An ACT-like Domain Participates in the Dimerization of Several Plant Basic-helix-loop-helix Transcription Factors*

Antje Feller1, J. Marcela Hernandez5, and Erich Grotewold1,4,2

From the 1Molecular, Cellular, and Developmental Biology Program, 2the Ohio State Biochemistry Program, and 3the Department of Plant Cellular and Molecular Biology and Plant Biotechnology Center, Ohio State University, Columbus, Ohio 43210

The maize basic-helix-loop-helix (bHLH) factor R belongs to a group of proteins with important functions in the regulation of metabolism and development through the cooperation with R2R3-MYB transcription factors. Here we show that in addition to the bHLH and the R2R3-MYB-interacting domains, R contains a dimerization region located C-terminal to the bHLH motif. This protein-protein interaction domain is important for the regulation of anthocyanin pigment biosynthesis by contributing to the recruitment of the C1 R2R3-MYB factor to the C1 binding sites present in the promoters of flavonoid biosynthetic genes. The R dimerization region bares structural similarity to the ACT domain present in several metabolic enzymes. Protein fold recognition analyses resulted in the identification of similar ACT-like domains in several other plant bHLH proteins. We show that at least one of these related motifs is capable of mediating homodimer formation. These findings underscore the function of R as a docking site for multiple protein-protein interactions and provide evidence for the presence of a novel dimerization domain in multiple plant bHLH proteins.

Proteins containing the basic-helix-loop-helix (bHLH)3 domain compose one of the largest transcription factor families in plants (1–4). The bHLH signature that defines the family is constituted by an N-terminal ~16-amino acid-long basic α-helix that binds DNA to the canonical E-box (CANNTG) (5) and a C-terminal helix-loop-helix (HLH) domain involved in homo- and/or heterodimerization (6). Some factors, however, such as the Id myogenic regulator, lack the basic region and function as inhibitors by forming heterodimers that cannot bind DNA (7). It is common for bHLH proteins to contain additional protein-protein interaction domains that contribute in unique ways to their regulatory function (8). For example, the Myc proto-oncoprotein forms heterodimers with Max through the respective bHLH and adjacent basic leucine zipper (bZip) domains (9). In contrast to Myc, which cannot homodimerize, Max can form homo- or heterodimers with several related proteins, including Mad1 and Mnt (10). The bHLH region of Myc mediates the interaction with Miz-1, a POZ transcription factor that permits Myc to bind and repress promoters lacking the CACGTG E-box (10, 11). Plant bHLH proteins are also characterized by the presence of several conserved domains in addition to the bHLH motif. For example, the analysis of the 133 Arabidopsis bHLH factors uncovered 40 or more domains present in three or more proteins (3). By and large the function of these domains is not known. Among the few for which functions have been identified are the N-terminally located small APB domain present in several phytochrome-interacting factors (12) and the region that mediates the interaction with R2R3-MYB factors (13, 14) central to providing R2R3-MYB transcriptional regulators with very similar DNA binding preferences with the ability to control distinct sets of target genes in vivo (15).

The bHLH/R2R3-MYB cooperation is best exemplified by the interaction between the N-terminal region of a member of the R/B group of bHLH regulators and the R2R3-MYB domain of C1 (13). The interaction with R is essential for the ability of C1 to activate all known flavonoid biosynthetic genes, resulting in anthocyanin pigment accumulation (16). A second function of R was uncovered using a mutant of the R2R3-MYB P1 regulator (P1*) (17). P1 normally regulates a subset of the C1/R-regulated genes (18). Among the P1-regulated genes is A1, which encodes dihydroflavonol reductase, an enzyme necessary for the formation of the anthocyanins (controlled by C1 + R) and the phlobaphenes (controlled by P1) (19). Different from P1, P1* interacts with R (17), and in the presence of R, the P1* regulatory activity is enhanced. This R-enhanced activity requires the anthocyanin regulatory element (ARE) (16) present in the promoter of A1 and other anthocyanin biosynthetic genes (20).

Based on the sequence homology of the bHLH domain and the presence of conserved N-terminal regions, R belongs to group III of the plant bHLH gene family (3). This group is shared by the Arabidopsis GL3 (GLABROUS3), EGL3 (ENHANCER OF GLABROUS3), TT8 (TRANSPARENT TESTA8) and AtMYC1 proteins, which participate in trichome and root hair formation and in the control of flavonoid pigments (21–25), functions that can be complemented by the maize R gene, Lc (26). Similar to the interaction of R with C1,
members of this group of bHLH proteins function by interacting with R2R3-MYB factors. For example, GL3 and EGL3 interact with the related R2R3-MYB GL1 (GLABROUS1) and WER (WERWOLF) proteins to control trichome or root hair production, respectively, or with PAP1 (PRODUCTION OF ANTHOCYANIN PIGMENT1) to control anthocyanin accumulation (27). Similarly, TT8 interacts with the R2R3-MYB TT2 (TRANSPARENT TESTA2) to activate proanthocyanidin accumulation in the seed coat (22). The picture that emerges from these and other studies is that the R2R3-MYB factors are responsible for providing the specificity for a particular process, whereas the bHLH factors play more pleiotropic roles, being shared between two or more cellular processes (16, 27). Given the multiplicity of conserved domains in R and R-related proteins and what is known on the mechanisms by which animal bHLH proteins function, the question remains as to whether regions in R participate in hetero- or homodimer formation and how these interactions contribute to R function.

