Novel Anther-Specific myb Genes from Tobacco as Putative Regulators of Phenylalanine Ammonia-Lyase Expression

Seungchan Yang1, Justin P. Sweetman2, Sasan Amirsadeghi, Medhi Barghchi, Alison K. Huttly, Won-Il Chung, and David Twellem

Department of Biological Sciences, Korea Advanced Institute of Science and Technology, 373–1, Kusondong, Yusong-gu, Taejon, Korea (S.Y., W.-I.C.); Department of Biology, University of Leicester, University Road, Leicester LE1 7RH, United Kingdom (J.P.S., S.A., D.T.); Department of Biological Sciences, DeMontfort University, Scraptoft Campus, Leicester LE1 9BH, United Kingdom (M.B.); and Institute for Arable Crops Research-Long Ashton Research Station, Department of Agricultural Sciences, University of Bristol, Long Ashton BS41 9AF, United Kingdom (A.K.H.)

Two cDNA clones (NtmybAS1 and NtmybAS2) encoding MYB-related proteins with strong sequence similarity to petunia (Petunia hybrida) PhMYB3 were isolated from a tobacco (Nicotiana tabacum cv Samsun) pollen cDNA library. Northern blot and in situ hybridization revealed that NtmybAS transcripts are specifically expressed in both sporophytic and gametophytic tissues of the anther including tapetum, stominium, vascular tissue, and developing pollen. Random binding site selection assays revealed that NtMYBAS1 bound to DNA sequences closely resembling consensus MYB binding sites MBSi and MBSII, with a higher affinity for MBSi. Transient expression analyses of the N-terminal MYB domain demonstrated the presence of functional nuclear localization signals, and full-length NtMYBAS1 was able to activate two different phenylalanine ammonia-lyase promoters (PALA and gPAL1) in tobacco leaf protoplasts. Similar analysis of truncated NtmybAS1 cDNAs identified an essential, C-terminal trans-activation domain. Further in situ hybridization analyses demonstrated strict co-expression of NtmybAS and gPAL1 in the tapetum and stominium. Despite abundant expression of NtmybAS transcripts in mature pollen, gPAL1 transcripts were not detectable in pollen. Our data demonstrate that NtMYBAS1 is a functional anther-specific transcription factor, which is likely to be a positive regulator of gPAL1 expression and phenylpropanoid synthesis in sporophytic, but not in gametophytic, tissues of the anther.

Male gametogenesis provides a tractable system to investigate cell differentiation, cell-cell interactions, and the regulation of cell-specific gene expression (McCormick, 1993; Twell et al., 1998). The male gametophytes, or pollen grains, of flowering plants develop within the anther and depend on complex interactions between sporophytic and gametophytic cells. The vital role of sporophytic anther tissues during early microgametogenesis is directly illustrated by experiments in which tapetal cells were genetically ablated (Mariani et al., 1990), and from the large number of male-sterile mutants arising from aberrant tapetal cell development (van der Meer et al., 1992; Chaudhury, 1993; Preuss, 1995; Aarts et al., 1997). In a similar manner, gametophytic expression of certain pollen and anther-specific genes is essential for pollen development and fertility (Muschietti, et al., 1994; Xu et al., 1995). Although a significant number of genes specifically expressed in anthers and/or pollen have been identified, details of their regulation in sporophytic and gametophytic tissues remain largely unknown (Twell, 1994).

Given the central role of transcription factors in the regulation of development, the identification of anther-/pollen-specific genes encoding transcription factors is an important step toward understanding regulatory mechanisms that operate during male gametogenesis. Several genes encoding putative transcription factors specifically expressed in anthers and/or pollen have been isolated; however, their functional activity and/or putative target genes remain to be demonstrated. These include pollen-specific genes encoding zinc finger proteins of the LIM class from sunflower and tobacco (Nicotiana tabacum cv Samsun; Baltz et al., 1992; Sweetman et al., 2000); anther and pollen-specific MADS box transcription factors, DEFH125 (Zachgo et al., 1997) and ZmMADS2 (Heuer et al., 2000), from snapdragon (Antirrhinum majus) and maize (Zea mays), respectively; and a family of anther-specific zinc finger proteins from petunia (Petunia hybrida; Kobayashi et al., 1998).

The myb gene family represents the largest transcription factor gene family in plants, with about

---

1 This work was supported by the Biotechnology and Biological Sciences Research Council (studentship to J.P.S.) and by Iran's Agricultural Research, Education, and Extension Organization (to S.A.). Work carried out in Long Ashton was supported by the British Council (fellowship to S.Y.).

2 These authors contributed equally to the paper.

* Corresponding author; e-mail twe@le.ac.uk; fax 44–0116–252–2791.

---
100 different \textit{myb} genes identified in Arabidopsis (Kranz et al., 1998). The MYB domain, which is composed of two or three imperfect repeats of about 50 amino acids with a predicted helix-turn-helix motif, binds DNA in a sequence-specific manner. Plant MYB proteins typically bind to one or more of two types of cis-elements known as MYB binding sites (MBSI, MBSII, or MBSIIG.), the consensus sequences of which are CNGTTR, GKTWGTTTR, and GKTWGGRTR, respectively (where \( n = A/T/C/G, R = A/G, K = G/T, \) and \( W = A/T; \) Romero et al., 1998).

Plant MYB proteins have been shown to regulate diverse developmental processes, as well as being involved in environmental signaling and secondary metabolism (Lipsick, 1996; Jin and Martin, 1999). The most well-established role of plant MYB proteins, however, is in the control of genes in the phenylpropanoid biosynthetic pathway (Quattrocchio et al., 1993; Franken et al., 1994; Martin and Paz-Ares, 1997). For example, \textit{PhMYB3} from petunia (Solano et al., 1995), \textit{AmMYB305} from snapdragon (Sablowski et al., 1994; Moyano et al., 1996), \textit{PcMYB26} from pea (Uimari and Strommer, 1997), and \textit{ZmMYB1} from maize (Franken et al., 1994) regulate levels of anthocyanins and flavonoids in flowers, whereas \textit{ZmMYBPL} from maize (Cone et al., 1994) performs a similar function in vegetative tissues. None of these \textit{myb} genes, however, are known to be expressed in anthers and/or pollen.

