Differences between Plant and Animal Myb Domains Are Fundamental for DNA Binding Activity, and Chimeric Myb Domains Have Novel DNA Binding Specificities*

(Received for publication, September 9, 1996, and in revised form, October 23, 1996)

Christopher E. Williams‡ and Erich Grotewold§
From Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724-2212 and Dowling College, Oakdale, New York 11768-1999

Several Myb domain proteins have been identified in plants, in which they play important regulatory roles in specific cellular processes. Plant and animal Myb domains have significant differences, but how these differences are important for function is not yet understood. The P gene encodes a Myb domain protein that activates a subset of flavonoid biosynthetic genes in maize floral organs. P and v-Myb bind different DNA sequences in vitro. Here we show that the Myb domain is solely responsible for the sequence-specific DNA binding activity of P, which binds DNA only in the reduced state. Differences in the DNA binding domains of v-Myb and P, which are conserved among animal and plant Myb domains, are fundamental for the high affinity DNA binding activity of these proteins to the corresponding binding sites but are not sufficient for the distinct DNA binding specificities of P and v-Myb. We conclude that significant structural differences distinguish plant from animal Myb domains. A chimeric Myb domain with a novel DNA binding specificity was created by combining Myb repeats of P and v-Myb. This approach could be used to artificially create novel Myb domains and to target transcription factors to genes containing specific promoters or to modify Myb-mediated interactions with other cellular factors.

In plants, a large number of Myb domain proteins have been identified and shown to participate in a variety of important cellular functions (11, 14, 21, 22). Petunia contains a Myb gene family with at least 40 members (12). Arabidopsis thaliana, more than 100 Myb-homologous sequences have been identified (23), and at least 20 different Myb domain genes are expressed in Arabidopsis flowers. The maize Myb domain proteins P and C1 regulate the accumulation of related, but different, flavonoid-derived pigments (24) through the activation of overlapping sets of flavonoid biosynthetic genes (25, 26). Maize flavonoid biosynthesis is probably one of the best described plant metabolic pathways, and studies of several putative target genes for the P and C1 proteins have yielded important information (27). In particular, P activates transcription of A1, one of the genes regulated by both P and C1, by binding to the P-responsive elements identified in the A1 gene promoter (26). Studies have also shown that the DNA consensus sequence CCTAACC recognized by P (26) differs substantially from the sequence C/TAACGG recognized by animal Myb proteins (26, 28, 29).

Myb domains are usually formed by two or three imperfect 51- or 52-residue repeats (R1, R2, and R3). Each repeat encodes three α helices, with the second and third helices forming a helix-turn-helix (HTH) structure when bound to DNA, similar to motifs found in the λ repressor and homeodomain proteins (30, 31–33). R2 and R3 are sufficient for sequence-specific DNA binding (31, 34, 35), and although the c-myb proto-oncogene contains all three repeats, retroviral versions of c-Myb (v-Myb) contain only R2 and R3 (36), as do most plant Myb domain proteins (12).

The structure of the mouse c-Myb R2-R3 DNA binding domain bound to DNA has been determined by NMR. The two Myb repeats R2 and R3 are closely packed in the major groove, with the third helix of each repeat making contact with DNA such that the recognition helices contact each other to bind to DNA in a cooperative manner (33). P and several other plant Myb domain proteins differ from v-Myb and many other animal Myb domain proteins in several aspects, which include the presence of an additional leucine residue between the second and third helices of R2 and residue differences at conserved positions in each of the two DNA recognition helices. Whereas the NMR studies of the c-Myb R2-R3 DNA binding domain offered important insight into Myb domain structure, they did not address the significance of the conserved differences between Myb domains of different kingdoms in Myb domain

*To whom correspondence should be addressed: Cold Spring Harbor Laboratory, 1 Bungtown Rd., Cold Spring Harbor, NY 11724. Tel.: 516-367-6860; Fax: 516-367-8369; E-mail: grotewol@cshl.org.

The abbreviations used are: R, repeat; HTH, helix-turn-helix; REDOX, reduced state; PCR, polymerase chain reaction; TBE, Tris borate/EDTA; wt, wild type; DTT, dithiothreitol; NEM, N-ethylmaleimide; SB, sonication buffer; WB, washing buffer.

This paper is available online at http://www-jbc.stanford.edu/jbc/

Printed in U.S.A.
structure or sequence-specific DNA recognition.

Here we show that the highly conserved Myb domains of P and v-Myb are sufficient for conferring differential DNA binding properties to these proteins in vitro. P requires both of its Myb repeats and must be in the reduced state (REDOX) to bind DNA efficiently. Differences in the DNA recognition helices of P and v-Myb, conserved among plant and animal Myb domains, respectively, play fundamental roles in permitting these proteins to bind to their corresponding target sites in vitro, providing the first experimental evidence that plant and animal Myb domains are structurally different. A chimeric Myb domain with a novel DNA binding specificity was created by combining Myb repeats of P and v-Myb in a specific orientation. This illustrates the possibility of artificially designing and creating novel Myb domains, useful for targeted gene regulation by binding to specific promoter elements or by Myb-mediated interactions with other cellular factors, providing convenient tools for both basic and applied studies.

**EXPERIMENTAL PROCEDURES**

**DNA Manipulations**—Standard cloning techniques were used (37). Mutant DNA sequences were generated by site-directed mutagenesis (38) in pBluescript vectors (Stratagene). All mutations were sequenced by the dideoxy method with Sequenseq (U. S. Biochemical Corp.).