Here we show that, similar to other bHLH proteins, R contains a dimerization domain that can direct the formation of homodimers in vitro and in vivo. This dimerization domain is necessary for the R regulatory activity, as demonstrated by mutants that exhibit a significant reduction of the activation of maize flavonoid genes and anthocyanin accumulation. The dimerization region of R is C-terminal to the bHLH motif and has structural similarity to the ACT domain involved in the allosteric regulation of many amino acid metabolic enzymes. Structural homology searches identified other bHLH factors with similar ACT domains, and we show that at least one of them can also mediate the formation of specific dimers. These findings highlight the role of a novel dimerization domain for the regulatory activity and provide evidence for the recruitment of eukaryotic transcription factors of protein-protein interaction domains characteristic of metabolic enzymes.

**MATERIALS AND METHODS**

**Plasmids Used in Transient Expression Experiments**—All plant expression vectors include the cauliflower mosaic virus 35S promoter, the tobacco mosaic virus Ω′ leader and maize first Adh1-S intron in the 5′-untranslated region, and potato proteinase II (pinII) termination signal unless otherwise specified. Previously described plasmids include p35S::C1, pBz1::Luc, pA1::Luc, pA1mPBS::Luc (18, 33), pA2::Luc (20), pBz2::Luc (47), and pBz1::Luc, pA1::Luc, pBz1::Luc, pA1mPBS::Luc, or pGal4::Luc were used in each bombardment. To normalize luciferase activity to GUS activity, 3 μg of pUBI::GUS was included in every bombardment. Each treatment was done at least in triplicate, and entire experiments were repeated at least twice. The assays for luciferase and GUS and the normalization of the data were done as described (18). Data are expressed as the ratio of arbitrary light units (luciferase) to arbitrary units of fluorescence (GUS).

**Yeast Two-hybrid Experiments**—The plasmid containing R411−610, R411−462, or R525−610 in the pBD-GAL4 (TRP+) vector and R411−610, R525−610, or RIF-1C cloned into pAD-GAL4 (LEU+) vector were cotransformed into yeast strain PJ69.4a (51) and plated on SC-LEU-TRP medium. Colonies were then screened for growth on SC-LEU-TRP, SC-LEU-TRP-HIS, and SC-LEU-TRP-HIS-ADH. β-Galactosidase assays were performed on the pl69.4 strain on three separate cultures carrying each set of plasmids (biological triplicates). Cells were grown in selective Leu−Trp media overnight and then used to inoculate 10 ml of yeast extract/peptone/dextrose-rich media. When cultures reached an A600 of ~0.8, the cells were collected and lysed in 0.1 M Tris-HCl, pH 8.0, 20% glycerol, and 1 mM dithiothreitol using glass beads on a Mini-Beadbeater (Biospec Products). β-Galactosidase assays were carried out essentially as described previously (52), and β-galactosidase units were calculated using the formula (A420 × 377.8)/(time of incubation × volume of extract × protein concentration in mg/ml).

**GST Pull-down Experiments**—The GST pull-down bait was made by cloning the R C-terminal region (residues 411−610) into the vector pGEX-KG (53) as a XhoI-HindIII fragment. The pGEX-KG vector was used as a negative control. The R cDNA was excised and cloned into pBluescript (Stratagene). The

---

ACT Dimerization Domains in Plant bHLH Factors

GST-R^{411–610} and pGEX-KG constructs were each transformed into *Escherichia coli* BL21 (DE3) *PlyS* cells for expression. Cultures were grown, induced with IPTG and purified essentially as described (54), with the following modifications: After induction of a 500 ml culture with 1 mM isopropyl-1-thio-β-D-galactopyranoside, the cells were harvested by centrifugation and stored at −80 °C until further use. The cells were resuspended in 10 ml of phosphate-buffered saline buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na2HPO4, 2 mM KH2PO4, 1 mM dithiothreitol, 1 mM isopropyl 1-thio-β-D-galactopyranoside) and passed twice through a French press. The cell lysate was centrifuged at 4000 × g for 20 min; the supernatant was filtered through two layers of Miracloth (Calbiochem).