Phenylpropanoids are known to play an important role in pollen fertility (for review, see Taylor and Hepler, 1997). Mutants deficient in chalcone synthase (CHS) expression in anthers produce flavonol-deficient pollen that is defective in germination or pollen tube growth (Mo et al., 1992; van der Meer et al., 1992). This defect can be rescued by the application of specific flavonols, which are thought to play a structural or signaling role in pollen tube growth (Ylstra et al., 1994; Napoli et al., 1999).

Based on the hypothesis that \textit{myb} family members may play important roles in male gametogenesis, we screened a mature pollen cDNA library of tobacco for \textit{myb}-related sequences and isolated two cDNA clones (\textit{NtmybAS1} and \textit{NtmybAS2}) encoding MYB proteins. These two genes were expressed specifically in sporophytic and gametophytic tissues of tobacco anthers. We show that \textit{NtmybAS1} encodes a functional regulatory protein, which binds both MBSI and MBSIIG, with a higher affinity for MBSI, and to MBS-containing fragments of tobacco \textit{PAL} promoters. Transcriptional activation of \textit{gPAL1} by \textit{NtMYBAS1} in leaf protoplasts and co-expression of \textit{NtmybAS1} and \textit{gPAL1} transcripts in young anther tissues strongly suggests that \textit{NtMYBAS1} is a regulator of \textit{PAL} expression and phenylpropanoid biosynthesis in sporophytic tissues of the anther.

**RESULTS**

**Cloning of \textit{NtmybAS1/2} cDNAs**

To identify putative pollen-expressed \textit{myb} genes, we screened a tobacco mature pollen cDNA library with a mixture of two MYB domain consensus 42 mer oligonucleotides according to Jackson et al. (1991). Sequence comparison of the cDNA insert sequences of four positive clones revealed two unique cDNA sequences. \textit{NtmybAS1} contains a 1,409-bp open reading frame encoding a 470-amino acid protein with an estimated molecular mass of 52.2 kD, whereas \textit{NtmybAS2} contains a 1,421-bp open reading frame encoding a 474-amino acid protein with an estimated molecular mass of 52.7 kD. Comparison of \textit{NTMYBAS1} and \textit{NTMYBAS2} amino acid sequences revealed that they share 92% sequence identity overall. Within their two contiguous R2 and R3 MYB repeats, which together constitute the MYB domain, this figure rises to 98% (Fig. 1). Southern-blot analysis showed that both genes are single copy in tobacco, with the parental \textit{Nicotiana sylvestris} and \textit{Nicotiana tomentosiformis} genomes contributing one gene copy each (data not shown). Therefore, these genes are likely to be orthologues.

The \textit{NtMYBAS1/2} proteins form a tightly related group with the petunia \textit{PhMYB3} protein (Avila et al., 1993). The 92% sequence identity in the R2 and R3 MYB repeats between \textit{NtMYBAS1/2} and \textit{PhMYB3} might suggest that they bind similar sequences. In addition to the N-terminal MYB domain, their C-terminal domains also shows regions of high identity (Fig. 1, marked by asterisks). Regions of high similarity exist between \textit{NtMYBAS1/2} and \textit{PhMYB3} in three potential trans-activation domains (amino acids 326–347, 357–378, and 413–432 in \textit{NtMYBAS1} and amino acids 330–351, 355–384, and 417–436 in \textit{NtMYBAS2}) by virtue of their acidity and potential to be α-helical. Sequence similarity in this region of the proteins could indicate that \textit{NtMYBAS1/2} share a similar biological function to \textit{PhMYB3}, which has been proposed to regulate CHS expression and thereby flavonoid biosynthesis in petunia petals (Avila et al., 1993; Solano et al., 1995).

According to Kranz et al. (1998), the MYB proteins in Arabidopsis were classified into 22 subgroups. \textit{NtMYBAS1/2} proteins belong to subgroup 18, which share the QRaGLPxYPxEx/S motif outside the MYB domain (Fig. 1, bold characters). Aligning the R2 and R3 MYB repeats of \textit{NtMYBAS} proteins with other MYB proteins, the three conserved Trp residues separated by 18 to 19 amino acids implicated in DNA binding (Saikumar et al., 1990; Ogata et al., 1994) are present in \textit{NtMYBAS1/2} sequences (Fig. 2a, marked as asterisks). As in other plant MYB domains, the first Trp of the R3 repeat in the \textit{NtMYBAS1/2} proteins is replaced by a Phe.

In Figure 2b, a dendrogram represents the degree of protein sequence similarity based on the align-
ments in Figure 2a. The NtMYBAS1/2 proteins are predicted to be relatively recently diverged from each other, and to cluster tightly with the PhMYB3 sequence relative to other plant sequences. HsMYBC and DmMYB from Homo sapiens and Drosophila melanogaster, respectively, form a separate deeply rooted branch in addition to the AtMYB101 and AtMYB81 proteins in this group (Kranz et al., 1998).

NtmybAS Expression Is Developmentally Regulated during Male Gametogenesis

Given the likelihood of NtMYBAS1 and NtMYBAS2 being orthologous, we continued our investigation with NtMYBAS1 only. The organ-specific expression pattern of NtmybAS shown by northern-blot analysis showed high expression of a 1.8-kb transcript in stamens and mature and germinating pollen (Fig. 3a). In contrast, young seedlings and other vegetative or floral organs did not show expression of NtmybAS transcripts.

To determine the temporal expression pattern of NtmybAS transcripts, total RNA was extracted from anthers at five developmental stages (10–12-mm, 12–16-mm, 17–30-mm, 30–40-mm, and 40–50-mm bud lengths) and from mature pollen. NtmybAS transcripts were first detectable at early and late stages of anther development with maximum levels in mature pollen (Fig. 3b).

To investigate the gametophytic expression pattern of NtmybAS, total RNA was extracted from purified populations of microspores and pollen from five different stages of development. NtmybAS transcripts were first detectable in late uninucleate microspores (12–15-mm buds). After pollen mitosis I (17–30-mm buds), the level of NtmybAS transcripts decreased to a basal level. Transcript abundance subsequently increased in mid- (30–40-mm buds) and late pollen-specific genes in tobacco (Twell, 1994; Sweetman et al., 2000).

