Evolution of Class B Floral Homeotic Proteins: Obligate Heterodimerization Originated from Homodimerization

Kai-Uwe Winter, Christof Weiser, Kerstin Kaufmann, Arend Bohné, Charlotte Kirchner, Akira Kanno, Heinz Saedler, and Günter Theißen
Max-Planck-Institut für Züchtungsforschung, Abteilung Molekulare Pflanzenzüchtung, Köln, Germany

The class B floral homeotic genes from the higher eudicot model systems Arabidopsis and Antirrhinum are involved in specifying the identity of petals and stamens during flower development. These genes exist in two different types termed DEF- and GLO-like genes. The proteins encoded by the class B genes are stable and functional in the cell only as heterodimeric complexes of a DEF- and a GLO-like protein. In line with this, heterodimerization is obligatory for DNA binding in vitro. The genes whose products have to heterodimerize to be stable and functional are each other’s closest relatives within their genomes. This suggests that the respective genes originated by gene duplication, and that heterodimerization is of relative recent origin and evolved from homodimerization. To test this hypothesis we have investigated the dimerization behavior of putative B proteins from phylogenetic informative taxa, employing electrophoretic mobility shift assays and the yeast two-hybrid system. We find that an ancestral B protein from the gymnosperm Gnetum gnemon binds DNA in a sequence-specific manner as a homodimer. Of the two types of B proteins from the monocot Lilium regale, the GLO-like protein is still able to homodimerize, whereas the DEF-like protein binds to DNA only as a heterodimeric complex with the GLO-like protein. These data suggest that heterodimerization evolved in two steps after a gene duplication that gave rise to DEF- and GLO-like genes. Heterodimerization may have originated after the gymnosperm-angiosperm split about 300 MYA but before the monocot-eudicot split 140–200 MYA. Heterodimerization may have become obligate for both types of flowering plant B proteins in the eudicot lineage after the monocot-eudicot split.

Introduction

The identity of the different floral organs during flower development is specified by the activity of floral organ identity genes. In the well-known “ABC model”, three classes of gene activities (homeotic functions) encoded by floral organ identity genes have been proposed, called A, B, and C (Weigel and Meyerowitz 1994). Within any one of the four flower whorls of the model plant Arabidopsis thaliana (thale cress; henceforth called Arabidopsis), expression of A alone specifies sepal formation. The combination AB specifies the development of petals, and the combination BC specifies the formation of stamens (male reproductive organs). Expression of the C-function alone determines the development of carpels, the female reproductive organs containing the ovules. The model also proposes that the A- and C-functions negatively regulate each other. Recently it has been demonstrated by a reverse genetics approach that yet another class of floral organ identity genes, termed class E genes (Theißen 2001), is involved in specifying petals, stamens, and carpels (Pelaz et al. 2000; Honma and Goto 2001).

In the flowers of Arabidopsis, the A-function is expressed in the first and second floral whorl, the B-function in the third and fourth whorl, and the E-function in whorls two, three, and four. Therefore, sepals, petals, stamens, and carpels are specified in whorls one, two, three, and four, respectively (for reviews see Weigel and Meyerowitz 1994; Theißen 2001).

Because the B-function genes, together with the A-, C- and E-function genes, are master control genes of petal and stamen development, loss of B-function leads to homeotic conversion of petals into sepaloid organs (expressing now only the A- and E-function) and of stamens into carpelloid organs (expressing now only the C- and E-function).

The genes providing the floral homeotic B-function have been cloned from several eudicotyledonous flowering plants such as Arabidopsis and Antirrhinum majus (snapdragon; henceforth called Antirrhinum) (for a review, see Weigel and Meyerowitz 1994; Theißen, Kim, and Saedler 1996; Theißen et al. 2000). In Antirrhinum, the B-function is provided by two different genes termed DEFICIENS (DEF) and GLOBOSA (GLO). In Arabidopsis the B-function is provided by APETALA3 (AP3), the putative ortholog of DEF, and PISTILLATA (PI), the putative GLO ortholog. The proteins encoded by these B-function genes are stable and functional in the cell only as heterodimers, i.e., DEF-GLO or AP3-PI complexes, respectively. These obligate heterodimers autoregulate their own expression at the transcriptional level.

