Planta (2008) 227:707–715
DOI 10.1007/s00425-007-0676-y

RAPID COMMUNICATION

The interaction domains of the plant Myc-like bHLH transcription factors can regulate the transactivation strength

Sitakanta Pattanaik · Claire H. Xie · Ling Yuan

Received: 26 October 2007 / Accepted: 1 December 2007 / Published online: 13 December 2007
© Springer-Verlag 2007

Abstract The N-terminal region of the plant Myc-like basic helix–loop–helix transcription factors (bHLH TFs) contains two domains. Approximately, 190 amino acids at the N-terminus comprise an interaction domain, a.k.a. Myb-interacting-region (MIR) for its primary function of interacting with Myb-like TFs. Following, the interaction domain is an activation (or acidic) domain responsible for transactivation. We have previously discovered that a lysine to methionine substitution (K157M) in the interaction domain of Myc-RP of *Perilla frutescens* leads to a 50-fold increase in transactivation activity. The result suggests that mutations in the interaction domain affect transactivation. The highly conserved nature of this lysine residue in many Myc-like bHLH TFs prompted us to explore the functional importance of this residue within the TF family and the influence of the interaction domain on the activation domain in transactivation. We found that the replacement of the equivalent lysine with methionine significantly affects the transactivation activities of two other Myc-RP homologues, Delila from snapdragon and Lc from maize. In addition to methionine, substitution with several other amino acids at this position has positive effects on transcriptional activity. A neighboring conserved alanine residue (A159 in Myc-RP, A161 in Delila and A172 in Lc) also affects transactivation. Substitution of this alanine residue to an aspartic acid abolished transactivation of both Myc-RP and Delila and severely reduced transactivation of Lc. Ectopic expression of a Myc-RP K157M mutant in transgenic tobacco resulted in increased anthocyanin accumulation compared to plants expressing the wild-type gene. Our study reveals the potential cooperation between functional domains of the bHLH TFs.

Keywords Basic helix–loop–helix transcription factors · Transcription factor domains · Directed evolution · Anthocyanin

Abbreviations

bHLH TF	Basic helix–loop–helix transcription factor
β-Gal	β-galactosidase

Introduction

Anthocyanins are a class of structurally conspicuous flavonoids that exhibit a wide range of colors. In plants they serve as visual signals that attract insects and animals for pollination and seed dispersal, and play a role in photoprotection in autumn foliage and rapidly growing shoots of tropical trees. The antioxidant activity of anthocyanins gives cause for a variety of medical usage including cancer prevention, anti-inflammation and anti-arteosclerosis (Mol et al. 1998; Saito and Yamazaki 2002). The anthocyanin biosynthetic pathway is one of the most extensively studied pathways of plant secondary metabolites. The majority of enzymes involved in this pathway, and the genes encoding them, have been isolated and characterized from a number of plant species including maize, *Arabidopsis*, petunia and snapdragon. In addition, many regulatory genes that directly control the transcription of the structural genes in the anthocyanin biosynthetic pathway have been identified from a variety of plants species (Koes et al. 2005; Mol et al. 1998). In maize three families of transcription factors

S. Pattanaik · C. H. Xie · L. Yuan (✉)
Department of Plant and Soil Sciences and the Kentucky Tobacco Research and Development Center, Lexington, Cooper and University Drives, KY 40546, USA
e-mail: lyuan3@uky.edu

(TFs), the bHLH, Myb and WD40, control the transcription of the structural genes in the anthocyanin biosynthetic pathway. The members of the bHLH gene family (*R, Lc, Sn, B*) and the Myb gene family (*C1* and *Pl*) encode basic helix–loop–helix transcription factors (bHLH TFs) and R2R3 MYB domain proteins, respectively. The third family of TFs associated with the bHLH and MYB proteins is the WD40 repeat proteins (pac1) which are characterized by the presence of a 40-residue core region containing random repeats of a glycine-histidine (GH) dipeptide and a tryptophan-aspartate (WD) dipeptide (Koes et al. 2005; Quattrocchio et al. 1998). Similar regulatory factors controlling anthocyanin biosynthesis have also been isolated from several dicot plants including petunia, snapdragon and perilla (*Perilla frutescens*) (de Vetten et al. 1997; Gong et al. 1999a, b; Martin et al. 1991; Quattrocchio et al. 1993, 1998; Sompornpailin et al. 2002).

The bHLH proteins belong to a large family of transcriptional regulators present in both animals and plants. This family of TFs is characterized by the signature bHLH domain comprising approximately 60 amino acids. The basic region of the domain binds to DNA containing the canonical E-box (CANNTG) sequence and the HLH region is involved in homo- and heterodimerization (Bailey et al. 2003; Heim et al. 2003; Li et al. 2006; Toledo-Ortiz et al. 2003). The plant Myc-like bHLH TFs contain an N-terminal interaction domain (Fig. 1). This domain, a.k.a. Myb-interacting region (MIR), has been shown to interact with an R2R3-MYB domain protein and affect transcription of anthocyanin biosynthetic pathway genes (Goff et al. 1992, 1990; Grotewold et al. 2000; Hernandez et al. 2004). The function of this N-terminal domain (herein referred to as interaction domain) is not fully understood. When deleted, as demonstrated in Lc and *Arabidopsis* GL3, the bHLH TFs are unable to interact with their partner Myb proteins (Grotewold et al. 2000; Payne et al. 2000). In addition to Myb interaction, however, the interaction domain may be involved in other functions. For example, an *Arabidopsis* bHLH TF ATR2 mutant, with a single mutation in the interaction domain, activates tryptophan regulation genes; however, the ATR2 mutant functions independently of ATR1, a Myb TF that regulates the tryptophan pathway. These observations suggest the possible involvement of the interaction domain in other regulatory mechanisms (Smolen et al. 2002).