NtmybAS Transcripts Are Expressed in Sporophytic and Gametophytic Anther Tissues

To define the cellular expression pattern of NtmybAS1, in situ hybridization was carried out on sections of flower buds from five developmental stages (8–9 mm, 10–12 mm, 13–15 mm, 16–20 mm, and 21–30 mm in length). The 243 bp (position 1,345–1,587) of NtmybAS1 cDNA was used as a probe. The antisense NtmybAS1 probe produced strong signals in sporophytic tissues, including vascular tissue, stomium, and tapetum, at early stages of development (Fig. 4a). As the anthers matured there was a decrease in sporophytic and an increase in gametophytic (pollen) hybridization (Fig. 4, b–e). The reduced sporophytic hybridization coincided with the degeneration of the tapetum and stomium. No significant hybridization signals were observed using the NtmybAS sense probe except for mature pollen, which showed lower background staining (Fig. 4, f–j).

Figure 1. Comparison of the amino acid sequence of NtMYBAS1/2 and PhMYB3. The R2 and R3 repeats of the MYB domains are boxed. Arrowheads correspond to conserved Trp residues in the MYB domain. The MYB protein subgroup 18 consensus motif (QRaGLPxYPxE/S; Kranz et al., 1998) is highlighted in bold. Identical amino acids are indicated below with asterisks.
in the northern analyses of anthers and isolated spores (Fig. 3). These data clearly demonstrate early sporophytic and late gametophytic expression of NtmybAS1 in developing anthers.

The N-Terminal MYB Domain Efficiently Targets Green Fluorescent Protein (GFP) to the Nucleus

The N-terminal MYB domain of NtMYBAS1 contains a basic motif (119-RLKKRRQR-125) resembling a monopartite nuclear localization signal. To determine whether the MYB domain containing this motif plays a role in nuclear import of NtMYBAS1 in vivo, pollen was bombarded with a control plasmid, pLAT52-sGFP, and a plasmid, pLAT52-sGFPM1, harboring the N-terminal 126 amino acids of NtMYBAS1 fused to sGFP. When sGFP was expressed alone (pLAT52-sGFP), the majority of the fluorescence signal was localized in the pollen cytoplasm (Fig. 5a). In contrast, the recombinant protein (sGFPM1) was exclusively localized to the pollen nucleus, with some localized accumulation in the nucleolus (Fig. 5b). We conclude that the first 126 amino acids of NtMYBAS1 are sufficient to target and retain sGFP in the nucleus.

To investigate the role of specific sequences in nuclear localization within this N-terminal domain, the localization of two additional mybAS1-sGFP fusion proteins were examined. Specific removal of the RLKRRQR motif from the N-terminal Myb domain (pLAT52-sGFPM2) severely reduced nuclear localization of sGFP, demonstrating its involvement in nuclear targeting (Fig. 5c). However, enhanced accumulation of sGFPM2 in the nucleus relative to sGFP alone suggests that the N-terminal domain (amino acids 1-119) retains partial nuclear localization function. In a similar manner, the RLKRRQR motif domain alone (pLAT52-sGFPM3) was not sufficient to exclusively direct sGFP to the nucleus, although enhanced accumulation of sGFPM3 was observed in the nucleus and nucleolus (Fig. 5d). These results demonstrate that the RLKRRQR motif is necessary but not sufficient for nuclear localization, and suggest that additional upstream sequences within the

Figure 2. Alignment (a) and dendrogram (b) of the deduced amino acid sequence of the MYB domain of NtMYBAS1/2 with those of related MYB proteins. Gene, genus nomenclatures, and accession nos. are as follows: NtMYBAS1, tobacco, AF198499; NtMYBAS2, tobacco, AF198498; PhMYB3, petunia, Z13998; AtMYB101, Arabidopsis, X90379; OsMYBGA, Oryza sativa, X98355; HvMYBGA, Hordeum vulgare, X87690; AtMYB33, Arabidopsis, AF062875; AtMYB65, Arabidopsis, AF062899; AtMYB81, Arabidopsis, AF062911; AtMYB103, Arabidopsis, AF048839; HsMYBC, Homo sapiens, M15024; and DmMYB, Drosophila melanogaster, X05939. Asterisks correspond to positions of conserved Trp residues. Boxes beneath the sequences indicate the three helices in each repeat including the recognition helix (black box). Sequence alignment and dendrogram were performed by the neighbor-joining algorithm of CLUSTALW, where the scale bar represents the degree of base pair substitution per million years (Thomson et al., 1994).
N-terminal Myb domain are also required for efficient nuclear targeting.

NtMYBAS1 Binds to MBS Motifs

To determine NTMYBAS1 DNA-binding specificity, random binding site selection experiments were performed with a C terminally truncated derivative, NtMYBAS1C1. This truncated protein (amino acids 1–208) containing the intact MYB domain, was expressed as an N-terminal poly-His fusion protein in Escherichia coli. NtMYBAS1C1 was incubated with oligonucleotides containing a random 23-bp core and was separated from unbound oligonucleotides by electrophoretic mobility shift assay (EMSA). After six rounds of screening, bound oligonucleotides were cloned and sequenced.

All of the oligonucleotides selected by NtMYBAS1C1 contained at least one AAC core motif characteristic of MYB binding sites (Fig. 6). The deduced consensus, (T/C) AACNGC, is similar to MBSI, (T/C) AAC(G/C) G, which is bound by animal MYB and some plant MYB proteins (Romero et al., 1998). NtMYBAS1C1 bound to these selected oligonucleotides with differing affinities in separate EMSAs (Fig. 6), with Y10 and Y18 showing the highest and lowest percentages of bound oligonucleotide, respectively.

Because all oligonucleotides contain the same AAC core motif, this suggests that sequence context is an additional factor affecting binding affinity.

To confirm the binding site specificity of NtMYBAS1C1, we tested the ability of NtMYBAS1 to bind to oligonucleotides containing either MBSI, or MBSIIG, and to mutated forms of these sequences. NtMYBAS1C1 bound strongly to MBSI oligonucleotides containing the sequence TAACCG (Fig. 7b). Extensive mutation of this sequence to TCCCCG in MBSIm2 abolished binding (Fig. 7b, lane 4), whereas changes outside the core sequence to TAACTG as in MBSIm1 resulted in a low level of binding (Fig. 7b, lane 3). This demonstrates that the core MYB binding site motif, AAC, is critical for NtMYBAS binding.