**Bacterial Expression of Proteins**—The Myb domain of P (P<sub>myb</sub>) was obtained by introduction of a BamHI site at the 5'-end primer for the cloning of Myb<sub>myb</sub>-coding sequences corresponding to residues 600–604 of E. coli-expressed P virtutively. To express the corresponding proteins in bacteria, the P<sub>myb</sub>-coding sequences were fused to a polyhistidine tag (N10His-P<sub>myb</sub>) to yield P<sub>N10His-P<sub>myb</sub></sub>. This was achieved by amplification of each Myb repeat independently, cloning into an intermediate plasmid and sequencing the inserts. The N10His-P<sub>myb</sub>-coding DNA fragment was then amplified from the intermediate plasmid and cloned into NcoI sites of pET19b (Invitrogen). The Myb domain of v-Myb (v-Myb<sub>myb</sub>) was obtained by PCR. The 5'-primer v-Myb<sub>5pET</sub> contains a sequence corresponding to residues 123–130 in v-Myb<sub>myb</sub>. Primer v-Myb<sub>3pET</sub> contains a sequence corresponding to the last 8 residues of the Myb domain (residues 123–130 in v-Myb<sub>myb</sub>) and 5'-primer pET<sub>19b</sub> contains a sequence corresponding to residues 600–604 of P<sub>myb</sub>.

**To generate Myb<sub>N10His-P<sub>myb</sub></sub>**, each repeat was independently amplified by PCR. For the generation of P1 (R<sub>2</sub> Myb repeat of P), a 5'-primer was used (p5pET), which has the same characteristics as v-Myb<sub>5pET</sub> but with P sequences (9) instead of v-Myb sequences, and a 3'-primer corresponding to residues 59–64 of P (9) in which the Asp<sup>39</sup> was changed to Glu to create an EcoRV site. Similarly, the second repeat of P<sub>myb</sub> was amplified using the same 3'-primer as before, but a 5'-primer corresponding to residues 72–77 (39), which contained a PmlI site, was introduced without any change in the amino acid sequence. The two PCR fragments were purified from agarose, digested with the corresponding enzymes, ligated, and used to generate the change in P of Leu<sup>55</sup> to Glu. The P<sub>myb</sub>-coding sequences were cloned as above.

**To express the corresponding proteins in bacteria**, the Xhol-BamHI inserts were introduced into the corresponding sites from pET19b (Novagen). The clones carrying the inserts were transformed into the BL21(DE3)PlyS Escherichia coli strain. Protein expression was induced by addition of inducers (I<sub> IPTG</sub>) of 0.3–0.5) with 1 ml isopropyl-1-thio-β-D-galactopyranoside for 2 h at 37 °C.

**To express the recombinant proteins** in bacteria, the Xhol-BamHI inserts were introduced into the corresponding sites from pET19b (Novagen). The clones carrying the inserts were transformed into the BL21(DE3)PlyS Escherichia coli strain. Protein expression was induced by addition of inducers (I<sub> IPTG</sub>) of 0.3–0.5) with 1 ml isopropyl-1-thio-β-D-galactopyranoside for 2 h at 37 °C.

**To generate the change in P of Leu<sup>55</sup> to Glu**, the first Myb repeat of P<sub>myb</sub> was amplified with a 5'-primer containing the sequence corresponding to residues 59–64 of P (9) in which the Leu<sup>55</sup> was replaced by the codon corresponding to Thr<sup>56</sup>.

**To generate the change in P of Leu<sup>55</sup> to Glu**, the first Myb repeat of P<sub>myb</sub> was amplified with a 5'-primer containing the sequence corresponding to residues 59–64 of P (9) in which the Leu<sup>55</sup> was changed to Glu to create an EcoRV site. Similarly, the second repeat of P<sub>myb</sub> was amplified using the same 3'-primer as before, but a 5'-primer corresponding to residues 72–77 (39), which contained a PmlI site, was introduced without any change in the amino acid sequence. The two PCR fragments were purified from agarose, digested with the corresponding enzymes, ligated, and used to generate the change in P of Leu<sup>55</sup> to Glu. The P<sub>myb</sub>-coding sequences were cloned as above.

**To express the corresponding proteins in bacteria**, the Xhol-BamHI inserts were introduced into the corresponding sites from pET19b (Novagen). The clones carrying the inserts were transformed into the BL21(DE3)PlyS Escherichia coli strain. Protein expression was induced by addition of inducers (I<sub> IPTG</sub>) of 0.3–0.5) with 1 ml isopropyl-1-thio-β-D-galactopyranoside for 2 h at 37 °C.
To generate all other residue substitutions, site-directed mutagenesis (38) of the corresponding Myb domains cloned in pBluescript plasmids was used. All the mutant Myb domains were sequenced before the corresponding fragments were ligated to PET19b for expression in E. coli.

**Purification of PolyHis Proteins—**Proteins were expressed in 150-ml cultures as described above, harvested by centrifugation, and lysed in 5 ml of SB (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, and 100 µg/ml phenylmethylsulfonyl fluoride) by sonication with a microtip at maximal power for three pulses of 30 s in ice. Samples were centrifuged for 10 min at 12,000×g and filtered through two layers of Miracloth. RNAse A and DNase I were added to a final concentration of 1 mg/liter, and samples were kept on ice for 10 min. 0.5 ml of a 50% slurry of Ni resin (Qiagen) equilibrated with SB was added to each sample and incubated with rocking at 4 °C for 1 h. The resin was loaded onto a column, washed three times with 3 ml of SB, followed by two washes with 1.5 ml of WB (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 1% Tween 20, 10% glycerol, 5 mM β-mercaptoethanol, and 10 mM EDTA). Elution was carried out by three sequential additions of 0.5 ml of 100 mM imidazol in WB. Protein levels were estimated by SDS-polyacrylamide gel electrophoresis by Coomassie Blue and silver staining and comparison with immunoglobulin standards. The yield of high-purity recombinant protein obtained in these conditions is between 0.5 and 5 mg of protein/liter of culture, depending on the particular Myb domain recombinant protein obtained in these conditions.