Key words: MADS-box gene, homeotic gene, gymnosperm, angiosperm, evolution, dimerization.

Address for correspondence and reprints: Günter Theißen, Max-Planck-Institut für Züchtungsforschung, Abteilung Molekulare Pflanzenzüchtung, Carl-von-Linné-Weg 10, D-50829 Köln, Germany. E-mail: theissen@mpiz-koeln.mpg.de.

level by binding to specific sequence elements in the promoters of the genes which encode them and are thus coexpressed in petals and stamens during flower development. In line with this, heterodimerization is absolutely required for DNA binding in vitro (Schwarz-Sommer et al. 1992; Riechmann, Krizek, and Meyerowitz 1996); although immunoprecipitation analyses suggested that AP3 and PI may be able to form homodimers, these complexes are unable to bind to DNA (Riechmann, Krizek, and Meyerowitz 1996).

Because of a strict dependence of floral organ identity on floral homeotic gene function, the phylogeny of the B-function genes must have played an important role in the evolution of petals and stamens (Kramer and Irish 1999; Theißen et al. 2000). More insights into the phylogeny of B-function genes may thus help us to better understand the evolution of flowers.

All B-function genes identified so far, like most other floral homeotic genes, belong to the family of MADS-box genes, encoding MADS-domain transcription factors (for recent reviews about MADS-box genes in plants, see Riechmann and Meyerowitz 1997; Theißen et al. 2000). In flowering plants MADS-box genes constitute a huge multigene family with more than 80 different loci in Arabidopsis alone (Riechmann et al. 2000). Members of the MADS-box multigene family control diverse developmental processes ranging from root to flower and fruit development. A thorough characterization of the MADS-box gene family in phylogenetic informative taxa is thus needed for a better understanding of the origin and evolution of the floral homeotic B-function.

Phylogeny reconstructions revealed that the MADS-box gene family is composed of several defined gene clades (reviewed by Theißen et al. 2000). The vast majority of plant MADS-box genes known so far are members of a monophyletic superclade of genes with a conserved structural organization, the so-called MIKC-type domain structure, including a MADS (M-), intervening (I-), keratin-like (K-) and C-terminal (C-) domain (Münster et al. 1997). The highly conserved MADS-domain is the major determinant of DNA binding, but it also performs dimerization and accessory factor binding functions. The relatively weakly conserved I-domain may constitute a key molecular determinant for the selective formation of DNA-binding dimers (Riechmann and Meyerowitz 1997). The K-domain is defined by a conserved regular spacing of hydrophobic residues, which is proposed to allow for the formation of an amphipathic helix involved in protein dimerization (Riechmann and Meyerowitz 1997). The very variable C-domain at the C-terminus of the MADS-domain proteins is involved in transcriptional activation or the formation of multimeric transcription factor complexes (Cho et al. 1999; Egea-Cortines, Saedler, and Sommer 1999).