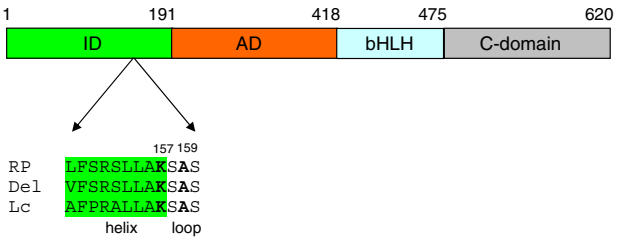

Fig. 1 Schematic representation of functional domains of a R-like bHLH protein. The protein contains an interaction domain (ID), activation domain (AD), bHLH domain and a C-terminal domain (C-domain). The amino acid numbering is based on the Myc-RP sequence. Partial sequence alignment of the interaction domains of Myc-RP (RP), Delila (Del) and Lc are highlighted to show the helix–loop region. The conserved lysine (K157) and alanine (A159) are in *bold*

In between the interaction domain and the bHLH domain is the activation domain which is rich in acidic amino acids (Fig. 1). It is generally believed that the activation domain forms a protein-binding surface that interacts with the RNA polymerase II machinery and subsequently initiates transcription. Upon protein–protein interaction, the activation domain undergoes structural changes from unstructured to an α-helical conformation (Roberts 2000). For plant Myc-like bHLH TFs, it remains unclear whether domains in addition to the activation-domain interface also contribute to the overall binding pocket for RNA polymerase II. Another hallmark of Myc-like bHLH TFs is the presence of multiple protein dimer interfaces. One such interface is an ACT-like domain located at the C-terminus. It has been demonstrated that the ACT domain of Lc mediates the formation of homodimers in vitro and in vivo (Feller et al. 2006).

Although the bHLH TFs form the largest plant TF family, little is known about the structure/function relationship of these proteins. We have applied laboratory directed evolution technologies (Chatterjee and Yuan 2006; Stemmer 1994; Yuan et al. 2005) to modify the transcriptional activity and promoter specificity of Myc-RP of Perilla and Delila of snapdragon (Pattanaik et al. 2006). One of the best characterized Myc-RP variants, M2-1, shows a 70-fold increase of transcriptional activity. Combinatorial site-directed mutagenesis of M2-1 revealed that a single lysine to methionine change (K157M) in the interaction domain is responsible for 80% of the improvement observed in this mutant (Pattanaik et al. 2006). This lysine is well conserved among the R-like bHLH TFs and resides in a predicted helix structure. Mutations in this helix and the loop structure that immediately follows both positively and negatively affect transactivation strength, indicating a role for these structures in transactivation. In addition to the K157M mutation, another mutation in the helix (S151P) also positively affects transactivation. In contrast, mutation of a conserved alanine in the loop structure of Delila to aspartic acid (A161 in Delila and A159 in Myc-RP; Fig. 1) abolishes transactivation (Gong et al. 2000). These observations led us to consider the possibility that the interaction domain folds close to the binding surface of the activation domain and small structural changes can significantly influence the recruitment of the RNA polymerase II complex by the activation domain.

In this report we show that the lysine to methionine change, equivalent to K157M of Myc-RP, in the interaction domains of two other bHLH TFs (Delila and Lc), also has significant influence on transactivation activity. In addition to methionine, substitutions of several other amino acids at this position have positive effects on transcriptional activity. Similar to the effect of A161D mutation in Delila, replacement of the equivalent alanine in Myc-RP and Lc with aspartic acid severely decreases transcriptional activity. These results support a general model in which the interaction domains of the bHLH TFs cooperate with the activation domain in transactivation. An evolved bHLH TF with improved transcriptional activity can result in phenotypic improvement in planta. Ectopic expression of the evolved Myc-RP (M2-1) in transgenic tobacco has resulted in increased anthocyanin accumulation in flowers compared to plants expressing the wild-type gene.

Materials and methods

Site-directed mutagenesis and saturation mutagenesis

The full-length cDNA of *Myc-RP*, *Delila* and *Lc* cloned in the yeast effector plasmid *pAS2-1* (Clontech, CA) were used as templates in PCR for site-directed and saturation mutagenesis. Site-directed mutagenesis was carried out using the method described previously (Zheng et al. 2004). Complementary primers with corresponding mutations were synthesized and used in PCR amplification of the *pAS2-1* containing the *Myc-RP, Delila or Lc* cDNA. The amplified plasmids were evaluated by agarose gel electrophoresis and gel purified using the QIAquick gel extraction kit (Qiagen, CA). Parental PCR template plasmid was removed from the mutant library by digestion with *Dpn*I prior to transformation into *E. coli* TB1 cells (New England Biolabs, MA). Sequencing to verify the mutation was performed using a Beckman Coulter Sequencer CEQ-8000 (Beckman, CA). In saturation mutagenesis, the targeted codon in the complementary primers was randomized, and the resulting amino acid residues were identified by DNA sequencing.

Yeast-one-hybrid experiment and β-galactosidase assay

The plasmids containing the mutants were transformed into yeast strain Y187 (*Saccharomyces cerevisiae* Y187:ura3-52, trp1-901, Ura3::GALUAS-Gal1TATA-LacZ) for transactivation of the *Gal1* promoter. Yeast transformations were conducted using the PEG/LiAc method as described in the Yeast Protocol Handbook (Clonetech, CA). The Y187 transformants were selected on synthetic dropout (SD) medium lacking tryptophan.

β-galactosidase activity was assayed by both colony-lift filter assay and liquid culture assay according to procedures described in the Yeast Protocols Handbook (Clonetech, CA). Substrates used for the filter and liquid methods were 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-GAL) and *o*-nitrophenyl-β-D-galactopyranoside (ONPG), respectively. One unit of β-galactosidase activity is defined as the hydrolysis of 1 μmol of ONPG to *o*-nitrophenol and D-galactose per min per cell.