**DNA Binding Reaction Incubations—**Incubations were carried out as described (26). Cloned fragments were labeled after cleavage with T4 DNA ligase and purified by preparative agarose gel electrophoresis (39). Ligated fragments were further separated by electrophoresis through 1.5-mm-thick 7% polyacrylamide gels (80:1 acrylamide:bisacrylamide) in 0.25 M TBE buffer at 40 V/cm for 55 min at 4 °C, unless otherwise indicated. Apparent Kd values were determined by incubation of fixed amounts of N10His-Pmyb protein with increasing amounts of APB1 probe (Fig. 1C) and determination of the bound and free oligonucleotide concentration by use of a PhosphorImager followed by Scatchard analysis. Studies for a particular protein-DNA complex were carried out independently at least three times.

Quantitations were done with a BAS2000 PhosphorImager (Fuji).

**Binding Site Selection—**The oligonucleotides used for these studies are the same as reported previously and were labeled in a similar way (26). For the first round of selection we made use of the polyhistidine tag present on all the proteins tested by incubating total E. coli extracts containing about 20 ng of the protein of interest with 0.4 µg of degenerate oligonucleotide. Once binding was completed, Ni²⁺ resin (Qiagen) equilibrated in binding buffer was added to the reaction, which was then incubated for 1 h in ice and washed three times with binding buffer. The resin was resuspended in 20 µl of water, and 1 µl was used for the generation of the probe for the next round by PCR (26). The four subsequent rounds of selection were done by excision of the shifted bands from preparative gel mobility shift assays as described (26) and amplification of the extracted DNA probes. After five rounds of selection, the PCR products were cut with BamHI and EcoRI and cloned into the corresponding sites of pBluescript KS−. Sequences were aligned on the basis of our previous studies of the sequences preferentially bound by P and v-Myb (26).

**RESULTS**

**DNA Binding Properties of the Myb Domain of P—**To determine whether the Myb domain is solely responsible for the DNA binding properties of P (26), the Myb domain of P (residues 1–122) was expressed as an amino-terminal polyhistidine fusion protein in E. coli (N10His-Pmyb). DNA binding activity of the fusion protein was then tested on wild-type and mutant forms of the previously identified element in the A1 promoter, which has two overlapping P binding sites (APB1; Ref. 26). Fig. 1A shows that N10His-Pmyb can bind effectively to APB1, and this binding is similar to that of the full-length P protein (26). Point mutations at either site I (APB2; Fig. 1C) or site II (APB3; Fig. 1C) slightly reduce binding (Fig. 1A, lanes 2 and 3, respectively), and binding to a mutant at both sites (APB5; Fig. 1C) was not detected (Fig. 1A, lane 4). Oligonucleotides con-
Structural Differences between Plant and Animal Myb Domains

taining only site I (PBS1; Fig. 1C), or site II (PBS2; Fig. 1C) efficiently compete binding of N10His-Pmyb to APB1 (Fig. 1B, lanes 4–7), as does APB1 itself (Fig. 1B, lanes 2 and 3), whereas oligonucleotides containing the corresponding mutant sites (PBS3 and APB5; Fig. 1C) do not compete the binding of N10His-Pmyb to APB1 (Fig. 1C, lanes 8–10). N10His-Pmyb binds APB1 in vitro with high affinity ($K_d$, 28 $\pm$ 3 nM; data not shown). Binding of the Myb domain of P to APB1 is severely reduced by mutations in the DNA recognition helices of the second or the third P Myb repeats, as well as by addition of monovalent cations such as NaCl to the reaction buffers (not shown). These results indicate that both Myb repeats of P are necessary and sufficient for high affinity sequence-specific DNA binding of P to the binding sites identified in the A1 promoter.

A highly conserved cysteine residue (C130 in c-Myb, C53 in v-Myb) is proposed to act as a molecular sensor. Efficient Myb-DNA binding was observed only when the proteins were in their reduced state, as indicated by the addition of excess DTT (Fig. 2, compare lanes 5 and 10). The irreversible effect of NEM on complex formation is only observed when the proteins are added in the presence of DTT. NEM is reversed by the addition of excess DTT (Fig. 2, compare lanes 6 and 16). The addition of NEM inhibited the formation of DNA-protein complexes, suggesting that NEM modifies the DNA recognition helices of both v-Myb and P (Fig. 2, lanes 9 and 10). The irreversible effect of NEM on complex formation is only observed when the proteins are in their reduced state. Further, the addition of NEM to proteins that have not been reduced previously with DTT, the inhibition by NEM is reversed by the addition of excess DTT (Fig. 2, compare lanes 6 and 16). Similar results were obtained using just the Myb domain of P and v-Myb (not shown). These results indicate that P must be in a reduced state to bind DNA, providing the first example of a plant Myb protein with similar REDOX requirements as animal Myb proteins (40–42).