The MIKC-type gene superclade can be further subdivided into several well-defined gene clades whose members share similar expression patterns and highly related functions. Of special interest here is the fact that all angiosperm B-function genes known so far fall into either one of two different clades, namely DEF- or GLO-like genes (Theißen and Saedler 1995; Theißen, Kim, and Saedler 1996; Theißen et al. 2000). These clades have been named after the first clade members that have been molecularly characterized (as outlined in Theißen, Kim, and Saedler 1996). DEF- and GLO-like genes are closely related within the MADS-box gene family because these two clades together also represent a well-supported gene clade (Doyle 1994; Purugganan et al. 1995; Theißen, Kim, and Saedler 1996; Theißen et al. 2000). Putative orthologs of these genes (termed GGM2- and DAL12-like genes) have also been isolated from diverse gymnosperm species, such as the gnetophyte Gnetum gnemon and the conifers Norway spruce (Picea abies) and Monterey pine (Pinus radiata) (Mouradov et al. 1999; Sundström et al. 1999; Winter et al. 1999). This indicates that the clade of DEF-, GLO-, GGM2-, and DAL12-like genes (termed “B genes” hereafter, irrespective of the function of these genes or the proteins they encode) originated more than 300 MYA. In contrast, B genes have not been found in ferns so far. This suggests that the B gene clade originated 300–400 MYA but does not completely rule out an earlier origin of B genes (Theißen et al. 2000). Anyway, the establishment of B genes in a common ancestor of extant seed plants by gene duplication, sequence diversification, and fixation was probably an important step towards the establishment of the floral homeotic B-function (Theißen, Kim, and Saedler 1996; Theißen et al. 2000). Phylogeny reconstructions and evaluation of exon-intron structures (K.-U. Winter and G. Theißen, unpublished data that are, however, available upon request) indicate that the gene duplication which led to distinct clades of DEF- and GLO-like genes occurred in the lineage that gave rise to extant angiosperms after the lineage that led to extant gymnosperms had already branched off. Within the eudicots, further duplications of DEF- and GLO-like genes occurred several times independently in different lineages (Kramer, Dorit, and Irish 1998; Kramer and Irish 1999, 2000).

Just one DEF-like gene has been found so far in diverse monocots such as lily (Lilium regale), wheat (Triticum aestivum), maize (Zea mays), and rice (Oryza sativa), so that there is no evidence that this gene has been duplicated during monocot evolution (Münster et al. 2001). In contrast, two different GLO-like genes have been reported from rice (OSMADS2 and OSMADS4) (Chung et al. 1995) and lily (LRGLOA and LRGLOB) (Theißen et al. 2000; A. Kanno-Shiotsuki and G. Theißen, unpublished data) and even three different genes from maize (ZMM16, ZMM18, ZMM29) (Münster et al. 2001).

We are intrigued by the observation that genes whose products have to heterodimerize to be stable and functional, such as DEF-GLO and AP3-PI, are each others closest relatives within their genomes (Theißen, Kim, and Saedler 1996; Theißen et al. 2000). This observation suggests that the respective genes originated by gene duplication and that heterodimerization is of relatively recent evolutionary origin and evolved from homodimerization. To test this hypothesis, we investi-
gated the dimerization behavior of B proteins from phylogenetic informative taxa, including a gymnosperm and a monocotyledonous flowering plant. Our data indicate that following a gene duplication within the B gene clade, obligate heterodimerization evolved in two steps from homodimerization via facultative heterodimerization. The importance of these findings for our understanding of the evolution of flowers is discussed.

Materials and Methods

In Vitro DNA-binding Assays

Plasmids derived from the pSPUTK in vitro transcription-translation vector (Stratagene) to produce AP3 and PI proteins have been described previously (Riechmann, Krizek, and Meyerowitz 1996). A vector to produce a C-terminal deletion derivative of AP3, termed ΔAP3 here, was also generated via PCR and subcloning into pSPUTK. pT7βSal-derivatives to produce ΔDEF and ΔGLO were provided by M. Egea-Cortines. The inserts had been subcloned into pSPUTK via PCR. The proteins produced from these vectors are truncated at their C-terminal ends. ΔDEF and ΔGLO proteins have a length of 152 amino acids, in contrast to the full length of the wild-type proteins, which is 227 (DEF) or 215 (GLO) amino acids. All primers used during PCR cloning procedures are available upon request.

In case of full-length GGM2, the NcoI-BamHI site of pSPUTK was used. All relevant regions were sequenced on both the strands by the MPIZ DNA core facility (ADIS) on PE Biosystems ABI Prism 377 and 3,700 sequencers using BigDye-terminator chemistry.