Construction of plant expression vectors for tobacco transformation and analysis of transgenic plants

The wild-type and variant Myc-RP (M2-1) genes were cloned into the binary vector pCAMBIA 2300 under the control of the mirabilis mosaic virus (MMV) promoter and rbcs terminator (Dey and Maiti 1999). The constructs were introduced into *Agrobacterium tumefaciens* strain C58C1:pGV3850 by freeze–thaw. Tobacco plants (*Nicotiana tabacum* cv Samsun NN) were transformed with the *Agrobacterium* as described previously (Pattanaik et al. 2004). Regenerated, kanamycin-resistant tobacco plants were grown under greenhouse conditions. Independent lines with single gene insertion were identified by analyzing the segregation ratio (Kan^R:Kan^S) of the kanamycin (*Kan*) marker gene.

Total cellular RNA from transgenic tobacco seedlings and young flower buds expressing the wild-type and variant Myc-RP (M2-1) construct was isolated using the RNeasy plant mini kit (Qiagen, Chatsworth, USA). Total RNA (2.5 μg) was treated with RNase-free DNase (Sigma, USA) and reverse transcribed in a total volume of 20 μl using Superscript II Reverse Transcriptase (Invitrogen, USA). For the no-reverse-transcriptase control, an individual reaction was performed in parallel without addition of reverse transcriptase. One μl of the RT products was used in the subsequent PCR reaction with appropriately designed forward and reverse primers to amplify the full-length wild-type or the variant Myc-RP transcript. The α-tubulin house-keeping gene was used as an internal control in RT reactions. The PCR reaction was performed in a total volume of 25 μl using GoTaq green master mix (Promega, USA) for 30 cycles (94°C for 30 s, 55°C for 30 s, 72°C for 2 min). As a negative control, each primer pair was tested against DNase-treated RNA to confirm cDNA dependence of amplification. PCR products were analyzed on an ethidium bromide-stained agarose gel.

Determination of anthocyanin content

Anthocyanin was extracted from fresh flowers collected from transgenic tobacco plants expressing wild-type Myc-RP or *M2-1*, or the empty vector in 1 ml of 1% HCl in

methanol in the dark overnight at 4°C. The extracts were diluted in 1% HCl in methanol and used for determination of absorbance at 530 nm. Anthocyanin content was expressed as absorbance units at 530 nm per gram tissue fresh weight as described previously (Gong et al. 1999a). At least three flowers from a single plant were used to obtain the mean anthocyanin contents.

Results

Mutations of a conserved lysine in the interaction domain affect transactivation

We have previously demonstrated that the transactivation activity of Myc-RP can be increased by recursive directed evolution (Pattanaik et al. 2006). Among the improved Myc-RP variants, M2-1, which has up to 70-fold increased transactivation activity, is particularly interesting. M2-1 contains four mutations, three in the activation domain and one in the interaction domain. Combinatorial site-directed mutagenesis revealed that the single mutation in the interaction domain (K157M), alone, increases transactivation by 50-fold (Pattanaik et al. 2006). Secondary protein structure prediction localizes K157 of Myc-RP to a α-helix in the interaction domain. This predicted α-helix is conserved among many Myc-RP homologues including Delila and Lc. In the present work conversion of the equivalent lysine of K157 to methionine resulted in two-fold increase of activity for Delila in a yeast one-hybrid assay (Fig. 2). However, the same mutation in Lc resulted in two-fold decrease of activity (Fig. 2). To determine whether methionine is the only amino acid substitution to K157 of Myc-RP that leads to increased transactivation or additional substitutions at the equivalent lysine of Lc (K170) also results in increased transactivation, we conducted saturation mutagenesis at this conserved residue in both Myc-RP and Lc. When K157 of Myc-RP was mutated to 19 other amino acids, 10 amino acid substitutions resulted in increased transactivation, of which five were significantly better than K157M (K157Q, K157H, K157F, K157W and K157Y; Fig. 3a). On the other hand, three amino acid substitutions (K170F, K170P and K170V) in Lc resulted in 4–5-fold increases of transactivation activity (Fig. 3b). These results indicate that mutations to the lysine residue in all three homologues can alter transactivation strength. Without three-dimensional structural data, the specific impact of the substitutions to the helix structure is difficult to explain. However, the side-chain properties of the substituted amino acids do not appear to be significant factors influencing the observed positive effects on transactivation. In contrast, substitutions of aromatic and hydrophobic residues in Lc appear to more positively affect transactivation.

A conserved alanine in the interaction domain can critically affect transactivation

Two amino acids away from K157 of Myc-RP is an alanine (A159, A161 and A172 in Myc-RP, Delila and Lc, respectively) that is conserved in the Myc-like bHLH TFs. Gong et al. (2000) demonstrated that replacement of A161 of Delila with 19 different amino acids results in significantly decreased or loss of transactivation activity in a yeast one-hybrid assay. To test whether the equivalent alanine residue in Myc-RP or Lc has similar effect on transcriptional activation, we replaced it with an aspartic acid by site-directed mutagenesis. The A159D mutation completely abolished the transactivation activity of Myc-RP in yeast one-hybrid assays (Fig. 4). The A172D mutant of Lc exhibited a two-fold decrease in transactivation activity compared to wild-type (Fig. 4). The saturation mutagenesis on A161 of Delila indicated that the majority of mutations lead to severe reductions rather than complete inactivation (Gong et al. 2000). We thus performed saturation mutagenesis on A172 of Lc. We observed significant reduction in transactivation activity with all 19 amino acid substitutions (data not shown). The severe reduction in activity resulting from replacement of the alanine with other amino acids in all