The GST pull-down protocol previously described was used (55). Glutathione-Sepharose beads (Novagen) were prewashed with NETN150 buffer (0.5% Nonidet P-40, 0.1 mM EDTA, 20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride) 4 times and equilibrated in NETN150 buffer. The beads were coated by incubating the bacterial cell extract (100–200 μl) containing GST (used as negative control) or GST-R^{411–610} with 20 μl of glutathione-Sepharose beads at 4 °C and nutating for 1 h. The beads were washed 4 times by nutating with 1 ml of NETN150 at 4 °C (20 min/ wash) and then resuspended in 20 μl of NETN150. Proteins were *in vitro* transcribed and translated in the presence of RediVue[^35]S-methionine *in vitro* translation grade (Amersham Biosciences) using the TNT T7/T3 coupled wheat germ extract system (Promega, Madison, WI). The pull-down reactions were done with 20-μl samples of coated beads to which 100 μl of NETN150 buffer was added together with 5 μl of labeled protein and incubated at 4 °C nutating for 1 h or overnight. Reactions included beads loaded with GST-R^{411–610} and GST, used as a negative control. The beads were washed 4 times (nutation at 4 °C for 5 min) with 1 ml of NETN150 buffer. The bound proteins were eluted by boiling in 10 μl of 2× SDS sample buffer and visualized using Coomassie Blue staining followed by autoradiography.

**Plant Transformation and Confocal Microscopy**—Plasmids corresponding to the GFP fusions were transformed into *Agrobacterium tumefaciens* strain GV3101, and infiltration was performed as described by Voinnet et al. (56) with modifications. In brief, bacteria strains carrying binary constructs were grown at 29 °C to stationary phase in LB media supplemented with antibiotics, acetosyringone, and MES. Cells were harvested and resuspended in media containing MgCl2, MES, and acetosyringone giving a final A_{600} of 0.9–1.1. After 3 h at room temperature, co-transfection of 3–4-week-old *N. benthamiana* plant leaves with *Agrobacterium* cells containing the R-GFP or R^{Δ532–560}-GFP plasmid and a plasmid expressing the tomato bushy stunt virus (pBlin61) silencing suppressor (56) were performed. After 3 days infiltrated leave areas were excised, and localization of GFP was determined by confocal laser scanning microscopy (Nikon Eclipse E600, Japan) using the same gain (2000).

**Analysis of Dimerization Domain Structure**—The structural analysis of the C-terminal dimerization domain of the R-like proteins and the other *Arabidopsis* proteins was done using the FUGUE Version 2.8.07 software (updated on May 2, 2002) (34). The sequences of the C-terminal domains of R, GL3, JAF13, AN1, NA11, At2g46810, At1g49770, At2g31210, At5g65640, At1g32640, At4g16430, At5g54680, At1g68810, At12g7740, At3g19500, At4g30980, and At1g73830 were individually analyzed with FUGUE, which used a fold library of size 6348. The length of the probe sequence varied between 79 and 88 amino acids. The probe divergence values were 0.577 for R and GL3, 0.656 for JAF13, 0.735 for NA11, 0.586 for AN1, 0.779 for At2g46810, 0.730 for At1g49770, 0.445 for At2g31210, 0.759 for At5g65640, 0.794 for At1g32640, 0.76 for At4g16430, 0.708 for At5g54680, 0.755 for At1g68810, 0.266 for At1g7740, 0.555 for At5g19500, 0.413 for At4g30980, and 0.302 for At1g73830. The recommended cutoff values were as follows: ZSCORE ≥ 6.0 (CERTAIN 99% confidence), ZSCORE ≥ 4.0 (LIKELY 95% confidence), ZSCORE ≥ 3.5 (MARGINAL 90% confidence), ZSCORE ≥ 2.0 (GUSS 50% confidence), and ZSCORE < 2.0 (UNCERTAIN). After the scores were obtained for all R-like proteins, three common structural hits were chosen and aligned with the dimerization domain sequences. The MEME/MAST system was used with the sequences mentioned above to generate a motif that was then utilized as a probe to search the nr data base available at the same site.

**RESULTS**

**R Contains a C-terminal Dimerization Domain**—To investigate whether the C-terminal region of R (residues 411–610, R^{411–610}, Fig. 1), which includes the bHLH domain, was capable of mediating the formation of dimers, we fused this region to the GAL4 DNA binding domain (GAL4DBD) and GAL4 activation domain (GAL4AD) and tested them for interaction in yeast. Growth in media lacking leucine and tryptophan (for the selection of the plasmids) and histidine and adenine (−Leu −Trp −His −Ade) is indicative of the formation of interacting partners (Fig. 2A, #1), providing evidence that the C-terminal region of R contains a dimerization domain. The equivalent regions of GL3 and EGL3 were also shown to interact with each other and to mediate homodimer formation in yeast two-hybrid experiments (27), suggesting that the presence of a C-terminal dimerization domain is a general feature of this group of bHLH proteins.