Proteins were synthesized using the TnT SP6 coupled reticulocyte lysate system (Promega) according to the manufacturer’s instructions. They were always [35S]methionine-labeled and were produced in the expected sizes and in similar amounts, as tested on 12% SDS-PAGE gels, with which X-ray films were exposed after gel run. Binding assays were carried out using double-stranded oligonucleotides containing CAR-g-box sequences from floral homeotic B-function genes. CAR-G1 (5'-GGC AAC TCT TTT CCT TTT TAG GTC GAA TAT GGT C-3') was derived from the Antirrhinum DEF promoter, CAR-G3 (5'-GGA TTA GGC AAT ACT TTC CAT TTT TAG TAA CT-3') was derived from the Arabidopsis AP3 promoter. CAR-g-box sequences are in italics. A deviation from the SRE-type CAR-G-box consensus (CC(A/T)nGG) in CAR-G3 is highlighted in bold. In tests for binding specificity, some CAR-G1 derivatives were employed. AntiCAR-G1 (5'-GGC AAC TCT TTT GGA TGC ATC CTC GCA TAT GGT C-3') contains several sequence deviations from CAR-G1 (bold) within the CAR-g-box (italics), most of which are not compatible with SRE-type CAR-g-box consensus (underlined). MutCAR-G1 (5'-GGC AAC TCT TTT CCT TGT TAG GTC GCA TAT GGT C-3') deviates from CAR-G1 only by a single nucleotide change (in bold). In MetCAR-G1 (5'-GGC AAC TCT TTT CTI TTT TAG GTC GCA TAT GGT C-3'), two single nucleotide substitutions at two sites (in bold) have transformed the SRE-type CAR-g-box into a N10-type CAR-g-box (consensus: C(A/T)nG). The double-stranded oligonucleotides contained a 5' protruding G nucleotide. They were labeled with 32P by Klenow fill-in reaction and purified by PAGE (10% gels) before they were used in DNA-binding experiments.

In vitro translated proteins were tested for DNA binding activity by gel retardation assays. Binding conditions were as described (Egea-Cortines, Saedler, and Sommer 1999). Total reaction volumes were always 12 µl, containing about 1 ng of labeled, double-stranded oligonucleotides, and roughly the same amounts of proteins. Within an experimental series, aliquots were used except for the components that varied. Reactions were incubated for 20 min on ice and loaded on a running 4% polyacrylamide:bisacrylamide (30:1) gel, which had been pre-run at 100 V in 1× TBE. The gel was run at 150 V. After run, the gel was transferred to paper, dried, and used to expose an X-ray film.

Yeast Two-hybrid Analysis

EcoRI-SalI fragments containing the full-length coding region of GGM2, GGM15, AP3 and PI were generated via PCR amplification and cloned into the binding domain vector pBDGAL4 and into the activation domain vector pADGAL4, yielding pBD(AGM2, pAD(AGM2, pAD(AGM15, pAD(AG3, and pADpγ. Both vectors were purchased from Stratagene. Proper fusion of the constructs was checked by sequence determination. The yeast strain YRG-2 (MATa, ura3-52, his3-200, ade2-101, trp1-901, leu2-3, 112, gal4-542, gal80-538, LYS2::GAL4, TRP1::GAL1-TATA, HIS3, URA3::GAL1-lact-ac, TATA-CYC1-lacZ) was purchased from Stratagene and transformed with pBD(AGM2, pAD(AGM2, pAD(AGM15, pAD(AG3, or pADpγ according to the manufacturer’s instructions. The single transformants were checked for HIS3 reporter gene expression by incubation on synthetic dropout (SD) medium (SD-Trp-His or SD-Leu-His, respectively) lacking histidine and either tryptophan or leucine, respectively. None of the single transformants was found to activate the HIS3 reporter gene expression, indicated by no growth on medium lacking histidine. The PBD transformants were used for transformation of 20 µg of the corresponding pAD plasmids, following the manufacturer’s instructions. The transformation mixture was plated on SD medium (SD-Trp-Leu and SD-Trp-Leu-His) to monitor the transformation efficiency and protein-protein interaction. Colonies growing on SD-Trp-Leu-His plates were transferred to SD-Trp-Leu-His plates to reconfirm HIS3 reporter gene expression and then transferred to nitrocellulose filters. These filters were used to analyze the second reporter gene expression by testing β-galactosidase activity. Plasmids from selected yeast colonies were recovered according to the manufacturer’s instructions, and inserts were amplified by PCR and sequenced.