The Myb Domains of P and v-Myb Are Sufficient to Confer Differential DNA Binding Activities to These Proteins—To understand whether the DNA binding properties of P are solely given by its Myb domain, or whether other regions of the domain contributed to sequence-specific DNA binding, we compared the binding specificities of the Myb domain of P with those of the intact P protein. The Myb domain of P and v-Myb were expressed in E. coli as polyhistidine fusion (N10His-Pmyb and N10His-cMyb, respectively). As previously determined for the intact P protein (26), the Myb domain of P does not bind the v-Myb binding sites present in the promoter of the mam-gene (APB1; Fig. 1C, lanes 9 and 10). The Myb domain of v-Myb does not bind the P binding sites of the promoter of the mam-gene (APB1; Fig. 1C, compare lanes 9 and 10). Further, the DNA binding capacities were compared using a site selection strategy on a random population of 26-mer oligonucleotides, as described previously (26). Fig. 3 shows the comparison of the sequences selected by N10His-Pmyb and P (Fig. 3, A and B) and by N10His-Mam-1 and v-Myb (Fig. 3, C and D). Overall, the corresponding Myb domains exhibit DNA binding preferences very similar to those of the intact proteins. These results demonstrate that the Myb domains of P and v-Myb are...
responsible for their different DNA binding preferences (26), and the contribution of the remainder of the respective proteins to the in vitro DNA binding properties is probably minimal. These results suggest that the highly conserved Myb domains of other plant Myb proteins are also sufficient for sequence-specific DNA binding.

Analysis of Differences between Plant and Animal Myb Domains—Having determined that the Myb domains of P and v-Myb are sufficient for their distinct DNA binding preferences, we investigated to what extent the differences present between plant and animal Myb domains, represented in P and v-Myb, have a role in DNA binding activity. The high degree of conservation in the DNA recognition helices among animal and plant Myb domains, as well as the differences between these two classes of Myb domain proteins, are illustrated in Fig. 4 (differences marked with dots and arrows). Several previous studies have investigated the effect of mutations of specific residues on the DNA binding activity of v-Myb and c-Myb (30, 31, 34, 43). The uniqueness of this study is that residue substitutions were designed based on the differences between plant and animal Myb domain proteins. Four residues are different in the R2 DNA recognition helices among animal and plant Myb domains, as well as the differences between these two classes of Myb domain proteins, are illustrated in Fig. 4 (differences marked with dots and arrows). Several previous studies have investigated the effect of mutations of specific residues on the DNA binding activity of v-Myb and c-Myb (30, 31, 34, 43). The uniqueness of this study is that residue substitutions were designed based on the differences between plant and animal Myb domain proteins. Four residues are different in the R2 DNA recognition helices among animal and plant Myb domains (Gln129, Glu132, His135, and Asn137; Fig. 4). We focused our studies on Glu132 of v-Myb, which corresponds to a leucine residue in all plant Myb proteins (Leu55 in P; Fig. 4) and which may play a role in base pair recognition, as well as in cooperative interaction between the two Myb repeats (33). The effect of mutations at Glu132, however, had not been previously studied. We investigated the effect of the change in v-Myb of Glu132 to Leu. The protein was expressed as a polyhistidine fusion (N10His-MybmybE132L) and tested for binding to P and v-Myb binding sites (Fig. 5, A and B). This mutation showed no significant effect on DNA binding to the mim-1 wt probe (Fig. 5A, bottom panel, compare lanes 2 and 3), indicating that its role in stabilizing the cooperative interaction of the two Myb repeats of v-Myb is minor, if significant at all. This mutant protein does not bind to the P binding sites present in the APB1 probe under the conditions tested (Fig. 5A, top panel, lane 3).

Site selection experiments carried out with N10His-MybmybE132L did not show any significant binding consensus difference between N10His-MybmybE132L and N10His-Mybmyb (not shown), supporting the idea that Glu132 does not play a role in sequence-specific DNA binding. The corresponding substitution introduced in the Myb domain of P (N10His-PmybL55E) inhibits DNA binding activity of P to the APB1 probe (Fig. 5A, top panel, compare lanes 7 and 8) and does not allow P to bind to the Myb binding sites (Fig. 5A, bottom panel, lane 8). In fact, the effect of this amino acid substitution in P has a more dramatic effect on DNA binding activity than the change of a conserved asparagine residue (Asn109, Fig. 4) to leucine (N10His-PmybN109L) (Fig. 5A, lane 9). Site selection experiments confirmed that N10His-PmybL55E has a general inability to bind DNA, since no sequences could be selected by this protein from a random population of oligonucleotides (not shown). We then investigated the effect of substitutions in the second recognition helix of v-Myb with the corresponding residues in P. Initially, we changed all three residues that differ between P and v-Myb in this DNA recognition helix (Ala180, Val181, and His184; Fig. 4, arrows) to the corresponding residues of P (Fig. 4). We focused on Ala180 in v-Myb with the corresponding Glu found in all plant Myb proteins (MybmybAVH-EIY). This substitution resulted in a complete loss of DNA binding activity by v-Myb to the mim-1 wt site (Fig. 5A, bottom panel, lane 4). The substitution of Ala180 in v-Myb with the corresponding Glu found in all plant Myb proteins or the substitution of Val181 and Tyr184 with the corresponding residues of P (Fig. 4) completely abolished binding of v-Myb to mim-1 wt, probe (Fig. 5A, bottom panel, compare lanes 2 and 3), indicating that its role in stabilizing the cooperative interaction of the two Myb repeats of v-Myb is minor, if significant at all. This mutant protein does not bind to the P binding sites present in the APB1 probe under the conditions tested (Fig. 5A, top panel, lane 3).
without a significant gain of DNA binding activity on APB1 (not shown). However, we have not tested whether these residue changes abolish binding of v-Myb only to the mim-1 wt site or whether they also affect the DNA binding activity of v-Myb in a more general way.