To determine whether this dimerization was mediated by the bHLH motif, residues 411–462 (bHLH, Fig. 1A) were fused to the GAL4DBD and tested for interaction with R^{411–610}, GAL4AD. Under the conditions used no interaction was observed (Fig. 2A, #3). To ensure that the R^{411–462}-GAL4DBD protein is properly expressed, we tested this protein for interaction with RIF-1C-GAL4AD, where RIF-1C corresponds to a maize R-interacting factor that interacts with the bHLH region of R. The growth observed in −Leu −Trp −His −Ade indicates that R^{411–462}-GAL4DBD is properly expressed (Fig. 2A, #4). Deletion analyses of R^{Δ532–610} demonstrated that the R^{525–610} region was necessary and sufficient for homodimer formation in yeast (Fig. 2A, #5) and does not activate on its own when fused to the GAL4DBD (Fig. 2A, #6).

To biochemically verify the ability of R^{411–610} to mediate homodimer formation, the full-length R protein was *in vitro* transcribed/translated in the presence of [35S]Met ([35S] R, Fig.
ACT Dimerization Domains in Plant bHLH Factors

JOURNAL OF BIOLOGICAL CHEMISTRY

in yeast can also happen (29), providing evidence that the dimerization observed (Fig. 2B) is a general phenomenon that occurs when transcriptional activation domains are fused to a heterologous DNA binding domain (29). When GST-R411–610 (Fig. 2B, lane 2) is used as a control, the pull-down experiment was performed in parallel with GST (Fig. 2B, lane 3), resulting in the efficient recovery of the ∼79-kDa R protein (Fig. 2B, lane 1) after autoradiography (Fig. 2B, top panel), providing evidence that the dimerization observed in yeast can also happen in vitro.

To establish whether the C-terminal region of R can mediate homodimer formation in maize cells, we took advantage of the presence of the acidic region in R (Fig. 1A), which can mediate transcriptional activation when fused to a heterologous DNA binding domain (e.g. GAL4DBD, not shown). p35S::GAL4DBD-R411–610 does not activate transcription from a promoter containing GAL4 binding sites, driving the luciferase reporter (Fig. 2C), while the GAL4DBD-R525–610 does not affect the nuclear localization of R, it appears to interfere with the formation of discrete speckles (Fig. 3A). The biological significance of this pattern for the R regulatory function is not known, but many plant proteins display a similar speckled nuclear localization (31).

Next, we investigated the effect of the 532–560 deletion on the ability of R (together with C1) to activate the anthocyanin pathway. The bombardment of p35S::R532–560-GFP (the same construct as used in Fig. 3A) into BMS cells resulted in the very rare formation of red cells (occasionally one or two) compared with the hundreds of red cells usually observed in similar experiments using p35S::R or p35S::R-GFP (Fig. 3B) in the presence of p35S::C1. The inability of p35S::R532–560-GFP to activate anthocyanins could not be compensated by increasing the amount of plasmid in the bombardments (Fig. 3B), suggesting that the lack of activity of p35S::R532–560-GFP is not due, for example, to the inefficient expression of this protein. As described later, the p35S::R532–560-GFP protein continues to be fully active on some of the other functions of R, indicating that the inability to induce anthocyanin formation in this experiment is not a consequence of, for example, no protein being formed.

We then compared the activity of p35S::R532–560-GFP and p35S::R-GFP for their ability to activate transcription together with p35S::C1 of four anthocyanin biosynthetic genes (A1, A2, Bz1, and Bz2) in transient expression experiments in maize

2B) and assayed in GST pull-down experiments for its ability to interact with bacterially expressed GST-R525–610 (Fig. 2A). As a control, the pull-down experiment was performed in parallel with GST (Fig. 2B). GST-R525–610 (Fig. 2B, lane 2), but not GST (Fig. 2B, lane 3), resulted in the efficient recovery of the ∼79-kDa R protein (Fig. 2B, lane 1) after autoradiography (Fig. 2B, top panel), providing evidence that the dimerization observed in yeast can also happen in vitro.

To establish whether the C-terminal region of R can mediate homodimer formation in maize cells, we took advantage of the presence of the acidic region in R (Fig. 1A), which can mediate transcriptional activation when fused to a heterologous DNA binding domain (e.g. GAL4DBD, not shown). p35S::GAL4DBD-R411–610 does not activate transcription from a promoter containing GAL4 binding sites, driving the luciferase reporter (Fig. 2C), while the GAL4DBD-R525–610 does not affect the nuclear localization of R, it appears to interfere with the formation of discrete speckles (Fig. 3A). The biological significance of this pattern for the R regulatory function is not known, but many plant proteins display a similar speckled nuclear localization (31).