Figure 3. mRNA expression analysis of NtmybAS. Northern blots of tobacco total RNA probed with radiolabeled NtmybAS1 probe. a, Spatial expression. 13D and 17D, Seedlings 13 and 17 d after germination; Rt, root; Yl, young leaf; Sp, sepal; Pt, petal; Pst, pistil; Stm, stamen; MP, mature pollen; GP, germinating pollen. b, Temporal expression in anthers. WL, Wild-type leaf; MP, mature pollen. Flower bud lengths shown in millimeters. c, Temporal expression in isolated spores. MP, Mature pollen; GP, germinating pollen.

Figure 4. Localization of NtmybAS mRNA during tobacco anther development. Paraffin-embedded sections were hybridized with digoxigenin-labeled NtmybAS1 probes and viewed under bright field. Scale bar = 400 μm. Magnification ×100 in 8- to 9-mm and 10- to 12-mm buds; and ×40 in 13- to 15-mm, 16- to 20-mm, and 21- to 30-mm buds. a through e, NtmybAS1 antisense probe with anthers from: a, 8 to 9 mm; b, 10 to 12 mm; c, 13 to 15 mm; d, 16 to 20 mm; and e, 21- to 30-mm buds. Staining represents hybridization signal. f through j, NtmybAS1 sense probe with anthers from: f, 8 to 9 mm; g, 10 to 12 mm; h, 13 to 15 mm; i, 16 to 20 mm; and j, 21- to 30-mm buds.
NtMYBAS1 also bound to the sequence GTTGGT, an MBSIG motif (Fig. 7b, lane 5). Extensive mutation of this sequence as in MBSIm2 (TTGCCG) abolished binding (Fig. 7b, lane 7), whereas a 2-bp mutation to TTGGAT in MBSIm1 had no effect (Fig. 7b, lane 6). The intensity of the shifted band in Figure 7b implies that NtMYBAS1 can bind both types of MBS sequences, but with a preference for MBSI. To further demonstrate preferential binding of NtMYBAS1 to MBSI, competition experiments were performed. Binding of NtMYBAS1 to MBSI was effectively reduced by the addition of unlabeled MBSI (Fig. 7c). In contrast, competition with MBSII, and with mutated derivatives of MBSI and MBSII, had little or no effect on binding.

NtMYBAS1 Binds Tobacco PAL Promoters in Vitro

Because several MYB proteins are known to play important roles in transcriptional regulation of phenylpropanoid biosynthetic genes (Martin and Paz-Ares, 1997) including trans-activation of the Phe ammonia lyase (PAL) promoter (Sablowski et al., 1994), we postulated that NtMYBAS1 similarly may regulate PAL expression specifically in anthers. To test whether NtMYBAS1 was capable of binding to a PAL gene promoter, EMSA was performed with two tobacco PAL genes, PALA (Pellegrini et al., 1994) and gPAL1 (Tomoko et al., 1996), which contain MBSs in their promoters. Four fragments (P1, P2, P3, and P4) of the PALA promoter, each containing multiple MBSs, were bound by NtMYBAS1 (Fig. 8a). Two
of the gPAL1 promoter fragments (T2 and T3) were also bound, whereas T1, which has no MBSs, failed to produce a shifted band (Fig. 8b). We conclude that NtMYBAS1 specifically binds tobacco PAL promoter regions containing MBSs.

**NtMYBAS1 Activates Transcription of PAL Promoters in Vivo**

To investigate whether the NtMYBAS1 protein could transcriptionally activate PAL, we cotransfected tobacco leaf protoplasts with an effector plasmid (35S::NtMYBAS1) harboring NtmybAS1 under the control of CaMV35S promoter and four different reporter plasmids. Reporter plasmids contained the tobacco PALA (Pellegrini et al., 1994) or gPAL1 (Tomoko et al., 1996) promoters, and two tobacco pollen-specific promoters, NPG1 (Tebbutt et al., 1994) and neIF4A8 (Brander and Kuhlemeier, 1995) fused to the GUS reporter gene. NtMYBAS1 specifically trans-activated both the PALA and gPAL1 promoters, whereas NPG1 and neIF4A8 promoters were not activated above basal levels of the reporter plasmids alone (Fig. 9b).

To identify the putative trans-activation domain within NtMYBAS1, we constructed five effector plasmids (M1–M5), harboring either the full-length or C-terminal truncations of NtMYBAS1 (Fig. 10a). These were tested for their ability to trans-activate a GUS reporter gene under the control of the two tobacco PAL promoters. The truncated protein M3 (amino acids 1–257) produced a significant decrease in the trans-activation of both PAL promoters, whereas truncation M4 (amino acids 1–322) did not affect trans-activation (Fig. 10b). Thus, sequences between 257 and 322 define an important transcriptional activation domain.

**gPAL1 Expression Coincides with NtmybAS1 Expression in Anthers**

From the DNA binding assays and transient expression analyses carried out it is possible to propose NtMYBAS1 as a regulator of PAL expression in anthers. To investigate this hypothesis, we studied the expression pattern of PAL in anthers and pollen. A northern blot previously probed with NtmybAS1 (Fig. 3b) was reprobed with a gPAL1 probe. gPAL1 was strongly expressed during the same early stages of anther development at which NtmybAS1 transcripts were highly expressed (Fig. 11a). The transcript abundance in 10- to 16-mm buds was higher than in leaves and at other stages of anther development (Fig. 11a). In contrast to NtmybAS1, there was no detectable gPAL1 expression in pollen.

Figure 7. EMSA with purified recombinant NtMYBAS1 protein. a, The sequences of double-stranded oligonucleotides used as probes. b, The NtMYBAS1 protein binds preferentially to MBSI (GTAACCG). The binding reaction fractionated in the first lane did not contain NtMYBAS1, whereas those in other lanes contained 200 ng NtMYBAS1. c, Competition for binding of NtMYBAS1 to the MBSI probe. The binding assay was performed by pre-incubating with unlabeled competitor before addition of 1 ng of 32P-labeled MBSI probe. The reaction fractionated in lane 1 contained no competitor, whereas those in lanes 2 through 7 contained 200 ng of competitor.
**DISCUSSION**

We have isolated two cDNAs clones (NtmybAS1 and NtmybAS2) from tobacco encoding MYB-related proteins, expressed specifically in sporophytic and gametophytic anther tissues. The 98% amino acid sequence identity between NtMYBAS1 and NtMYBAS2 within their N-terminal DNA binding domains suggests that their biological roles may be similar. Furthermore, based on the finding that in tobacco both NtmybAS1/2 are single copy, and inherited from the genome of each parental line, we speculate that these genes are likely to be orthologous. When compared with all other available MYB protein sequences, NtMYBAS1/2 show the highest similarity to petunia PhMYB3. According to the classification of MYB proteins from Arabidopsis (Kranz et al., 1998), this places NtMYBAS1/2 in subgroup 18, although the closest MYB family members in Arabidopsis are relatively divergent. PhMYB3 is expressed in petal epidermal cells and is able to trans-activate the promoter from a CHS gene, suggesting an involvement in the regulation of flavonoid biosynthesis (Solano et al., 1995). The function of the other MYB proteins in this group remains unknown.