We next investigated whether a compensatory substitution in R₂ is required for the proteins containing substitutions in R₃ to bind DNA. In particular, we found striking the observation that animal Myb domains have the charged Glu132 residue in R₂ that corresponds to a leucine residue in plants, whereas most plant Myb domains have a charged residue at a similar position in R₃, replaced by a hydrophobic residue in animal Myb domains (Ala180 in v-Myb; Fig. 4). Thus, we combined the E₁₃₂L substitution with the A₁₈₀E substitution (N₁₀His-Myb₁₂₈₂E₁₃₂L₁₈₀E) and the E₁₃₂L substitution with the substitution of the three residues different between P and v-Myb in the DNA recognition helix of R₃ (N₁₀His-Myb₁₂₈₂E₁₃₂L₁₈₀E). No significant binding to the mim-1 wt or the APB₁ probes was detected when these mutants were tested for DNA binding activity (Fig. 5, A, lanes 5 and 6, and B), indicating that spacing between helices 2 and 3 in R₂ of v-Myb is critical for DNA binding to the mim-1 wt sequence.

**Generation of Myb Domains with Altered DNA Binding Preferences by Exchange of Myb Repeats between P and v-Myb—**

The Myb domain can be thought to be composed of two modules, each represented by one Myb repeat. According to the current understanding of the structure of the Myb domain (33), each module (i.e. Myb repeat) would contribute to sequence-specific DNA binding by recognizing half-sites of the binding site (30). Unlike other bipartite or bimodular DNA binding domains (for example, the POU domain; Ref. 44), each Myb repeat does not bind DNA with high affinity by itself, and cooperative interactions between the repeats are required (33). We investigated whether Myb proteins with novel DNA binding specificities could be created by generating chimeric Myb domains with repeats from Myb proteins with different DNA binding properties. For that purpose, we combined R₂ of P with R₃ of v-Myb (see "Experimental Procedures") (N₁₀His-Myb₁₂₈₂P₁₈₀E; Fig. 6A). This chimeric Myb domain binds APB₁ at least as efficiently as N₁₀His-P₁₈₀E (Fig. 6B, compare lanes 1 and 3). Binding of N₁₀His-Myb₁₂₈₂P₁₈₀E to mim-1 wt is also observed under conditions in which N₁₀His-P₁₈₀E binding is undetectable (Fig. 6B, compare lanes 1 and 11), although this binding is weaker than that of N₁₀His-Myb₁₂₈₂P₁₈₀E (Fig. 6B, compare lanes 11 and 12). Mutant versions of the corresponding binding sites are not bound by this chimeric Myb domain, indicating that binding is specific (Fig. 6B, lanes 7 and 15).

To demonstrate further that the N₁₀His-Myb₁₂₈₂P₁₈₀E chimeric Myb domain has DNA binding preferences different from P and...
v-Myb, site selection experiments were carried out. The percentages of nucleotides in the DNA sequences selected by the chimeric Myb domain are shown in Fig. 7A. The consensus sequence recognized by the chimeric Myb appears to be a composite of P and v-Myb consensus binding sites. This composite binding site does not have an A at position -4, as is preferred by P and its Myb domain (Fig. 3, A and B). The lack of nucleotide preference at that position, however, is reminiscent of the consensus binding site of v-Myb (Fig. 3, C and D). The lack of a marked preference for any nucleotide at position -3 also resembles the v-Myb binding site. At positions -2 and -1, the binding preference of the chimeric Myb domain is very similar to that of v-Myb. In fact, no sequences selected by N10His-Myb12M62 contained a T at position -1, in contrast to the sequences selected by P (Fig. 3, A and B) (26). The composite site at positions 0 and +1 (A and C residues, respectively) of the composite site resembles both the P and v-Myb binding sites, which are the same at those two positions. At position +2, N10His-Myb12M62 slightly prefers a C residue, similar to the preference seen for P; v-Myb was shown previously to be flexible at position +2 (29). In our experiments, however, v-Myb seems to prefer a G at this site (Fig. 3, C and D). At position +3, the chimeric Myb domain is much like P, showing no particular preference.

These results agree with the current understanding of the way in which the Myb domain contacts DNA. The R1 Myb repeat recognizes positions -1, 0, and +1 in a very sequence-specific fashion, whereas the R2 Myb repeat recognizes position +3 (33, 35). Our results, however, indicate that the influence of the R2 Myb repeat extends to position -4, whereas both the R2 and R4 Myb repeats seem to have some influence on position +2. The strong binding of the chimeric N10His-Myb12M62 protein to APB1 is expected, since one of the P binding sites present in APB1 (ACCAACCT; Fig. 1C) matches perfectly the deduced binding consensus for the chimera (Fig. 7A). The weaker binding to mim-1 wt is probably due to the low preference of the chimera for a G residue at position +2.

These results also suggest that there may be DNA sequences that can be bound by the chimeric Myb domain but not by P or v-Myb. One such sequence, DI8, was identified from the clones selected by N10His-Myb12M62 (Fig. 7D). DI8 is recognized by N10His-Myb12M62 but not by N10His-v-Myb or N10His-Pmyb (Fig. 7B, lanes 1–3). The binding of N10His-Myb12M62 can be specifically competed by wild-type versions of the P or v-Myb binding sites but not by mutant counterparts, indicating the specificity of the binding (Fig. 7C). The N10His-Myb12M62 protein does not bind DI8 (Fig. 7B, lane 4), indicating that the leucine residue found in the P1Myb2 chimera (L55) is not responsible for this new binding specificity of P1Myb2.