Next, we investigated the effect of the 532–560 deletion on the ability of R (together with C1) to activate the anthocyanin pathway. The bombardment of p35S::R532–560-GFP (the same construct as used in Fig. 3A) into BMS cells resulted in the very rare formation of red cells (occasionally one or two) compared with the hundreds of red cells usually observed in similar experiments using p35S::R or p35S::R-GFP (Fig. 3B) in the presence of p35S::C1. The inability of p35S::R532–560-GFP to activate anthocyanins could not be compensated by increasing the amount of plasmid in the bombardments (Fig. 3B), suggesting that the lack of activity of p35S::R532–560-GFP is not due, for example, to the inefficient expression of this protein. As described later, the p35S::R532–560-GFP protein continues to be fully active on some of the other functions of R, indicating that the inability to induce anthocyanin formation in this experiment is not a consequence of, for example, no protein being formed.

We then compared the activity of p35S::R532–560-GFP and p35S::R-GFP for their ability to activate transcription together with p35S::C1 of four anthocyanin biosynthetic genes (A1, A2, Bz1, and Bz2) in transient expression experiments in maize
BMS cells. Consistent with the decreased anthocyanin accumulation specified by p35S::R/H9004532–560-GFP (Fig. 3B), the activity of p35S::R/H9004532–560-GFP was significantly reduced on the pA1::Luc, pA2::Luc, pBz1::Luc, and pBz2::Luc promoter constructs, with the most dramatic effect on pA1::Luc (75% reduction) (Fig. 3C). Taken together, these results indicate that the dimerization domain is necessary for the transcriptional activity of R.

The Dimerization Domain Is Required for Only a Subset of the R Activities—Previously, we utilized a version of the P1 protein that can interact with R [P1* (17)] to identify two distinct regulatory activities for R (16). The R-enhanced activity is evidenced by a 2–3-fold increase in the activation of pA1::Luc when in the presence of R (16). To investigate whether the dimerization region participates in the R-enhanced activity, we tested P1* with p35S::R-GFP and p35S::R/H9004532–560-GFP for the activation of the pA1::Luc reporter construct (Fig. 4A). As previously observed for p35S::R, p35S::R-GFP increases the expression of the luciferase reporter from the pA1::Luc construct about 2-fold (Fig. 4A, P1*–R-GFP). When the 532–560 region was deleted, R continued to enhance the P1* activity (Fig. 4A, compare P1* with P1*/H9004532–560-GFP). Thus, the dimerization domain is not required for the R-enhanced activity.

We also showed previously that the R-enhanced activity requires the ARE element (16), a cis-regulatory motif conserved
in the promoters of several flavonoid biosynthetic genes (20) and located between the distal low affinity and high affinity P1 binding sites (laPBS and haPBS, respectively). The ARE is necessary for the C1 + R regulation of A1 but not for the activation by P1 (32). Consistent with this, a mutant A1 promoter lacking the laPBS and haPBS (pA1mPBS::Luc) shows no activation by P1. However, pA1mPBS::Luc continues to be modestly activated by C1 + R (33), presumably by the recruitment of the complex to the DNA by R through the ARE, a mechanism similar to which P1* was proposed to activate Bz1 (16).

To determine whether the R dimerization is necessary for the activation of the Bz1 promoter (pBz1::Luc) by P1*, we tested it when together with R-GFP or R^532–560-GFP. Similar to P1, P1* is unable to activate pBz1::Luc in the absence of R (Fig. 4A, P1*).

However, the activation of pBz1::Luc is similar when P1* is co-bombarded with R-GFP or with R^532–560-GFP, indicating that the dimerization region is not required for the regulation of Bz1 by P1*.

Next, we compared the ability of R-GFP and R^Δ532–560-GFP to cooperate with C1 to activate an A1 promoter lacking the laPBS and haPBS (Fig. 4B, pA1mPBS::Luc). Both p35S::R-GFP and p35S::R^Δ532–560-GFP similarly activated the pA1mPBS::Luc reporter when together with p35S::C1 (Fig. 4B, C1). Thus, the absence of R dimerization did not affect the R-enhanced activity, represented by the recruitment of the complex through the ARE. Note, however, that the activation of pA1mPBS::Luc is significantly lower than that of pA1::Luc, highlighting the importance of the laPBS and haPBS in the regulatory activity by C1 + R. Taken together, these results indicate that the dimerization region participates in just a subset of the activities displayed by R, likely to include the formation of a stable complex of C1 with the C1 binding sites (represented by the laPBS and haPBS).