Northern-blot analysis revealed that NtmybAS transcripts are initially abundant in young anthers when the tapetum is intact and subsequently decrease in line with tapetal cell degeneration (Koltunow et al., 1990). However, NtmybAS transcripts also accumulate gametophytically in mature and germinating pollen similar to the “late class” of pollen-expressed genes (Mascarenhas, 1990; Twell, 1994). In situ hybridization revealed strong sporophytic expression in tapetum, stomium, and vascular tissues of young anthers, and confirmed abundant expression in maturing pollen. These data suggest that NtMYBAS1/2 are likely to regulate gene expression

**Figure 8.** EMSA with plant promoter fragments. Binding of the NMYBAS C1 to PALA (a) and gPAL1 promoter fragments (b). The locations of the fragments in the promoters relative to the transcriptional start site are given as follows: PALA P1, −1,115 to −876; P2, −875 to −626; P3, −625 to −396; P4, −396 to −356; gPAL1 T1, −658 to −475; T2, −474 to −204; and T3, −203 to −42. ψ represents the presence of MBSs.

**Figure 9.** Trans-activation of tobacco promoters by NtMYBAS1. a, Effector and reporter constructs used in cotransfection experiments. The 35S::NtMYBAS1 effector plasmid consisted of the coding region of NtmybAS1 under control of the CaMV 35S promoter. The reporter plasmids consist of two tobacco PAL promoters, PALA (Pellegrini et al., 1994) and gPAL1 (Tomoko et al., 1996); tobacco polygalacturonase NPG1 promoter (Tebbutt et al., 1994) and tobacco nefI4A8 promoter (Brander and Kuhlemeier, 1995) fused to the GUS reporter gene. b, Response of reporter constructs to NtMYBAS1 in transient expression assays. Tobacco leaf protoplasts were cotransfected with each reporter plasmid together with 35S::NtMYBAS2. Error bars represent SD (n > 7).
in both sporophytic anther tissues and in maturing pollen.

Several anther- and pollen-specific genes encoding putative transcription factors have been isolated—(Baltz et al., 1992; Zachgo et al., 1997; Chung et al., 1998; Kobayashi et al., 1998; Schieflhaler et al., 1999; Yang et al., 1999; Heuer et al., 2000), including a tapetum-specific myb-related gene from Arabidopsis (Li et al., 1999). NtmybAS1/2 appear to be unique in that they show both sporophytic and gametophytic expression. However, there are several examples of anther-specific genes, which show this overlap in expression. For example, Bcp1 is expressed in tapetum and microspores and is functionally important in both cell types (Xu et al., 1995). Several lipid biosynthetic genes are also expressed both in tapetal cells and pollen grains (Piffanelli et al., 1997). Therefore, such sporophytic/gametophytic overlap of gene expression in the anther may be regulated by transcription factors that are expressed in the same manner as NtMYBAS1/2.

The N-Terminal Myb Domain of NtmybAS1 Contains a Functional Nuclear Localization Signal (NLS)

Localization of plant Myb proteins in the nucleus has previously been shown by immunocytochemical analysis for petunia PhMyb3 (Avila et al., 1993) and by transient expression for Arabidopsis CCA1 (Wang et al., 1997). DNA binding assays showed that NtMYBAS1 binds preferentially to a type I MYB binding site, but that it could also bind weakly to a type II G site (Romero et al., 1998). Random binding site selection has been successful in deriving consensus target sequences for MYB proteins, which are very similar to their presumptive native targets (Solano et al., 1995; Gubler et al., 1999). Therefore, the results described here may reflect the in vivo binding properties of NtMYBAS1.

The N-terminal 126 amino acids of NTMYBAS1 localize sGFP exclusively to the pollen nucleus. A putative NLS (119-RLKRRQR-125) in NtMybAS1 resembling monopartite NLSs (Raikhel, 1992; Merkle and Nagy, 1997) is present within this region. Deletion of the RLKRRQR motif from this region (sGFP2) severely reduced, but did not abolish, the nuclear localization of sGFP, demonstrating that this motif is necessary for efficient nuclear import. Analysis of surface probability (MacVector, Oxford Molecular Group PLC, Oxford) revealed that the Arg at position 122 (within the 119-RLKRRQR-125 motif) is predicted to have the highest surface probability along the length of NTMYBAS1. This implies that the six amino acids surrounding this Arg have the highest probability of being exposed and recognized by the nuclear import machinery.

The partial nuclear accumulation of sGFP2 suggests the presence of additional sequences that contribute to nuclear localization. Further analysis demonstrated that the RLKRRQR motif alone also was not sufficient for exclusive nuclear localization of sGFP, but showed enhanced accumulation in the nucleolus. These data suggest that the NTMYBAS1 NLS is not simply monopartite but depends on the cooperative function of the RLKRRQR motif with an upstream element in the DNA binding domain. The cooperative functions of bipartite and multiple NLSs have been reported for the potyviral NiA protein (Carrington et al., 1991) and B-Myb (Takekomo et al., 1994). Two separate NLSs, a monopartite (NLS1) in the central region and a bipartite (NLS2) in the C-terminal basic region were shown to be necessary for efficient nuclear targeting of B-Myb (Takekomo et al., 1994).

It is interesting that the basic RLKRRQR motif resembles the nucleolar targeting signals of the RNA-binding rat ribosomal protein L31 (LRKRKR, Quaye et al., 1996). Furthermore, Arg-rich motifs are known to bind RNA and their binding affinity is dramatically enhanced by the presence of a single Gln (Q) or Asn (N; Tan and Frankel, 1998). This raises the possibility that the partial localization of sGFP (containing the RLKRRQR motif alone) to the nucleolus may be mediated by a similar mechanism via RNA binding.