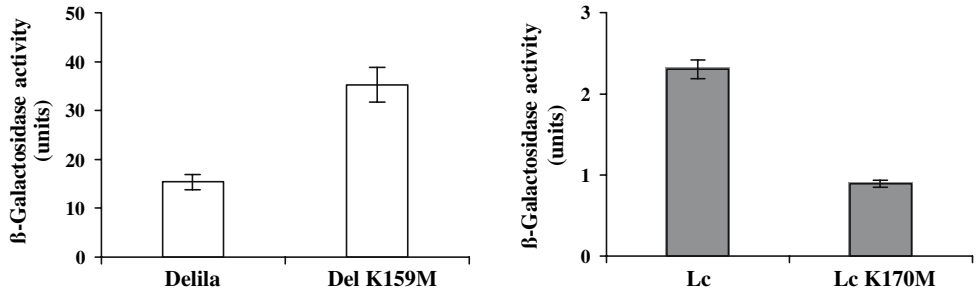

Fig. 2 Transactivation activities of the lysine to methionine mutants of Delila and Lc. Mutations were generated by site-directed mutagenesis, and mutants were assayed in yeast using the pLacZ-GAL1 reporter. The reporter only control gave no detectable *β*-Gal activity. One unit of enzyme activity is defined as the hydrolysis of 1 µmol of ONPG to *o*-nitrophenol and D-galactose per min per cell. Each value is the mean of at least three independent assays

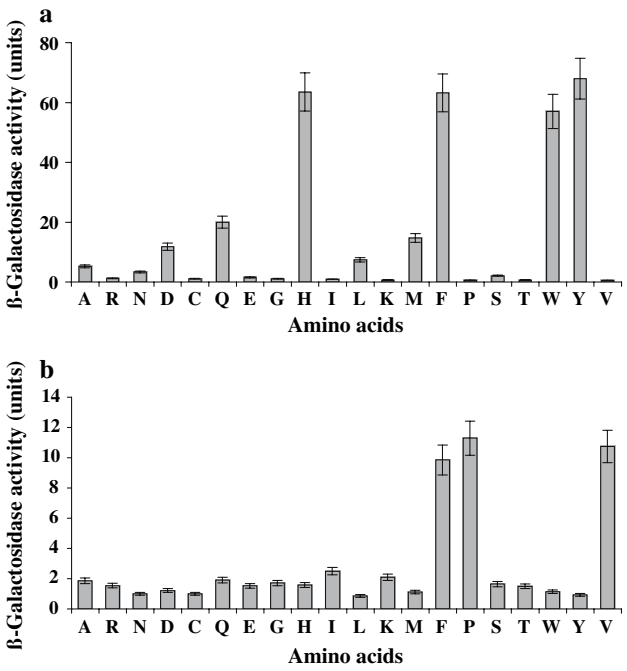

Fig. 3 Site-saturation mutagenesis of K157 of Myc-RP **a** and K170 of Lc **b** The lysine residue of each TF was replaced by all 19 possible amino acid substitutions. Each amino acid is labeled by the one-letter code. The transactivation activities of wild-type and mutants were assayed in yeast using the pLacZ-GAL1 reporter. The reporter only control gave no detectable *β*-Gal activity. One unit of enzyme activity is defined as the hydrolysis of 1 µmol of ONPG to *o*-nitrophenol and D-galactose per min per cell. Each value is the mean of at least three independent assays

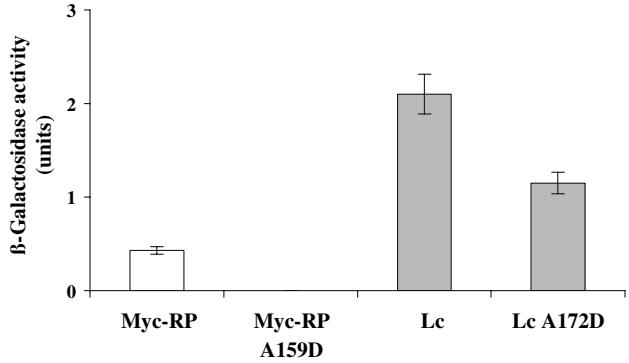

Fig. 4 Transactivation activities of the alanine to aspartic acid mutants of Myc-RP and Lc . The mutations were generated by site-directed mutagenesis, and the mutants were assayed in yeast using the pLacZ-GAL1 reporter. The reporter only control gave no detectable *β*-Gal activity. One unit of enzyme activity is defined as the hydrolysis of 1 µmol of ONPG to *o*-nitrophenol and D-galactose per min per cell. Each value is the mean of at least three independent assays

three proteins underscores the critical role of the alanine in maintaining optimal transactivation for this class of bHLH TFs. Secondary structure prediction localizes this alanine residue to a loop adjacent to the α-helix where K157 of Myc-RP resides (Fig. 1). The location of this helix–loop structure towards the end of the interaction domain and near the activation domain suggests a possibility that it acts as a hinge controlling the proximity between the two domains. As such, mutation of the alanine increases the rigidness of the loop. In severe cases, e.g. A159D of Myc-RP or A161D of Delila, alteration to the loop leads to a masking of the activation domain by a portion of the interaction domain that prevents binding of general transcriptional machinery. Deletion of the N-terminal 136 amino acids from the Delila A161D mutant has been observed to recover approximately 10% of the transcriptional activity (Gong et al. 2000), providing supporting evidence for this theory.