Structural Analysis of the Dimerization Domain of R—The last 86 amino acids of R correspond to a region that is conserved among multiple R-related bHLH transcription factors (Fig. 1B). BLAST analyses using this region of R as a query identifies a number of other R-like proteins in the database but no evidence for the presence of this domain in proteins other than closely related bHLH factors. As a first step in establishing whether this region of R and R-related proteins have structural similarity to any known domain structure, we utilized the FUGUE sequence-structure homology recognition program (34). The top hit obtained (probe divergence values were 0.577 for R and GL3, and their z-scores were 3.74 and 4.84, respectively) corresponded to the ACT domain, a fold present in many proteins that participate in the ligand-mediated allosteric regulation of several biosynthetic enzymes (35). One such enzyme is phosphoglycerate dehydrogenase (PGDH), which catalyzes the first reaction in the L-serine biosynthetic pathway. The pathway is feedback-regulated by L-serine, inhibiting PGDH by binding to the allosteric site, which resides within the ACT domain. This domain also mediates homodimerization of the protein. The interface of the two ACT domains is formed by a 6- or 8-strand β sheet composed of three or four β-strands from each subunit (36, 37) (Fig. 5A). Two L-serine binding sites are formed in the interface of the two ACT domains, each of which includes His-342 and Asn-346 from one subunit and Val-363, Asn-364, and Ile-365 from the other subunit, thus creating 2 binding sites in each homodimer (36).

The predicted secondary structure for the R dimerization domain is in agreement with the secondary structure of the E. coli PGDH ACT domain (Fig. 5B). The three β-strands (blue arrows in Fig. 5B) involved in the homodimerization of the PGDH ACT domain align well with three predicted β strands in R (green arrows in Fig. 5B). Moreover, the region deleted in the R^Δ532–560 mutant (in orange in Fig. 5B) includes one of the β strands involved in the dimerization interface.

Several single amino acid mutations (indicated with a red star in Fig. 1B) to Ala failed to affect the dimerization of the ACT-
like domain (not shown). However, when Ser-560, Gln-562, and Ser-564 were simultaneously replaced by Ala, a significant ($p < 0.01$) reduction in the ability of R$^{411-610}$ to dimerize was observed, evidenced by the reduced $\beta$-galactosidase activity in quantitative yeast two-hybrid assays (Fig. 5C). These residues form part of a $\beta$-strand central for the dimerization of ACT domains (Fig. 5A) and predicted to be conserved in the R ACT-like domain (Fig. 5B), providing additional evidence for the presence of a similar fold in R. Interestingly, quantitative yeast two-hybrid interactions also showed that the 525–610 region of R dimerizes significantly weaker than the 411–610 region (Fig. 5C), a difference not evident from the growth in selective media (compare #1 and #5 in Fig. 2A). Although we cannot rule out the possibility that $R^{525-610}$-GAL4DBD or $R^{525-610}$-GAL4AD is expressed at a significantly lower level than the $R^{411-610}$-GAL4DBD and $R^{411-610}$-GAL4AD proteins, these results may suggest an involvement of the bHLH region of R in participating or stabilizing the dimerization provided by the ACT-like domain.

Identification of ACT Dimerization Domains in Other bHLH Proteins—The structural similarity of the dimerization region of R with the ACT domain prompted us to investigate whether similar domains might be present in other plant bHLH proteins outside of the group formed by R and related factors (Fig. 1B). It should be noted that the dimerization domain was not among the conserved regions identified in a previous analysis of Arabidopsis bHLH proteins (3). We reasoned that ACT-like domains could be present in other bHLH proteins, but due to extensive sequence divergence, we would not be able to detect them by using simple sequence alignments. Therefore, to investigate whether the ACT domain was also present in the C terminus of other bHLH families in Arabidopsis, we used the FUGUE program with representative C-terminal sequences from several of the bHLH subclasses described (3). We found that ACT domains were present in proteins belonging to groups Ia (e.g. At2g46810), IIIa (e.g. At4g21330), IIb (e.g. At5g65640), IIc (e.g. At1g10610), IIIId (e.g. At4g16430), IIle (e.g. At1g32640), IVa (e.g. At2g22770), and Vb (e.g. At1g68810) (Fig. 6A). We verified that all the members of these groups contained the ACT-like domain by BLAST. We found this to be the case with
FIGURE 6. **ACT-like domains are present in several other plant bHLH proteins.**

A, relationships between bHLH proteins containing ACT-like domains based on the phylogenetic reconstruction of the *A. thaliana* bHLH family generated by Toledo-Ortiz et al. (4). The subgroups identified by Heim et al. (3) were mapped onto the tree, and those groups containing the ACT-like domain are indicated in **bold**, with the number of group members that contained the domain, compared with the total number of group members, indicated in **parentheses**. For those groups in which an ACT-like domain was found, a graphical representation of the domain structure of the proteins is indicated including a **dark gray box** for the bHLH domain and a **light gray box** for the ACT-like domain.