Results

Sequence-specific DNA-binding of GGM2 Protein Homodimers to CAR-g-boxes

The MADS-domain proteins from flowering plants that have been analyzed so far bind as dimers or mul-
FIG. 1.—GGM2 alone is able to bind to CArG-box DNA sequence elements. DNA binding of in vitro–translated proteins as revealed by gel retardation assays is shown. In these experiments, radioactively labeled DNA probes, termed CArG1 or CArG3, as indicated, were incubated with in vitro–translated proteins as shown above the lanes and then subjected to gel electrophoresis. Empty vector symbolizes negative controls where the expression vector did not contain a cDNA insert, but all other experimental conditions were identical to the reactions where in vitro translated MADS-domain proteins were generated. Lysate indicates an additional negative control where unprogrammed reticulocyte lysate was incubated with the probes.

FIG. 2.—A) Competition gel retardation assay: Lysate or empty vector was incubated with in vitro–translated GGM2 followed by gel retardation assay. Lysate indicates an additional negative control where unprogrammed reticulocyte lysate was incubated with the probes. AP3 and PI used as a control are only able to bind DNA as a heterodimeric complex (fig. 1). Very similar results were obtained with probe CArG3, a stretch from the Arabidopsis AP3 promoter comprising a CArG-box termed CArG3 (fig. 1). CArG3 deviates at one position from the SRE-type CArG-box consensus (see Materials and Methods).

Sequence-specificity of GGM2 binding to CArG-boxes was tested in two different experiments, employing sequence variants of CArG-boxes. In gel retardation assays using mutated CArG-box sequences, different sequence changes that lead to a deviation from the SRE-type CArG-box consensus, including a single nucleotide exchange abolished DNA binding beyond recognition (fig. 2A). Under these conditions also no binding to an N10-type CArG-box (termed MefCarG) was observed (fig. 2A). This type of CArG-box (consensus: C[A/T]$_6$G) is bound by some animal MADS-domain proteins, such as MEF2A from human, a MEF2-type MADS-domain protein involved in muscle development (Shore and Sharrocks 1995).

Under the previously used conditions, only stable, long-term binding of proteins to DNA can be observed; therefore, we also carried out a more sensitive test for DNA binding. In this experiment, GGM2 binding to the wild-type probe CarG1 oligonucleotide was assayed in the presence of about 125-fold surplus of unlabeled wild-type or mutated CarG-box DNA. Binding to the unlabeled DNA was inferred from the reduction of the binding to the labeled DNA. Whereas strong competition for GGM2 binding was observed with the wild-type CarG1 sequence, a highly mutated sequence termed AntiCarG1 apparently could not compete in binding GGM2 (fig. 2B), demonstrating that the DNA binding of GGM2 is sequence-specific. A weaker competition than with wild-type CarG1 sequence was observed with MutCarG1, a CarG1 derivative where one nucleotide exchange leads to a deviation from the SRE-type consensus sequence (fig. 2B). Even weaker, but still clearly visible competition, indicating weak GGM2 binding to the unlabeled, mutant DNA, was found when MefCarG1 was used, a CarG1 derivative where two nucleotide substitutions lead to an N10-type CarG-box (fig. 2B). These data demonstrate that the more a sequence deviates from the SRE-type CarG-boxes consensus, the less it is able to compete for GGM2 binding. Thus GGM2 binds preferentially to SRE-type CarG-box sequences, like most MADS-domain proteins from flow-
FIG. 2.—GGM2 binds to the CArG-box of probe CArG1 in a sequence-specific manner. Lysate and Empty vector denote negative controls (see legend to fig. 1). A, Binding of GGM2 to radioactively labeled probe CArG1, or mutant derivatives thereof, as indicated below the lanes. B, Binding of GGM2 to radioactively labeled probe CArG1 was challenged by a 125-fold surplus of unlabelled CArG1 or mutant derivatives thereof (marked by "\^"), as indicated above the lanes.