NtMYBAS1 as a Regulator of gPAL1 Expression

DNA binding assays showed that NtMYBAS1 binds preferentially to a type I MYB binding site, but that it could also bind weakly to a type II G site (Romero et al., 1998). Random binding site selection has been successful in deriving consensus target sequences for MYB proteins, which are very similar to their presumptive native targets (Solano et al., 1995; Gubler et al., 1999). Therefore, the results described here may reflect the in vivo binding properties of NtMYBAS1.

PAL gene promoters typically contain multiple potential MBS elements, which are important for
expression of these genes (Hatton et al., 1996). NtMYBAS1 was shown to bind specifically to fragments from two tobacco PAL gene promoters that contain MBSI and MBSII sites. In addition, NtMYBAS1 was able to trans-activate the same PAL gene promoters in tobacco protoplasts, whereas it failed to activate two tobacco pollen-specific promoters. Transient expression of truncated NtMYBAS1 derivatives revealed an internal region (amino acids 257–322) necessary for trans-activation. In contrast, other MYB proteins have been shown to contain acidic C-terminal domains that are required for trans-activation (Facchinetti et al., 1997; Gubler et al., 1999). The internal trans-activation region in NtMYBAS1 is not acidic and is not predicted to form a-helix. It is interesting that this region contains putative protein kinase C (TTR, amino acids 264–266) and caseine kinase C phosphorylation sites (SIPE, amino acids 293–303; TSSD, amino acids 306–309), which could be important posttranslational regulatory sites. These results collectively suggest NtMYBAS1 is a good candidate for a positive regulator of PAL transcription in tobacco anthers.

NtMYBAS1/2 are most closely related to PhMYB3 from petunia, which binds both MBSI and MBSII sites and weakly to MBSII (Solano et al., 1995, 1997). The evidence thus far suggests PhMYB3 to be involved in phenylpropanoid metabolism, possibly in anthocyanin biosynthesis due to its ability to trans-activate a CHS gene promoter in tobacco protoplasts. We have not tested the ability of NtMYBAS1/2 to trans-activate the CHS promoter, nor has PhMYB3 been tested against a PAL promoter, but this may be evidence for structurally similar proteins from different species not necessarily activating the same target genes. This situation is true for the highly conserved MYBs ZmMYBC1 and PhMYBAN2, which have been shown genetically to control different genes in maize and petunia, respectively (Quattrochio et al., 1999). This also may not be too surprising given that the introduction of a single amino acid change in the R3 repeat of PhMYB3 has the effect of changing the binding site preference from MBSI and MBSII, to MBSI alone in the mutated form (Solano et al., 1997). On the other hand, structurally dissimilar MYB proteins can bind to similar DNA sequences (Solano et al., 1997; Romero et al., 1998) and therefore could have potentially similar target genes and functions.

NtMYBAS1 is not the only MYB protein to be shown to trans-activate a PAL gene in vivo. AmMYB305 from snapdragon, an evolutionarily distant MYB protein to NtMYBAS1/2, will trans-activate a PAL gene promoter in tobacco protoplasts...
ever, the strong induction of expression and the decay of the tapetal cells. How-coincided precisely with the suppression of expression in anthers at mid-bicellular stage anthers.

$gPAL1$ NtMYBAS1 as an activator of provides further compelling evidence for the role of tapetum at early stages of anther development pro-
ducer of PAL expression and therefore phenyl-
panoid pathway, is an important regulatory point (Hahlbrock et al., 1976; Bate et al., 1994). PAL activity in immature anthers has also been maximal in this regard, PAL enzyme activity was maximal in immature anthers and immunolocalization showed that PAL was predominantly localized to the tapetal cells. PAL activity in immature anthers has also been positively correlated with pollen fertility in male-fertile and cytoplasmic male-sterile strains of Brassica oleracea (Kishitani et al., 1993). Furthermore, tapetum-specific sense and antisense expression of a sweet potato PAL cDNA in tobacco resulted in partial male sterility and abnormal pollen grains devoid of flav-
onoids (Matsuda et al., 1996). In sharp contrast, flavonoids do not play an essential role in pollen fertility in Arabidopsis (Burbulis et al., 1996), so the flavonoid requirement for male gametophyte develop-
ment can differ between species.

$gPAL1$ Expression in Tapetum as a Regulator of Phenylpropanoid Synthesis

Several lines of evidence demonstrate that specific flavonols derived from the phenylpropanoid pathway play an essential role in pollen fertility (for review, see Taylor and Hepler, 1997). PAL, which catalyzes the first committed step in the phenylpropanoid pathway, is an important regulatory point (Hahlbrock et al., 1976; Bate et al., 1994). PAL activity in the tapetum therefore is expected to be an important regulator of flavonol synthesis in the anther. In this regard, PAL enzyme activity was maximal in immature anthers and immunolocalization showed that PAL was predominantly localized to the tapetal cells. PAL activity in immature anthers has also been positively correlated with pollen fertility in male-fertile and cytoplasmic male-sterile strains of Brassica oleracea (Kishitani et al., 1993). Furthermore, tapetum-specific sense and antisense expression of a sweet potato PAL cDNA in tobacco resulted in partial male sterility and abnormal pollen grains devoid of flavonoids (Matsuda et al., 1996). In sharp contrast, flavonoids do not play an essential role in pollen fertility in Arabidopsis (Burbulis et al., 1996), so the flavonoid requirement for male gametophyte development can differ between species.

Given the importance of PAL in providing phenylpropanoid precursors for flavonol synthesis in the tapetum, identification of transcriptional regulators of PAL is equally important. The demonstration of binding of NtMYBAS1 to PAL promoters in vitro and trans-activation of PAL-GUS in protoplasts strongly supports the hypothesis that NtMYBAS1 may be a regulator of PAL expression and therefore phenylpropanoid biosynthesis in tobacco anthers.

The colocalization of NtmybAS and gPAL1 in the tapetum at early stages of anther development provides further compelling evidence for the role of NtMYBAS1 as an activator of gPAL1 expression in the tapetum. The onset of repression of gPAL1 expression in anthers at mid-bicellular stage anthers coincided precisely with the suppression of NtmybAS expression and the decay of the tapetal cells. However, the strong induction of NtmybAS mRNA in pollen at later stages of anther development further suggests that NtMYBAS1/2 might perform other regulatory functions in mature pollen.