When a chimeric protein containing the first Myb repeat of v-Myb and the second Myb repeat of P (N10His-Myb1P2; Fig. 6A) was tested in identical conditions, no binding to either APB1 or mim-1 wt was observed (Fig. 6B, lanes 2 and 10). When this protein was tested on site selection experiments, no sequences were selected (not shown). This indicates that N10His-Myb1P2 is unable to bind DNA in general, not just the P and v-Myb binding sites. We cannot rule out, however, that the structure of Myb1P2 is affected by the combination of these Myb repeats, explaining its inability to interact with DNA.

**DISCUSSION**

Our studies were aimed at studying the structure-function relationship of Myb domains separated by over 600 million years of evolution. We used the maize Myb domain protein P and the avian myeloblastosis virus v-Myb protein as models to investigate how similarities and differences between plant and animal Myb domains affect their DNA binding properties and how novel Myb domains can be created by combining regions from these structurally divergent Myb domains. Previous studies showed that the maize Myb domain protein P has a different DNA binding preference from v-Myb and other animal Myb proteins, providing convenient tools to assay the sequence-specific DNA binding activities of these proteins (26).

The Myb domain of P, as well as the one from v-Myb and from most other plant Myb domain proteins, is formed by two HTH motifs, one corresponding to each Myb repeat. Our studies show that both HTH motifs of P are necessary and sufficient...
for high affinity DNA binding, similar to animal Myb proteins (33, 35) but different from a plant Myb protein containing a single Myb repeat (45). The Myb domain of P is sufficient for sequence-specific DNA binding, and regions outside the Myb-homologous DNA binding domain have little influence on the DNA sequences preferentially bound by P in vitro. Moreover, the Myb domains of P and v-Myb display similar differences in DNA binding preference, as originally demonstrated for the intact proteins (26), indicating that differences within the Myb domain are responsible for the different DNA binding specificities of these two proteins. Our studies indicate that P requires a reduced state to bind to DNA in vitro. This is similar to what has been found with v-Myb and c-Myb, in which a conserved cysteine residue (Cys130 in c-Myb) could function as a molecular sensor for a REDOX regulatory mechanism (33, 40–42). We do not know whether a REDOX mechanism regulates P function in vivo (26), but it is tempting to speculate that such a mechanism could control the tissue-specific function of other P alleles (46). Alternatively, this could reflect a general requirement of Myb domains to bind DNA and not a specific mechanism that regulates in vivo activity of particular Myb proteins.

Plant and animal Myb domains differ in a number of aspects, which include several residue differences in the DNA recognition helices of R2 and R3 (Fig. 4). We approached this problem by using a strategy fundamentally different from previous studies (30, 31, 34, 43), which consisted of substituting specific residues in the Myb domain of v-Myb for the corresponding residues present in P and other plant Myb domain proteins, to study their effect on DNA binding activity. The bottom line of these experiments is clear. With the exception of Glu132, all the residue substitutions tested (Fig. 4) had a dramatic inhibitory effect on the DNA binding activity of v-Myb to its binding sites. Moreover, none of the single or multiple residue substitutions tested allowed v-Myb to bind to the P binding sites, suggesting that they are not sufficient for altering DNA binding specificity of Myb domains. Perhaps the specificity of DNA binding by P and v-Myb is influenced by regions outside of the DNA recognition helices, as happens in other proteins containing helix-turn-helix motifs (1). Alternatively, unique interactions could be required between the residues that are conserved among plant or animal Myb domain protein family members, respectively, which are different between the two types of Myb domains. We do not know whether the effect of these residue substitutions is the result of a structural change that prevents DNA binding or whether they affect residues involved in the interaction with DNA. In any case, these findings indicate that these residues play fundamental roles in the DNA binding activity of Myb domains, adding fundamental information to previous structural studies (33, 47, 48). Glu132 provided an exception in many aspects. Glu132 is the only residue difference between plant and animal Myb domain DNA recognition helices that the NMR structure of the c-Myb R2–R3 DNA binding domain implicated in possible base pair contact and cooperative interaction between Myb repeats (33). Yet, replacing Glu132 with the leucine found in all plant Myb domains (Fig. 4) has no effect on the sequence-specific binding of v-Myb to DNA, whereas the corresponding change in P of Leu52 to Glu abolishes the DNA binding activity of P. These results indicate that Glu132 is not important in stabilizing the cooperative interaction between R2 and R3 of c-Myb, or in specific base pair contacts, as was proposed previously (33). In addition, these findings provide evidence that the presence of a conserved leucine, and not glutamic, residue at this position in plant Myb domains is fundamental for DNA binding activity.

All plant Myb domain proteins characterized to date contain an additional leucine residue inserted between the first and second helices of the R3 repeat (Fig. 4), which is absent in animal Myb domains. We investigated the significance of the additional leucine residue by adding it to the Myb domain of v-Myb at the equivalent position. The addition of this residue has an inhibitory effect on the DNA binding activity of v-Myb to its binding sites in the mim-1 wt sequence (Fig. 5, A and B). Since this residue is outside the DNA recognition helices, we conclude that the spacing between the helices in plant and animal Myb domains is critical. The three helices in the first Myb repeat of plant proteins are probably packed differently from the corresponding helices of animal Myb domains.