Comparison of the gel electrophoretic mobility of the GGM2-DNA complex with that of the AP3-PI-DNA complex strongly suggests that GGM2 binds as a homodimer to both CArG1 and CArG3 (fig. 1). The ability of GGM2 to homodimerize even without binding to a CArG-box was demonstrated by yeast two-hybrid assays. YRG-2 yeast colonies, transformed with a bait-and prey-vector, both containing the full-length GGM2 cDNA fused in frame to the GAL4 activation domain or the binding domain, respectively, grow on synthetic media (SD-Trp-Leu-His) lacking histidine and therefore activate HIS3 reporter gene expression (fig. 3C). Colonies transferred on nitrocellulose filters and tested for activation of the second reporter gene β-galactosidase showed strong β-galactosidase activity (data not shown). To summarize these data, GGM2 very likely binds CArG-box DNA as homodimer in a sequence specific manner, whereas class B proteins from higher eudicots like Arabidopsis and Antirrhinum usually bind DNA as heterodimers.

In order to test the potential of GGM2 to interact with class B proteins of eudicotyledonous flowering plants, gel retardation assays with cotranslated GGM2 and truncated angiosperm class B proteins were carried out. These experiments indicated that GGM2 prefers to homodimerize, rather than to heterodimerize with AP3 (fig. 3B), DEF, or GLO (fig. 3A), at least under our in vitro conditions. This is indicated by the fact that no bands reflecting gel electrophoretic mobilities between those of the complexes containing only gymnosperm or angiosperm proteins were obtained. Our data demonstrate that the homodimerization of GGM2 is not simply because of a lack of any other B protein, but is a more specific event.

In line with this, in yeast two-hybrid assays, the GGM2 protein fused to the GAL4 DNA-binding domain does not interact with either PI or AP3, each fused to the activation domain of GAL4, indicated by no growth on synthetic media lacking histidine (SD-Trp-Leu-His) (fig. 3C). This suggests that a primary inability in protein-protein interactions is at least one of the reasons why GGM2 does not form DNA-binding complexes.
FIG. 3.—GGM2 homodimerizes rather than heterodimerizes with angiosperm B proteins. In (A) and (B) DNA binding of in vitro translated proteins as revealed by gel retardation assays is shown. Proteins with shortened C-terminal domains are indicated by Δ. Lysate and Empty vector denote negative controls (see legend to fig. 1). The probes used are denoted below the lanes. A band representing a ΔDEF-ΔGLO heterodimer is marked by “*.” A band representing a GGM2 homodimer is marked by “<.” In (C) protein dimerization as indicated by yeast two-hybrid analyses is shown. Yeast colonies cotransformed with plasmids containing GGM2 cDNA fused to the binding domain of GAL4 and either AP3, PI, GGM15 or GGM2 (as indicated) fused to the activation domain of GAL4 all grow on synthetic media lacking tryptophan and leucine (SD-Trp-Leu). But only colonies containing GGM2 fused to the binding domain as well as GGM2 fused to the activation domain of GAL4 are able to grow on synthetic media also lacking histidine (SD-Trp-Leu-His), indicating a homodimerization of the two GGM2 proteins.
with Pl and AP3. Also, no interaction with GGM15, the second class of B proteins from Gnetum (DAL12-like gene) (Becker et al. 2000), was observed (fig. 3C), which makes it even more likely that GGM2 functions as a homodimer also in planta.