Based on the proposal that NtMYBAS1 acts as a positive regulator of $gPAL1$, we can postulate potential mechanisms that could account for the specific activation of $gPAL1$ in sporophytic anther tissues. First, transcription of $gPAL1$ in sporophytic tissues may be mediated by binding of NtMYBAS1 to MBSs in the gPAL1 promoter. In gametophytic cells, including mature pollen, suppression of $gPAL1$ may result from the presence of a repressor binding either to the $gPAL1$ promoter or to NtMYBAS1. Because some MYB proteins are known to cooperate with other transcription factors (Tice-Baldwin et al., 1989; Burk et al., 1993), NtMYBAS1-mediated induction of $gPAL1$ in sporophytic tissues could also be indirect via a second factor, and the absence of this factor in pollen could prevent $gPAL1$ transcription. Posttranslational modifications such as phosphorylation and acetylation, which are known to regulate MYB protein activity, could also modulate NtMYBAS1 binding to the gPAL1 promoter in a tissue-specific manner (Johnson et al., 1999; Tomita et al., 2000). Among plant MYB proteins CCA1 has been shown to interact with, and is phosphorylated by, a regulatory subunit (CKB3) of the Ser/Thr protein kinase CK2 (Sugano et al., 1998).

Confirmation of the primary role of NtMYBAS1 as a positive regulator of PAL expression in sporophytic tissue of the anther and a demonstration of its other regulatory functions in the anther and pollen requires further work. One approach to this will be to examine the effect of tissue-specific reduction of NtMYBAS expression in tapetum and pollen and to map the structurally important cis-elements in the tobacco gPAL1 promoter.

MATERIALS AND METHODS

Plant Materials

Tobacco (Nicotiana tabacum cv Samsun) plants were grown under standard greenhouse conditions (16-h day, 18°C–28°C).

cDNA Library Construction and Screening

RNA isolation and construction of a UniZAPII cDNA library from mature pollen poly(A)− RNA of tobacco were carried out as previously described (Sweetman et al., 2000). Approximately 4 × 10⁵ pfu were screened using a mixture of two 42 mer oligonucleotides, O1 and O2 (Jackson et al., 1991). Oligonucleotide labeling and hybridization procedures performed as previously described (Jackson et al., 1991).

DNA Sequencing and Analysis

Sequencing was performed using the automated ABI Prism procedure (Applied Biosystems, San Jose, CA). Se-
sequence data was analyzed using the Sequence Editor v1.0.3, GeneJockey II (Biosoft, Cambridge, UK) and protein secondary structure predictions were compiled using the Chou Fasman and Robson Garnier algorithm (MacVector, Oxford Molecular Group PLC).

The nucleotide sequences of cDNA clones NtmybAS1 and NtmybAS2 are available in GenBank as accession nos. AF198499 and AF198498, respectively. Clones are freely available from the corresponding author (twe@le.ac.uk) for non-commercial research purposes subject to a material transfer agreement with the University of Leicester.

Northern and Southern Analysis

Genomic DNA and total RNA isolations, gel-blot studies, and hybridization were carried out as previously described (Sweetman et al., 2000).

RNA in Situ Hybridization

For in situ hybridization, probes were labeled with digoxigenin-11-UTP using a nucleic acid labeling kit (Boehringer, Indianapolis). The 243-bp (position 1,345–1,587) NtmybAS1 cDNA was amplified using two oligonucleotide primers, 5′-GGATGCCACCCACAAAGCCACTGCAA-3′ and 5′-GATTTGGACAAATGTGATTGTATTTGGT-3′.

229-bp (position 85–313) gPAL1 cDNA was amplified using two oligonucleotide primers, 5′-GCTGAAATCTTAAAGGAGGTAC-3′ and 5′-CAGTTCCTTTATTCA-3′. The PCR-amplified cDNAs were cloned into pGEM-T (Promega, Madison, WI) and used as templates for synthesizing digoxigenin-labeled RNA probes. Sense and antisense NtmybAS1 probes were made using T7 and SP6 RNA polymerases after digestion with SaeI and ApaI, respectively. Sense and antisense gPALI probes were made using SP6 and T7 RNA polymerases after digestion with ApaI and SaeI, respectively.

Flower buds at five developmental stages (8–9 mm, 10–12 mm, 13–15 mm, 16–20 mm, and 21–30 mm in length) were fixed and paraffin embedded according to Dixon et al. (1995). Sectioning, hybridization, and detection of hybridization signals were performed (Sweetman et al., 2000).

Construction of sGFP and NLS Fusion Plasmids

The control plasmid pLAT52-sGFP, used to direct expression of GFP in pollen, consisted of the pollen-specific lat52 promoter (Twell et al., 1989a), the tobacco etch virus translational enhancer sequence (TEV-L), the sGFP coding sequence (Chiu et al., 1996), and the cauliflower mosaic virus 35S polyadenylation/terminator sequence (C3′) in pUC19 (precise construction details available on request). An NtMYBAS1-GFP fusion plasmid (pLAT52-sGFPM1) was constructed as follows. A 392-bp fragment of NtmybAS1 cDNA (position 176–567) encoding the N-terminal 126 amino acids of NtMYBAS1 was amplified using oligonucleotide primers (5′-GGAATCATGACACCAGATGGAGGAGCAG-3′ and 5′-GGCATGATCCTGTCTGGACACAC-3′) with BspHI restriction sites at both ends. The amplified fragment was digested with BspHI, gel purified, and cloned into the NcoI site of pLAT52-sGFP fusing the first 126 amino acids of the NtMYBAS1 protein (position 181–558) to sGFP.

The fusion plasmid pLAT52-sGFPM2 was constructed to determine the effect of removing the putative NtMYBAS1 NLS (RLKRRQR motif) on nuclear targeting of sGFP. A 371-bp fragment of the NtmybAS1 cDNA (position 176–546) encoding the N-terminal 119 amino acids of NtMYBAS1 lacking the RLKRRQR motif was amplified by PCR. The same forward primer used in the construction of pLAT52-sGFPM1 and a reverse primer (5′-CATTTCATGATTGTGATCAGCTATTCTC3′) with a BspHI site at 3′ end were used. The PCR-amplified fragment was digested with BspHI, gel purified, and cloned into the compatible NcoI site of pLAT52-sGFP.