Ectopic expression of Myc-RP variant M2-1 in tobacco increases anthocyanin accumulation in flowers

One of our long-term goals is the engineering of "master" TFs capable of regulating complete plant metabolic pathways to generate desired phenotypes. We have created a number of TF variants with modified functions through the application of directed evolution technologies. These TF mutants are also advantageous for structure/function studies because, in addition to gained novel functions, many of the proteins contain only a few mutations. During directed evolution of Myc-RP, the TF mutants were first selected using yeast one-hybrid transactivation assays. Selected variants were then characterized further using a tobacco protoplast transient expression assay. The variant M2-1, in which the K157M mutation is responsible for 80% of its improvement in transactivation activity, shows marked improvement in the protoplast transactivation assay (Pattanaik et al. 2006). M2-1 was subsequently tested to determine whether the improvement in activity translated phenotypically in transgenic plants. Stable transgenic tobacco plants expressing *M2-1*, wild-type *Myc-RP* and empty-vector were generated. Twelve independent transgenic lines were developed for each construct. Seedlings showing a segregation ratio of 3:1 for the kanamycin marker gene were used for further analysis. RT-PCR analysis was conducted using total RNA isolated from seedlings and flower buds to amplify the full-length Myc-RP transcript. A ~1.9 kb transcript was detected in plants expressing the wild-type or the mutant Myc-RP both in the vegetative tissues as well as in flower buds. The same transcript was not detected in plants expressing the empty vector (Fig. 5a). As an internal control, a house-keeping gene, *α-tubulin* was used in the RT-PCR experiments and a corresponding 1.2 kb band was detected in all transgenic lines tested (Fig. 5a). Ectopic expression of the wild-type *Myc-RP* has been shown to elevate the red pigments in tobacco flowers (Gong et al. 1999a). Transgenic tobacco lines expressing *M2-1* pro-

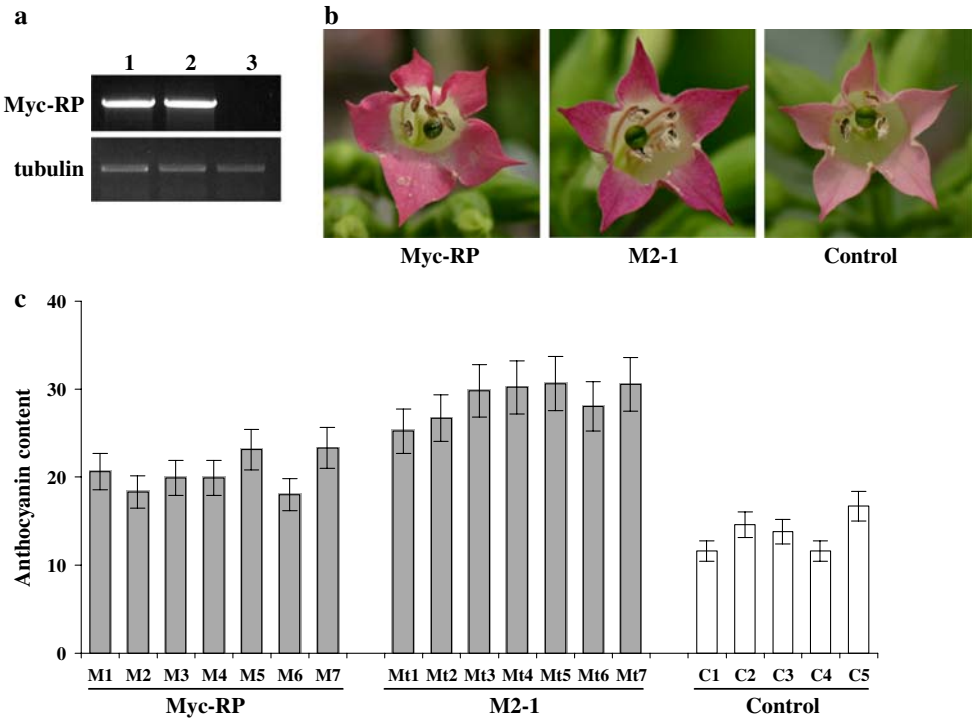

Fig. 5 Enhanced anthocyanin production in flowers of transgenic tobacco expressing the wild-type *Myc-RP* and variant *M2-1*. **a** RT-PCR analysis of total RNA isolated from plants expressing *Myc-RP* (*lane 1*), *M2-1* (*lane 2*) and empty vector (*lane 3*). **b** Flowers from transgenic tobacco plants expressing *Myc-RP*, *M2-1* and empty vector. **c** Anthocyanin content in flowers of different transgenic lines. Anthocyanin was extracted from fresh flowers with 1% HCl in methanol in the dark. Each bar represents the mean anthocyanin concentration of three flowers from an independent, single-gene insertion line. The anthocyanin concentration is expressed as the absorbance units at 530 nm per gram fresh tissue weight

duced flowers with more intense red color than those expressing wild-type *Myc-RP* or the empty vector (Fig. 5b). Increased accumulation of anthocyanin pigments in the petals and anther filaments of flowers was even more clearly visible in *M2-1*-expressing lines compared to those from the wild-type *Myc-RP* or the empty vector lines. Quantitatively, the mean anthocyanin content of the empty-vector control plants was approximately 15 units, whereas, the mean anthocyanin content of *Myc-RP* and *M2-1* plants was determined to be approximately 20 and 25 units at absorbance at 530 nm per gram fresh weight, respectively (Fig. 5c). Flowers were analyzed from independent, single-gene insertion lines; therefore, these results show that twice the amount of anthocyanin is produced in tobacco flowers through ectopic expression of *M2-1* than the wild-type *Myc-RP*. Ultimately, the feasibility of using engineered TFs to control plant metabolic pathways will require verification in transgenic plants. Our data indicate that in stable transgenic tobacco plants, M2-1, a Myc-RP mutant engineered to increase transcriptional activity, generates substantially more anthocyanin. Previously, we have demonstrated that the improved transactivation properties of M2-1 are not a result of increased protein production (Pattanaik et al. 2006). Increased *M2-1* mRNA levels compared to wild-type *Myc-RP* were not observed in transgenic tobacco lines. Therefore, it is unlikely that the phenotypic improvement observed in the transgenic plants is a consequence of increased accumulation of the engineered TF.