**B Class Proteins from Lily Bind DNA as Homo-as Well as Heterodimers**

Similar to AP3 from Arabidopsis and DEF from Antirrhinum, the DEF-like protein from the monocotyledonous plant lily (L. regale) LRDEF does not bind the CArG3 probe without an appropriate heterodimerization partner (fig. 4). In contrast, the lily GLO-like proteins LRGLOA and LRGLOB are both able to bind the CArG3 probe alone (fig. 4). DNA binding is also observed when LRDEF and PI are cotranslated. Because LRDEF alone as well as PI alone are not able to bind to the probe, one has to postulate a heterodimer i.e., a protein complex of the lily DEF-like protein LRDEF and the GLO-like protein PI of Arabidopsis. DNA binding is also observed when LRDEF is cotranslated with one of the GLO-like proteins from lily LRGLOB (fig. 4). In this case, two bands with similar gel electrophoretic mobilities are visible, representing the LRGLOB homodimer and LRGLOB-LRDEF heterodimer, respectively (fig. 4).

**Discussion**

Interesting questions arise by the finding that two relatively closely related transcription factors such as the class B floral organ identity proteins of higher eudicotyledonous flowering plants require heterodimerization to be functional. What are the mechanistic implications of such a system? And how did it originate and evolve?

The identification of putative orthologs of class B genes in monocots and even in diverse gymnosperms provided a starting point for the studies on the origin of the obligate heterodimeric system. Recent phylogeny reconstructions strongly support the view that the DEF- and GLO-like genes are closely related paralogous gene lineages (Doyle 1994; Purugganan et al. 1995; Theissen, Kim, and Saedler 1996; Theissen et al. 2000) which originated by a gene duplication event near the base of the angiosperms, and that the B genes from gymnosperms are thus ancestral to both clades of B genes in angiosperms (see Introduction). As the relationships of the
The evolutionary scenario outlined in figure 5 raises the question as to the functional significance of the observations made in vitro or in yeast with respect to what happens inside of the plant nucleus, where MIKCC-type transcription factors are assumed to carry out their function. There are a number of cases where we have information about dimerization or even multimerization of MIKCC-type MADS-domain proteins as observed by gel retardation assays and the yeast two-hybrid system on the one hand and genetic data about gene interaction on the other. Examples are the heterodimerization of DEF-GLO and AP3-PI and the formation of SQUA-SQUA-DEF-GLO and AP3-PI-AG-SEP multimers. In all these cases we have an impressive compatibility between the in vitro and yeast in vivo data and the in planta data, as revealed by genetics (Schwarz-Sommer et al. 1992; Tröbner et al. 1992; Riechmann, Krizek, and Meyerowitz 1996; Egea-Cortines, Saedler, and Sommer 1999; Honma and Goto 2001; Theißen 2001; Theißen and Saedler 2001). Therefore, we are quite confident that the dimerization behavior that we have observed for the B proteins reflects their in planta behavior and is thus of functional significance. Direct evidence for that, however, awaits future studies. One should also take into consideration that inside of the plant nucleus, multimers rather than dimers may be formed (Egea-Cortines, Saedler, and Sommer 1999; Honma and Goto 2001; Theißen 2001; Theißen and Saedler 2001), a feature of MIKCC-type proteins that is beyond the scope of the studies reported here.