B, a yeast two-hybrid experiment showing homodimerization of At2g22770–314 but no heterodimerization with R525–620 in the yeast strain PJ69.4a (51) containing the **HIS3** and **ADE2** genes under the control of the GAL4 binding sites.

C, alignment of plant ACT-like domains present in bHLH proteins, as identified in this study. Based on these sequences, a consensus was derived (Consensus bHLH) and compared with the consensus obtained for ACT domain present in enzymes (Consensus enzymes) previously described (45) following their criteria of big amino acid (FILMVWYKREQ) (**b**), charged amino acid (DEHKR) (**c**), hydrophobic amino acid (ACFILMVWY) (**h**), branched amino acid (ILV) (**i**), polar amino acid (DEHKNQRST) (**p**), and small amino acid (ACSTDNGP) (**s**). To appear in the consensus, an amino acid should be present in at least 8/17 sequences.
the exception of one protein (At3g06120) in Group Ia and four proteins in group III (At3g06120, At2g28160, At5g57150, At4g29930). We verified that these were the only bHLH subclass to contain the ACT-like domain by using the MEME algorithm for the discovery of conserved motifs (38) using a training set that included the ACT-like motifs of the R-like bHLHs as well as those of the proteins mentioned above in subclasses Ia, III, IVa, and Vb. The resulting motif was used to search the entire non-redundant data base, and we were only able to detect the motif in the same bHLH subclasses that we had obtained with the previous method.

The presence of an ACT domain in other bHLH proteins does not necessarily imply that they can mediate similar protein-protein interactions, as we established for the dimerization of R (Fig. 2). Thus, to determine whether the presence of the ACT domain correlates with dimerization, we selected the corresponding region of At2g22770 as a case study. In yeast two-hybrid experiments we tested the interaction of At2g22770233–314 with itself and with R525–610 (Fig. 6B). The robust growth in -Leu - Trp - His - Ade media of a strain expressing At2g22770233–314-GAL4AD and At2g22770233–314-GAL4BD (Fig. 6B, #1) and the absence of autoactivation of At2g22770233–314-GAL4AD (Fig. 6B, #4) provide strong evidence that the corresponding region in At2g22770 mediates homodimer formation. In contrast, this region cannot mediate heterodimer formation with R525–610 (Fig. 6B, #2 and #3), suggesting a clear specificity in the dimerization preference of proteins containing this conserved domain and indicating a correlation between the presence of this domain and the ability to form homodimers in yeast.

**DISCUSSION**

We describe here the presence of a novel protein-protein interaction domain in a group of plant bHLH transcription factors. This region has structural similarity with the ACT domain present in several metabolic enzymes. The corresponding domain in R is central to its ability to cooperate with the R2R3-MYB transcription factor C1 to control anthocyanin pigment formation in maize. These findings provide additional support for the importance of R in mediating multiple protein-protein interactions in the regulation of transcription of flavonoid biosynthetic genes.

Previous studies identified R alleles derived from the excision of the Ds transposon from the r-m9 allele, inserted immediately upstream of the sequence encoding the identified dimerization domain. One of these derivative alleles, R-v24, carrying a 7-bp insertion, resulted in a frameshift mutation with reduced aleurone anthocyanin pigmentation (39). RNA expression analyses on the mutant v24 showed a 50–70% reduction in the mRNA steady-state levels of C2 and A1, two anthocyanin biosynthetic genes (39). Because the v24 mutation affected the R C-terminal nuclear localization signal (Fig. 1A), the results were interpreted to indicate that the reduced nuclear localization of R in the mutant was at least in part responsible for the decreased anthocyanin accumulation (39). Strikingly, however, we found a similar reduction in the activation of the A1 promoter when we delete the 532–560 region of R (Fig. 3C). R532–560 displays similar nuclear localization as the wild type R when fused to GFP (Fig. 3A). Taken together, we interpret these results to indicate that the main effect of the R-v24 mutation is the absence of the dimerization domain and is not a problem in nuclear localization. The significantly reduced accumulation of anthocyanin pigments in the R-v24 allele is consistent with the decreased ability of R532–560 to induce anthocyanin formation in maize cells (Fig. 3B), providing *in vivo* evidence for the significance of the dimerization domain for R function.

Our studies show that the R525–610 region can mediate homodimerization in yeast (Fig. 2A), *in vitro* (Fig. 2B), and in maize cells (Fig. 2C). However, we are cautious to propose that the main role of this region of R is to mediate homodimer formation *in vivo* or that the observed reduction on the R regulatory activity (Fig. 3, B and C) of deletions of this region is due to the absence of R homodimerization. The C-terminal regions of GL3 and EGL3 (including the bHLH) domains have been shown to heterodimerize (27), and we have established by yeast two-hybrid that the corresponding regions of R and B can interact as well (not shown). However, in maize R functions in the absence of B and, similar to GL3 and EGL3, R and B have very similar functions (40); thus, it is unlikely that the role of this region is solely to mediate R/B heterodimer formation. We cannot yet rule out the possibility that this region of R mediates the formation of heterodimers with other as-yet unidentified proteins. Yeast two-hybrid screens are currently under way to try to determine the identity of such partners if they exist.