We investigated the possibility of generating Myb domains with novel DNA binding characteristics by combining components of the Myb domains from proteins with different DNA binding activities. The chimera Myb domain P1Myb2, containing the R2 of P and R3 of v-Myb, binds both the P binding sites present in APB1 that are normally bound by P, but not by v-Myb, as well as the Myb binding sites in mim-1 wt, normally bound by v-Myb, but not by P. The results obtained from site selection experiments carried out with the P1Myb2 Myb domain suggest that this chimeric protein recognizes a site that is a composite between the DNA binding consensus determined for v-Myb and the one determined for P but different from either one. Thus, the P1Myb2 protein has a new DNA binding preference. The prediction that the P1Myb2 chimeric Myb domain should bind DNA sequences that are not recognized by either P or v-Myb was fulfilled by the identification of clone DI8. DI8, which contains the core sequence CTTAACTC, is bound with high affinity only by P1Myb2 but not by P or v-Myb. Contrary to what happens with P1Myb2, Myb1P2 containing the first Myb repeat of v-Myb and the second of P has no in vitro DNA binding activity. These results indicate that within certain limitations, new Myb domains with novel DNA binding specificities can be created by combining Myb repeats of proteins with different DNA binding preferences. In addition to the described DNA binding activity, Myb domains can also mediate protein-protein interactions (49, 50). Thus, novel Myb domains generated in this fashion could be used to target transcription factors to genes containing specific promoters and to modify Myb-mediated interactions with other cellular factors, providing useful tools for both basic and applied studies.

In light of the effect of the previously described mutations in the DNA binding activity of v-Myb, it is interesting that the P1Myb2, but not the Myb1P2, Myb domain chimera binds DNA efficiently. The observation that certain combinations of Myb repeats are unable to bind DNA indicates that corresponding Myb repeats of P and v-Myb are not functionally equivalent. Several reports indicate that the carboxyl-terminal region of R2 of animal Myb domains has a disordered structure, which adopts a helical conformation only on binding to DNA (42, 47, 48). We do not know the structure of the corresponding region of the Myb domain of P or other plant Myb proteins, but interestingly, the helical propensity (53) of the carboxyl-terminal portion of R2 from P is significantly higher than the corresponding region of c-Myb and comparable with the DNA recognition helices of the R2 Myb repeat from plants or animals. We can imagine that Myb domains containing the R2 Myb repeat from P (such as P1Myb2) would require less energy to adopt the double helix-turn-helix structure and will have a stronger DNA binding affinity than Myb domains containing the first repeat of v-Myb (such as Myb1P2). Whereas this is certainly the situation for the chimeric proteins tested in this study, intramolecular interactions specific for each type of Myb domain could play additional stabilizing roles in the formation of stable protein-DNA complexes, interactions probably not conserved in the Myb1P2 chimera.
Together, these results provide the first experimental evidence that plant and animal Myb domains have significant structural differences, despite extensive sequence similarity. In addition, they indicate a need to structurally analyze plant Myb domains to understand the evolutionary mechanisms involved in the high conservation of this domain in a large number of plant transcription factors.

Acknowledgments—We thank Ben Bowen for helpful discussions throughout this work, Michele Cleary and Leemor Joshua-Tor for help with computer structural analyses, Hong Qian for the helical propensity program, George Tokiwa for help with PhosphorImager analyses, Michele Cleary and Patty Springer for comments and critical reading of the manuscript.

