., Gerrero,M.R., Mok,M., Aggarwal,A. and Rosenfeld,M.G. (1993) *Genes Dev.*, **7**, 2483–2496.
Lüscher,B. and Eiseman,R.N. (1990) *Genes Dev.*, **4**, 2235–2241.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499–560.
Miller,J. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Myrset,A.H., Bostad,A., Jamin,N., Lirsac,P.N., Toma,F. and Gabrielsen,O.S. (1993) *EMBO J.*, **12**, 4625–4633
Noda,K.I., Glover,B.J., Lindstead,P. and Martin,C. (1994) *Nature*, **369**, 661–664.
Ogata,K., Hojo,H., Aimoto,S., Nakai,T., Nakamura,H., Sarai,A., Ishii,S. and Nishimura,Y. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 6428–6432.
Ogata,K., Morikawa,S., Nakamura,H., Sekikawa,A., Inoue,T., Kanai,H., Sarai,A., Ishii,S. and Nishimura,Y. (1994) *Cell*,**79**, 639–648.
Oppenheimer,D.G., Herman,P.L., Sivakuaran,S., Esch,J. and Marks,M.D. (1991) *Cell*, **67**, 483–493.
Ording,E., Kvavik,W., Bostad,A. and Gabrielsen,O.S. (1994) *Eur. J. Biochem.*, **222**, 113–120.
Parets-Soler,A., Cañas,L.A. and Beltrán,J.P. (1993) *Plant Sci.*, **92**, 57–67.
Paz-Ares,J., Ghosal,D., Wienand,U., Peterson,P.A. and Saedler,H. (1987) *EMBO J.*, **6**, 3553–3558.
Paz-Ares,J., Ghosal,D. and Saedler,H. (1990) *EMBO J.*, **9**, 315–321.
Piñero,M., Garcia-Olmedo,F., and Díaz,I. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 3867–3871.
Pollock,R. and Treisman,R. (1990) *Nucleic Acids Res.*, **18**, 6197–6204.
Power,J.B., Davey,M.R., McLellan,M. and Wilson,D. (1989) *Laboratory Manual of Plant Tissue Culture*. University of Nottingham, UK.
Ptashne,M. (1988) *Nature*, **335**, 683–689.
Ramsay,R.G., Ishii,S. and Gonda,T.J. (1992) *J. Biol. Chem.*, **267**, 5656–5662.
Rohde,W., Dörr,S., Salamini,F. and Becker,D. (1991) *Plant Mol. Biol.*, **16**, 1103–1106.
Sablowski,R.W.M., Moyano,E., Culiañez-Macia,F.A., Schuch,W., Martin,C. and Bevan,M. (1994) *EMBO J.*, **13**, 128–137.
Sakura,H., Kanei.Ishii,C., Nagase,T., Nakagoshi,H., Gonda,T.J. and Ishii,S. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 5758–5762.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sanger,F., Nicklen,S. and Coulsen,A.R. (1977) *Proc. Natl Acad. Sci. USA.*, **74**, 5463–5467.
Sarai,A., Uedaira,H., Morii,H., Yasukawa,T., Ogata,K., Nishimura,Y. and Ishii,S. (1993) *Biochemistry*, **32**, 7759–7764.
Sommer,H. and Saedler,H. (1986) *Mol. Gen. Genet.*, **202**, 429–434.
Studier,F.W., Rosemberg,A.H., Dunn,J.J. and Dubendorff,J.W. (1990) *Methods Enzymol.*, **185**, 60–89.
Tanikawa,J., Yasukawa,T., Enari,M., Ogata,K., Nishimura,Y., Ishii,S. and Sarai,A. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 9320–9324.
Tice-Baldwin,K., Fink,G.R. and Arndt,K.T. (1989) *Science*, **246**, 931–935.
Tullius,T.D. and Dombroski,B.A. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 5469–5473.
Urao,T., Yamaguchi-Shinozaki,K., Urao,S. and Shinozaki,K. (1993) *Plant Cell*, **5**, 1529–1539.
Vieira,J. and Messing,J. (1982) *Gene*, **19**, 259–268.
Weston,K. (1992) *Nucleic Acids Res.*, **20**, 3043–3049.

Received on August 24, 1994; revised on January 19, 1995