Discussion

The bHLH TFs are multidomain proteins that arise evolutionarily through protein domain shuffling (Morgenstern and Atchley 1999). The modular structures of TFs, as implied by the domain names such as activation domain or DNA-binding domain, suggest that these proteins are composed of independent functional domains. The multidomain characteristic of TFs has been the basis for widely used molecular biological techniques. One such technique is the yeast one-hybrid system in which the GAL4 DNA-binding domain is fused with other potential activation domains and the fusion protein is assayed for transcriptional activity. While many domain-swapping experiments have been applied successfully to generate novel functions for TFs, little is known regarding the functional relationships between individual domains. Nevertheless, evidence supporting domain cooperation within the bHLH TFs exists. For instance, maximal DNA-binding strength of the bHLH domain of the E47 TF requires the cooperation of other regions of the protein (Lu and Sloan 2002). In another example, bHLH/DNA binding complexes mediated by the heterodimers between the aryl hydrocarbon receptor nuclear translocator and dioxin receptor or hypoxia inducible factor are stabilized by the PAS domain (Chapman-Smith et al. 2004). Several lines of evidence indicate that the N-terminal interaction domains of the plant Myc-like bHLH TFs are critical for their transcriptional activities (Goff et al. 1992; Gong et al. 1999a; Pattanaik et al. 2006;

Payne et al. 2000; Smolen et al. 2002). Goff et al. (1992) have demonstrated that the bHLH domain and the majority of the C-terminus of the maize TF B can be deleted with only partial loss of transactivation activity in both yeast and plant cells. In contrast, deletions within the interaction domain of B severely reduce transactivation. Similar observations were also made for Delila (Pattanaik et al. unpublished data) and the *Arabidopsis* GL3 (Payne et al. 2000).

Our present study provides additional information on cooperation between domains in plant Myc-like bHLH TFs. We have demonstrated that single amino acid substitutions in the helix–loop structure of the interaction domain results in significant changes in transcriptional activity, i.e. the 50-fold increase of activity by the K157M mutation in Myc-RP and complete inactivation by A159D and A161D mutations in Myc-RP and Delila, respectively. These observations have led us to propose that the interaction domain regulates the activation domain geometrically via control of a helix–loop hinge. Our working model to explain the influence of the interaction domain on the recruitment of the transcriptional initiation complex (TIC) by the activation domain is summarized in Fig. 6. TFs generally function by increasing local protein concentrations of limiting factors at target promoters (Fig. 6a). Specifically the activation domains of TFs directly interact with the TIC which in turn initiates transcription (Fig. 6b). In this model, structural changes in the interaction domain that arise as a consequence of mutation K157M (and other substitutions) optimally stabilize the TIC-activation domain complex and result in a significant increase in transactivation strength (Fig. 6c). The A159D mutation causes the interaction domain to mask activation domain and thereby prevents interaction with the TIC (Fig. 6d). The K157M/A159D double mutant cannot compensate for the loss of activity (data not shown). This model suggests that a portion of the ID is structurally close to the binding surface of the activation domain or perhaps even participates as part of the overall binding pocket. This model, however, does not intend to suggest that the function of the activation domain completely depends on the interaction domain. The interaction domain-less Myc-RP or Delila is capable of initiating transactivation in a yeast one-hybrid system (Pattanaik, et al. unpublished data; Gong et al. 1999a), suggesting the presence of the basal activity of the activation domain. The primary function of the interaction domain of the Myc-like bHLH TFs is to mediate protein-protein interaction with other co-regulators in the flavonoid biosynthetic pathway. The expression of an interaction domain-truncated *Myc-RP* in transgenic plants does not produce an anthocyanin phenotype (Gong et al. 1999a). Nonetheless, here we present evidence that the interaction domains of Myc-RP, Delila and Lc can influence the transactivation strength of the TFs. Because the targeted amino acids in this study are highly conserved within this family

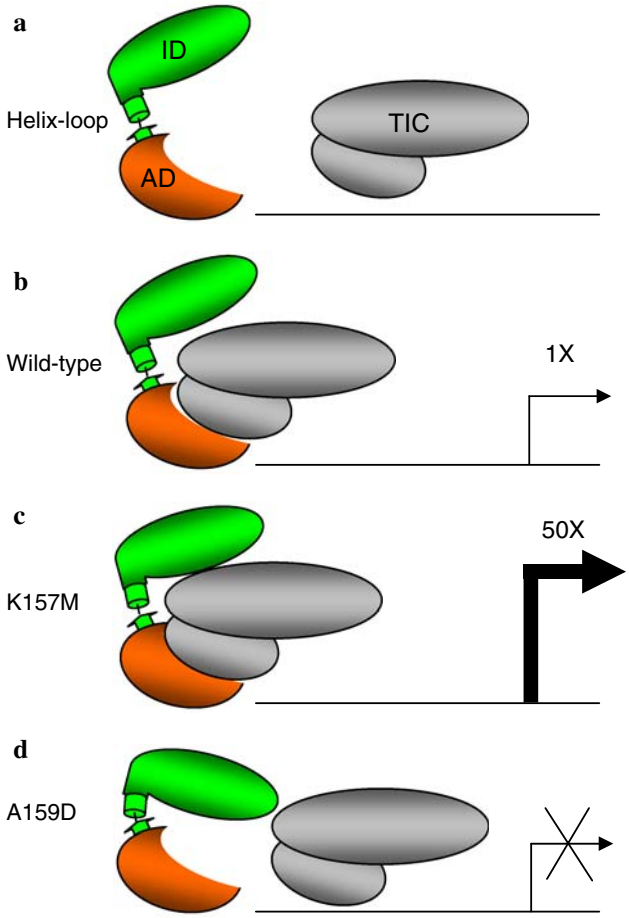

Fig. 6 Proposed model for effects of interaction domain mutations on transactivation; **a** the interaction domain (ID; green) and activation domain (AD; red) of wild-type bHLH TF and the transcriptional initiation complex (TIC). The amino acid sequence forming the helix–loop structure that contains K157 and A159 of Myc-RP is shown. For clarity the rest of the bHLH protein is not shown. **b** initiation of transcription by the bHLH TF upon binding of TIC (*arrow* indicates transcriptional strength); **c** K157M substitution resulting in optimized ID conformation that further stabilizes the AD-TIC complex. **d** A159D substitution resulting in ID masking AD

of Myc-like bHLH TFs; it is likely that similar effects of the amino acid substitutions are also present in other family members.