The fusion plasmid pLAT52-sGFPM3 was constructed to investigate the ability of the RLKRRQR motif alone to direct sGFP to the pollen nucleus. A 30-bp fragment of the NtmybAS1 cDNA encoding the MTRLKRRQR1 motif (position 259–558) was synthesized using two oligonucleotide primers, 5′-CATGAGAAGCTTACAAAAAAGAACACAGACAGCGC3′ and 5′-CATGGTCTTGGTCTTGGCTATGCTTGGT-3′. BspHI and NcoI sites at 5′ and 3′ ends of these oligonucleotides were introduced to avoid replacement of Thr (T) and Ala (A) residues, at positions 2 and 10, with other amino acids. Oligonucleotides were kinased, annealed, and the double-stranded fragment cloned into the compatible NcoI site of pLAT52-sGFP.

Particle Bombardment and Microscopy

Plasmids were bombarded into pollen as described by Twell et al. (1989a) with some modifications. The pollen suspension (30 mg mL−1) prepared in pollen germination medium (Tupy et al., 1991) and 0.4 mL of the suspension was spread onto the surface of a 2-cm Hybond membrane square. Each bombardment was performed in duplicate and the experiment repeated five times. Localization of sGFP was visualized after 16 h using a Nikon Optiphot microscope with a 470- to 490-nm excitation filter and a 480-nm barrier filter. Images were captured directly using a CCD camera (JVC KYF55B) and Imagegrabber software (Neotec, Southampton, UK).

Production and Purification of Recombinant Protein

To introduce restriction enzymes sites at 5′ and 3′ ends of the coding region of NtmybAS1, two 27 mer oligonucleotides were synthesized. Forward primer containing a BamHI site (TGCCTGCAGCTACGCTAGCAGGAC) and reverse primer containing a SalI site (GGCCGTAGTGATTCTGCTGAGAG) were used for PCR amplification of the NtmybAS1 cDNA. The PCR product was ligated into the BamHI/Sall site of the pET 24b(+) vector (Novagen, Madison, WI). As material for EMSA, part of the sequence (amino acids 209–470) encoding the potential activator domain was removed. To remove this region, pET:
NtmybAS1 was cut with HindIII. The HindIII site was religated such that the intact MYB domain and part of the activation domain remained. This recombinant protein, NtMYBASΔC1, was purified from IPTG-induced Escherichia coli cells over an Ni⁺-agarose column under native conditions according to the manufacturer’s instructions (Qiagen, Valencia, CA). Purified protein was eluted from the column using elution buffer (1× Tris-buffered saline with 500 mM imidazole) and dialysis was done in a buffer containing 50% (w/v) glycerol, 1× Tris-buffered saline, and 1 mM phenylmethylsulfonyl fluoride at 4°C overnight.

PCR-Assisted Binding Site Selection

Binding site selection was performed essentially according to Blackwell et al. (1990). The oligonucleotide 5’-TCATGATCCATAGGCAGNN21AGTATGATGCCTTAGACGGT-3’ was synthesized and used for PCR. Approximately 2 pmol of radiolabeled double-stranded DNA was incubated with 100 ng of purified recombinant protein in binding buffer [10 mM Tris-Cl (pH 7.6), 50 mM NaCl, 1 mM dithiothreitol, 1 mM EDTA, 5% (w/v) glycerol, 0.05 μg mL⁻¹ phenylmethylsulfonyl fluoride, 12.5 μg mL⁻¹ poly (dI-dC) overnight.

Results in each experiment. Relative binding affinities presented in Figure 6 represent mean values. For the optimization of competition assays, a range of excesses (50× to 500×) of unlabeled specific competitor were used. A 200× excess was selected because this concentration was sufficient for complete and effective competition for MBSI. Three independent experiments were performed with qualitatively similar results. A representative example is shown in Figure 7. Binding assays with PAL promoters were also repeated three times and gave qualitatively consistent results. A representative example is shown in Figure 8.

Construction of Effector Plasmids

Effector plasmids used in transfection were generated using PCR-based construction. C-terminal deleted cDNA fragments were amplified from full-length NtmybAS1 cDNA using the following synthetic oligonucleotides containing a BamHI or SacI site at their 5’ ends: 5’-TGCGGATCCAA-TGGCAACCGATGGAG-3’ (amino acid 1), 5’-GGAG-CTCTTCTTTTTAGCCCTTG-3’ (amino acid 123), 5’-GGCGCTGTTGTGAGATAGATGAGAA-3’ (amino acid 200), 5’-GGCGAGCTCTATTTGAGGCGA-3’ (amino acid 322), and 5’-GGCGAGCTCCTGCGTAACATC-3’ (amino acid 470). PCR fragments were digested with BamHI and SacI and purified fragments cloned into pBI221 (CLONTECH, Palo Alto, CA) by replacing the BamHI/SacI fragment of the GUS coding region.

Construction of Reporter Plasmids

To construct reporter plasmids, plant promoters were amplified from tobacco genomic DNA using the following synthetic oligonucleotides containing HindIII or Xhol sites at their 5’ and 3’ ends, respectively: gPAL1 promoter region, 5’-GGCGAGCTCTGATCCCGGAGAAGCTCA-3’ (for position −1,115) and 5’-GGCGTCTAGTGAATAAGGTTGTGAGGA-3’ (for position −56); PAL promoter region, 5’-GGCGAAGCTTGCACCTGCAGGTCAAC-3’ (for position −659) and 5’-GGCGTCTAGTGAATAAGGTTGTTT-3’ (for position −42); ntlF4A8 promoter region, 5’-GGCGAGCTTAACTAAAAATCTGG-3’ (for position −2,138) and 5’-GGCGTCTAGTGAATAAGGTTGTTT-3’ (for position −825) and 5’-GGCGTCTAGATGCCCCCTCCAC-CTCC-3’ (for position −6). PCR products were digested with HindIII and Xhol and purified fragments cloned into pBI221 (CLONTECH) replacing the CaMV 35S promoter region.

Protoplast Isolation and Transfection

Isolation and transfection of tobacco leaf protoplasts was performed according to Goodall et al. (1990). Expression of reporter genes was monitored using the fluorimetric assay for GUS activity (Jefferson, 1987). Protein concentrations
ACKNOWLEDGMENTS

We thank Cathie Martin for providing samples of oligonucleotides O1 and O2 and Graham Benskin and June Saddington for maintaining plants.

Received February 20, 2001; returned for revision April 10, 2001; accepted May 16, 2001.

LITERATURE CITED


Tice-Baldwin K, Fink GR, Arndt KT (1989) BAS1 has a MYB motif and activates HIS4 transcription only in combination with BAS2. Science 246: 931–935