The capacity of GLO-like proteins from monocotyledonous flowering plants to homodimerize makes it conceivable that also these homodimers, and not only B gene heterodimers, have a function in monocots, as suggested earlier (Münster et al. 2001). In line with this, a considerable disjunction of the expression patterns of DEF- and GLO-like genes can be observed in monocots, basal angiosperms, and even lower eudicots (Moon et al. 1999; Kramer and Irish 2000; Münster et al. 2001). In maize and rice, for example, the GLO-like genes are expressed in carpels as well as stamens and lodicules, whereas the expression of the single DEF-like genes in these species is restricted to stamens and lodicules. These findings may indicate that upregulation of the GLO-like genes in these grasses does not require the B protein heterodimer. Moreover, they make it conceivable that homodimers of GLO-like proteins have a function in diverse angiosperms outside of the higher eudicots (Krämer and Irish 2000; Münster et al. 2001). In maize and rice, this function would be operational in carpels, and thus would be different from the classical B-function, in grasses specifying lodicules (rather than petals),
and stamens (reviewed by Theißen et al. 2000). However, moderate differences in their expression domains are also known for the class B genes from *Arabidopsis* and *Antirrhinum* (Samach et al. 1997), indicating that the differences between higher eudicots and other angiosperms are quantitative rather than absolute.

Why did obligate heterodimerization evolve at all? The redundancy implied by the duplication of an ancestral B gene may have led to relaxed constraints in sequence conservation; thus mutations may have occurred upon which the ability to form DNA-binding homodimers may have been impaired. For example, one protein may have lost the ability to form homodimers, whereas the other lost the ability to bind DNA alone. Thus the system may have been locked into a state in which the two B genes became mutual interdependent. This raises the question as to whether obligate heterodimerization remained just a neutral change that became fixed or whether it provided a selective advantage? Because the phenomenon is found in the most successful group of flowering plants, the eudicots, one is tempted to assume that the latter is true. This appears quite possible if obligate heterodimerization is considered in combination with positive autoregulation, i.e., the well-known phenomenon that, e.g., DEF-GLO heterodimers upregulate the transcription of both the DEF and the GLO gene by binding to the promoters of both the genes. The selective advantage of a combination of both could be the fixation of class B gene expression patterns and thus the spatial domain within the flower of the floral homeotic B-function during evolution. Mutational changes in the promoter region of only one gene may not lead to changes in the expression domain of the class B genes because expression of the other partner, and hence upregulation, is missing in the ectopic expression domain. Only parallel changes in both types of class B genes, which are certainly less likely than changes in single genes, will lead to ectopic expression of the B-function under the assumption of obligate heterodimerization and strong autoregulation.

It thus appears as an intriguing hypothesis that obligate heterodimerization may have evolved in parallel, or even as a prerequisite, of the canalization of floral development and thus standardization of floral structure at the base of higher eudicots (Endress 1987; Kramer and Irish 2000). The vast majority of higher eudicots has sepaloid organs in the first floral whorl and petaloid organs in the second whorl (with these terms referring to organ identity, not to the position in the flower). In contrast, the flowers of the basal angiosperms, monocots, and lower eudicots are much more plastic with respect to sepaloidy and petaloidy of their organs, suggesting frequent changes in the expression domain of the B-function during evolution (Albert, Gustaffson, and Di Laurenzio 1998; Kramer and Irish 2000). Less strict requirements for heterodimerization or autoregulatory upregulation may facilitate spatial shifts in the B-function in taxa outside of the eudicots. It appears thus as one interesting possibility that coevolution with highly specialized pollinators may have selected for more standardized flowers and thus an evolutionarily more stable expression of the class B genes. This may have been the selective force that locked the interaction of class B proteins into the mode of obligate heterodimerization.

**Acknowledgments**

We are grateful to Marcos Egea-Cortines for help with the gel retardation assays, Thomas Münster for organizational lab work, and Annette Becker for valuable comments on the manuscript. We thank the Automatic DNA Isolation and Sequencing team of our institute for sequencing and Britta Grosardt for excellent technical assistance. We are grateful to two anonymous reviewers who provided a number of helpful suggestions for improvement on an earlier version of the manuscript. Financial support from the DFG to G.T. (Th 417/3-2) is highly acknowledged. A.K. was supported by a Japanese Society for the Promotion of Science (JSPS) Postdoctoral Fellowship for Research Abroad.

**LITERATURE CITED**


WILLIAM MARTIN, reviewing editor

Accepted October 8, 2001