Given that the interaction domain is known for the Myb-interaction; it is reasonable to expect that the mutational effects on transactivation are consequences of altered interactions between the interaction domains and a Myb-TF. While this possibility cannot be ruled out, two lines of evidence suggest this is unlikely. First, the mutants were initially identified and characterized by yeast one-hybrid assays in which the yeast GAL4 DNA-binding domain is fused with the bHLH TFs, and the GAL4 domain binds to the GAL response elements upstream of the promoter in the reporter construct. Such a system does not require the bHLH TF to interact with a yeast Myb TF in transactiva-

tion; thus the increase or decrease of activity by different interaction domain mutants are not likely to be consequences of altered interaction with an endogenous yeast Myb protein. Further supporting evidence comes from the observations that interaction domain-truncated Myc-RP, Delila and GL3 activate reporters in yeast one-hybrid assays, suggesting that the transactivation does not depend on the Myb-interaction by the bHLH TFs (Pattanaik et al. unpublished results; Gong et al. 1999a; Payne et al. 2000). Therefore, while the interaction domain mutants described here may have altered strengths for Myb-interaction (especially in plant cells as discussed below), the dramatic changes of transactivation activity by the mutations in the helix–loop structure are more likely due to the intra-molecular cooperation of the interaction domains in transactivation. An important implication of these results is that the domains of the plant bHLH TFs do not function as completely independent modules, and the strength of a particular domain may be regulated by the other regions of the protein.

In this study we have used an engineered bHLH TF to improve a plant phenotype. Utilization of TFs for the control of metabolic pathways has emerged as an effective approach for metabolic bioengineering. The results reported here confirm our previous studies of achieving novel TF functional properties through protein engineering. We have demonstrated that the function of a TF can be evolved by selecting libraries of TF variants using a high throughput, heterologous selection system and the improved protein function is confirmed in plants. Through extensive site-directed and saturation mutagenesis of multiple members of the Myc-like bHLH TF family, the present work also extends our knowledge of domain cooperation for this class of TFs.

A number of mechanisms may be responsible for the phenotypic improvement of M2-1 over the wild-type Myc-RP. First, the enhanced anthocyanin production may be due to the increased transactivation strength of M2-1. Interestingly, when comparing the transgenic tobacco plants expressing *Myc-RP* or *Delila*, two highly homologous TFs with marked different transactivation activity, the more active Delila generates a stronger anthocyanin phenotype (Pattanaik et al. unpublished results). Second, the stronger anthocyanin phenotype by M2-1 may be the combined outcome of increased transactivation strength and improved interaction with a tobacco Myb protein. The tobacco Myb TF interacting with Myc-RP has not been identified, however, it is possible that the interaction with this unidentified protein is improved in M2-1. Alternatively, because the bHLH, Myb and WD-40 factors work in combination, M2-1 may form a more stable bHLH-Myb-WD40 complex, which results in enhanced anthocyanin production. Nevertheless, although the exact mechanism remains undetermined; we have engineered a Myc-like bHLH TF that is more superior in up-regulating a plant metabolic pathway. While artificial DNA-binding regulatory proteins have been created and demonstrated functional in planta (Ordiz et al. 2002; Sanchez et al. 2006), our work, to the best of our knowledge, represents the first example by which a native TF has been modified and used to improve the production of a valuable plant secondary metabolite.

Acknowledgements This work is supported by a grant from the Kentucky Tobacco Research and Development Center, University of Kentucky (to L.Y.). We would like to thank Professor K. Saito for providing *Myc-RP* cDNA and Professor Susan Wessler for *Lc* cDNA, and express our appreciation to the John Innes Research Center for supplying the *Delila* cDNA. We also thank Dr. I. Maiti for providing the MMV promoter and helpful suggestions, Dr. Soumendra K. Naik for the assistance in molecular analysis of the transgenic plants, and K. Shen for the critical reading of the manuscript.

References

Bailey PC, Martin C, Toledo-Ortiz G, Quail PH, Huq E, Heim MA, Jakoby M, Werber M, Weisshaar B (2003) Update on the basic helix–loop–helix transcription factor gene family in *Arabidopsis thaliana*. Plant Cell 15:2497–2502

Chapman-Smith A, Lutwyche JK, Whitelaw ML (2004) Contribution of the Per/Arnt/Sim (PAS) domains to DNA binding by the basic helix–loop–helix PAS transcriptional regulators. J Biol Chem 279:5353–5362

Chatterjee R, Yuan L (2006) Directed evolution of metabolic pathways. Trends Biotechnol 24:28–38

de Vetten N, Quattrocchio F, Mol J, Koes R (1997) The an11 locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals. Genes Dev 11:1422–1434

Dey N, Maiti IB (1999) Structure and promoter/leader deletion analysis of mirabilis mosaic virus (MMV) full-length transcript promoter in transgenic plants. Plant Mol Biol 40:771–782

Feller A, Hernandez JM, Grotewold E (2006) An ACT-like domain participates in the dimerization of several plant basic-helix–loop–helix transcription factors. J Biol Chem 281:28964–28974

Goff SA, Klein TM, Roth BA, Fromm ME, Cone KC, Radicella JP, Chandler VL (1990) Transactivation of anthocyanin biosynthetic genes following transfer of B regulatory genes into maize tissues. Embo J 9:2517–2522