REFERENCES

1. Pabo, C. O., and Sauer, R. T. (1992) Annu. Rev. Biochem. 61, 1053–1095
2. Tice-Baldwin, K., Pink, G. R., and Arndt, K. (1989) Science 246, 931–935
3. Stober-Grasser, U., Brydolf, B., Bin, X., Grasser, F., Pirtel, R. A., and Lipsick, J. S. (1992) Oncogene 7, 589–596
4. Oh, R., Mecocci, D., Hirani, B., Denkhaese, G. J., Zhang, X., Burke, J. D., Turner, K., and Gould, K. L. (1994) EMBO J. 13, 471–483
5. Amaravadi, L., and King, M. W. (1994) Oncogene 9, 971–974
6. Peters, C. W. B., Sippel, A. E., Vingron, M., and Klempnauer, K. H. (1987) EMBO J. 6, 3085–3090
7. Paz-Ares, J., Ghosal, D., Sivakumaran, S., Esch, J., and Marks, M. (1992) Cell 70, 481–496
8. Peters, C. W. B., Sippel, A. E., Vingron, M., and Klempnauer, K. H. (1987) EMBO J. 6, 3553–3558
9. Marrocco, A., Wissenbach, M., Becker, D., Paz-Ares, J., Sauder, H., Salamini, F., and Rohde, W. (1989) Mol. & Gen. Genet. 216, 183–187
10. Grotewold, E., Athma, P., and Peterson, T. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 4587–4591
11. Jackson, D., Cullinan-Macia, F., Prescott, A. G., Roberts, K., and Martin, C. (1991) Plant Cell 3, 115–125
12. Oppenheimer, D. G., Herman, P. L., Sivakumaran, S., Esc, J., and Marks, M. D. (1991) J. Biol. Chem. 266, 453–463
13. Avila, J., Nieto, C., Cañas, L., Benito, M. J., and Paz-Ares, J. (1993) Plant J. 3, 553–562
14. Li, S. F., and Parish, R. W. (1995) Plant J. 8, 963–972
15. Gabler, F., Kulla, R., Roberts, J. K., and Jacobsen, J. V. (1995) Plant Cell 7, 1879–1891
16. Gerondakis, S., and Bishop, M. (1986) Mol. Cell. Biol. 6, 3677–3684
17. Gonda, J. T., Gough, N. M., Dunn, A. B., and de Blaquiere, J. (1985) EMBO J. 4, 2003–2008
18. Slamon, D. J., Boone, T. C., Murdock, D. C., Keith, D. E., Press, M. F., Larson, R. A., and Souza L. M. (1986) Science 233, 347–351
19. Nemura, N., Takahashi, M., Matsui, M., Ishii, S., Dute, T., Sasamoto, S., and Ishizaki, R. (1988) Nucleic Acids Res. 16, 11075–11089
20. Thompson, M. A., and Ramsay, R. G. (1995) Bioessays 17, 341–350
21. Ness, S. A., Markwell, A., and Graf, T. (1989) Cell 59, 1115–1125
22. Noda, K.-I., Glover, B. J., Linstead, P., and Martin, C. (1994) Nature 369, 661–664
23. Sablowski, R. W. M., Moyano, E., Cullinan-Macia, F. A., Schuch, W., Martin, C., and Bevan, M. (1994) EMBO J. 13, 128–137
24. Solano, R., Nieto, C., Avila, J., Canas, L., Diaz, I., and Paz-Ares, J. (1995) EMBO J. 14, 1773–1784
25. Koes, R. E., Quattrocchio, F., and Mol, J. N. M. (1994) Bioessays 16, 123–132
26. Styles, E. D., and Ceska, O. (1989) Nucleic Acids Res. 34, 227–237
27. Grotewold, E., Drummond, B., Bowen, B., and Peterson, T. (1994) Cell 76, 543–553
28. Donner, H. K., Robbins, T. P., and Jorgensen, R. A. (1991) Annu. Rev. Genet. 25, 173–199
29. Howe, K. M., and Watson, R. J. (1991) Nucleic Acids Res. 19, 3913–3919
30. Westen, K. (1992) Nucleic Acids Res. 20, 3043–3049
31. Frampton, J., Gibson, T. J., Ness, S. A., Doderlein, G., and Graf, T. (1991) Protein Eng. 4, 891–901
32. Peters, C. W. B., Sippel, A. E., Vingron, M., and Klempnauer, K. H. (1987) EMBO J. 6, 3085–3090
33. Avila, J., Nieto, C., Canas, L., Benito, M. J., and Paz-Ares, J. (1993) Plant Cell 3, 115–125
34. Oppenheimer, D. G., Herman, P. L., Sivakumaran, S., Esc, J., and Marks, M. D. (1991) J. Biol. Chem. 266, 453–463
35. Koes, R. E., Quattrocchio, F., and Mol, J. N. M. (1994) Mol. & Gen. Genet. 242, 243–250
36. Klempnauer, K.-H., Ramsay, G., Bishop, J. M., Moscovich, M. G., Moscovic, C., McGarth, J. P., and Levinsen, A. D. (1983) Cell 32, 345–355
37. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
38. Kunzel, T. D., Roberts, J. D., and Zakour, R. A. (1987) Methods Enzymol. 154, 367–382
39. Garcia, A., LaMontagne, K., Reavis, D., Stober-Grasser, U., and Lipsick, J. S. (1991) Oncogene 6, 265–273
40. Grotewold, E., LaMontagne, K., Whittaker, L., Stohr, S., and Lipsick, J. S. (1991) Oncogene 7, 1005–1009
41. Guehmann, S., Vorbrueggen, G., Kalkbrenner, F., and Moelling, K. (1992) EMBO J. 11, 3458–3463
42. Myrset, A. H., Bostad, A., Jamin, N., Lirsac, P.-N., Toma, F., and Gabrielsen, O. S. (1993) EMBO J. 12, 4635–4638
43. Katsuki-Ishii, C., Sarai, A., Sawazaki, T., Nakagoshi, H., Hirono, M., Ogata, K., Nishimura, Y., and Ishii, S. (1990) J. Biol. Chem. 265, 19990–19995
44. Sturm, R., and Herr, W. (1988) Nature 336, 601–604
45. Baranowski, N., Frohberg, C., Prat, S., and Willmitzer, L. (1994) EMBO J. 13, 5383–5392
46. Chopra, S., Athma, P., and Peterson, T. (1996) Plant Cell 8, 1149–1158
47. Jamin, N., Gabrielsen, O. S., Gilles, N., Lirsac, P.-N., and Toma, F. (1993) Eur. J. Biochem. 216, 147–154
48. Carr, M. D., Wollborn, U., McIntosh, T. A., McCormick, J. E., Bauer, C. J., Klempnauer, K.-H., and Feene, J. (1996) Eur. J. Biochemistry 235, 721–735
49. Goff, S. A., Cone, K. C., and Chandler, V. L. (1992) Genes Dev. 6, 864–875
50. Mink, S., Kerber, U., and Klempnauer, K.-H. (1996) Mol. Cell. Biol. 16, 1316–1325
51. Klempnauer, K.-H., Gonda, T. J., and Bishop, J. M. (1982) Cell 31, 453–463
52. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
53._LESSFOO
Differences between Plant and Animal Myb Domains Are Fundamental for DNA Binding Activity, and Chimeric Myb Domains Have Novel DNA Binding Specificities
Christopher E. Williams and Erich Grotewold

J. Biol. Chem. 1997, 272:563-571.
doi: 10.1074/jbc.272.1.563

Access the most updated version of this article at http://www.jbc.org/content/272/1/563

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 52 references, 14 of which can be accessed free at http://www.jbc.org/content/272/1/563.full.html#ref-list-1