The finding that the dimerization region of R cannot interact with the corresponding region of At2g22770, which also forms homodimers (Fig. 6B), suggests that the interactions mediated by this novel domain are specific. At2g22770 corresponds to the recently described NAI1 gene, which controls the formation of the ER bodies (41), endoplasmic reticulum-derived structures involved in the transport of proteins to the vacuole (42). The possibility that the regulatory activity of NAI1 requires the identified dimerization domain has not been previously explored.

The regulation of A1 transcription by C1 and R requires at least three *cis*-regulatory elements, the haPBS, the laPBS, and the ARE (16). In the absence of C1 DNA binding, either because of a mutation in one of the C1 DNA-recognition helices or because of mutations in the haPBS and laPBS, C1 can be recruited in an R-dependent fashion to the A1 promoter by the ARE (16, 33). Similarly, R can recruit the P1* protein to the Bz1 promoter, which is normally not regulated by P1 (16, 17). Our results indicate that the deletion of the dimerization domain of R does not affect the ARE-mediated tethering of the C1/P1* proteins to the corresponding promoters (Fig. 4). Rather, it is apparent that the dimerization of R is necessary for the binding of C1 to the haPBS, to the laPBS, or to both. The finding that the deletion of the dimerization domain of R is important for just one aspect of the R regulatory function is of importance from at least two perspectives. First, it provides strong confirmation that the deletion of the 532–560 region does not affect the overall folding or stability of R, since other R-dependent activities remain unharmed. Second, it may help to understand how a stable C1-DNA complex is formed given the very low affinity of C1 for DNA (33).
It is common for bHLH transcription factors to have other protein-protein interaction domains, and domain shuffling was proposed as one mechanism to explain the domain multiplicity in members of this large family of regulatory proteins (8). The presence of an ACT-like domain in the dimerization region of one of the best-described groups of plant bHLH proteins provides a new addition to the diversity of protein-protein interaction domains associated with this class of transcription factors. ACT domains have been typically associated with metabolic enzymes, including PGDH (43) and phenylalanine hydroxylase (44), where they often play a regulatory role by providing allosteric regulation through the binding of specific small molecule ligands (35). In addition to PGDH, ACT domain-mediated protein-protein interactions have been described in several other factors (35). The high conservation of several residues among the ACT-like domain present in plant bHLH proteins (Fig. 6A) and their position outside the dimerization interface in the E. coli PGDH ACT structure (Fig. 5A) suggest that they could be involved in other functions such as, for example, ligand recognition. It should be noted, however, that no plant bHLH proteins have been yet shown to interact with a small molecule ligand. An alignment of 17 ACT-like domains present in plant bHLH factors permitted us to deduce a loose consensus (Fig. 6C) that may facilitate the identification of a similar fold in other proteins. Interestingly, however, when this consensus is compared with that of several ACT-domain containing enzymes (45), significant differences are evident (Fig. 6C).

The presence of an ACT-like domain in various distinct groups of bHLH proteins is intriguing and suggests an ancient function in the plant bHLH family. The occurrence of an ACT domain in a regulatory protein has, however, precedents in the glycine binding domain of the E. coli transcriptional accessory protein GcvR, which participates in the regulation of the gcvTHP operon (46). The presence of ACT domains in eukaryotic transcription factors provides to our knowledge a novel finding. It will be of interest to establish whether they are limited to the plant bHLH family or whether the presence of ACT-like domains is a more general feature of other transcription factor families.

Taken together, the results presented here demonstrate the presence of a novel ACT-like dimerization domain in R and other plant bHLH proteins. This domain plays an important function in the ability of R to cooperate with the R2R3-MYB factor for the activation of maize pigmnet biosynthesis. These findings highlight the growing need to combine protein fold recognition with functional analyses in discovering novel functional protein domains.

REFERENCES
ACT Dimerization Domains in Plant bHLH Factors

Antje Feller, J. Marcela Hernandez and Erich Grotewold

Basic-helix-loop-helix Transcription Factors

An ACT-like Domain Participates in the Dimerization of Several Plant

doi: 10.1074/jbc.M603262200 originally published online July 25, 2006

Access the most updated version of this article at doi: 10.1074/jbc.M603262200

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 53 references, 29 of which can be accessed free at
http://www.jbc.org/content/281/39/28964.full.html#ref-list-1