Goff SA, Cone KC, Chandler VL (1992) Functional analysis of the transcriptional activator encoded by the maize B gene: evidence for a direct functional interaction between two classes of regulatory proteins. Genes Dev 6:864–875

Gong ZZ, Yamagishi E, Yamazaki M, Saito K (1999a) A constitutively expressed Myc-like gene involved in anthocyanin biosynthesis from *Perilla frutescens*: molecular characterization, heterologous expression in transgenic plants and transactivation in yeast cells. Plant Mol Biol 41:33–44

Gong ZZ, Yamazaki M, Saito K (1999b) A light-inducible Myb-like gene that is specifically expressed in red *Perilla frutescens* and presumably acts as a determining factor of the anthocyanin forma. Mol Gen Genet 262:65–72

Gong ZZ, Yamazaki M, Saito A (2000) Critical role of alanine-161 in Delila protein involved in regulation of anthocyanin pigmentation for transcriptional activation in yeast. Plant Biotech 17:309–314

Grotewold E, Sainz MB, Tagliani L, Hernandez JM, Bowen B, Chandler VL (2000) Identification of the residues in the Myb domain of maize C1 that specify the interaction with the bHLH cofactor R. Proc Natl Acad Sci USA 97:13579–13584

Heim MA, Jakoby M, Werber M, Martin C, Weisshaar B, Bailey PC (2003) The basic helix–loop–helix transcription factor family in plants: a genome-wide study of protein structure and functional diversity. Mol Biol Evol 20:735–747

Hernandez JM, Heine GF, Irani NG, Feller A, Kim MG, Matulnik T, Chandler VL, Grotewold E (2004) Different mechanisms participate in the R-dependent activity of the R2R3 MYB transcription factor C1. J Biol Chem 279:48205–48213

Koes R, Verweij W, Quattrocchio F (2005) Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. Trends Plant Sci 10:236–242

Li X, Duan X, Jiang H, Sun Y, Tang Y, Yuan Z, Guo J, Liang W, Chen L, Yin J, Ma H, Wang J, Zhang D (2006) Genome-wide analysis of basic/helix–loop–helix transcription factor family in rice and *Arabidopsis*. Plant Physiol 141:1167–1184

Lu J, Sloan SR (2002) The basic helix–loop–helix domain of the E47 transcription factor requires other protein regions for full DNA binding activity. Biochem Biophys Res Commun 290:1521–1528

Martin C, Prescott A, Mackay S, Bartlett J, Vrijlandt E (1991) Control of anthocyanin biosynthesis in flowers of *Antirrhinum majus*. Plant J 1:37–49

Mol J, Grotewold E, Koes R (1998) How genes paint flowers and seeds. Trends Plant Sci 3:212–217

Morgenstern B, Atchley WR (1999) Evolution of bHLH transcription factors: modular evolution by domain shuffling? Mol Biol Evol 16:1654–1663

Ordiz MI, Barbas CF 3rd Beachy RN (2002) Regulation of transgene expression in plants with polydactyl zinc finger transcription factors. Proc Natl Acad Sci USA 99:13290–13295

Pattanaik S, Dey N, Bhattacharyya S, Maiti IB (2004) Isolation of full-length transcript promoter from the Strawberry vein banding virus (SVBV) and expression analysis by protoplasts transient assays and in transgenic plants. Plant Sci 167:427–438

Pattanaik S, Xie CH, Kong Q, Shen KA, Yuan L (2006) Directed evolution of plant basic helix–loop–helix transcription factors for the improvement of transactivational properties. Biochim Biophys Acta 1759:308–318

Payne CT, Zhang F, Lloyd AM (2000) GL3 encodes a bHLH protein that regulates trichome development in *Arabidopsis* through interaction with GL1 and TTG1. Genetics 156:1349–1362

Quattrocchio F, Wing JF, Leppen H, Mol J, Koes RE (1993) Regulatory genes controlling anthocyanin pigmentation are functionally conserved among plant species and have distinct sets of target genes. Plant Cell 5: 1497–1512

Quattrocchio F, Wing JF, van der Woude K, Mol JN, Koes R (1998) Analysis of bHLH and MYB domain proteins: species-specific regulatory differences are caused by divergent evolution of target anthocyanin genes. Plant J 13:475–488

Roberts SG (2000) Mechanisms of action of transcription activation and repression domains. Cell Mol Life Sci 57:1149–1160

Saito K, Yamazaki M (2002) Biochemistry and molecular biology of the late-stage of biosynthesis of anthocyanin: lessons from *Perilla frutescens* as a model plant. New Phytol 155:9–23

Sanchez JP, Ullman C, Moore M, Choo Y, Chua NH (2006) Regulation of *Arabidopsis thaliana* 4-coumarate:coenzyme-A ligase-1 expression by artificial zinc finger chimeras. Plant Biotechnol J 4:103–114

Smolen GA, Pawlowski L, Wilensky SE, Bender J (2002) Dominant alleles of the basic helix–loop–helix transcription factor ATR2 activate stress-responsive genes in *Arabidopsis*. Genetics 161:1235–1246

Sompornpailin K, Makita Y, Yamazaki M, Saito K (2002) A WD-repeat-containing putative regulatory protein in anthocyanin biosynthesis in *Perilla frutescens*. Plant Mol Biol 50:485–495

Stemmer WP (1994) Rapid evolution of a protein in vitro by DNA shuffling. Nature 370:389–391

Toledo-Ortiz G, Huq E, Quail PH (2003) The Arabidopsis basic/helix–loop–helix transcription factor family. Plant Cell 15:1749–1770

Yuan L, Kurek I, English J, Keenan R (2005) Laboratory-directed protein evolution. Microbiol Mol Biol Rev 69:373–392

Zheng L, Baumann U, Reymond JL (2004) An efficient one-step site-directed and site-saturation mutagenesis protocol. Nucleic Acids Res 32:e115