Available online at www.sciencedirect.com
ScienceDirect
Biochimica et Biophysica Acta 1769 (2007) 139–148
Biochimica et Biophysica Acta
BBA
www.elsevier.com/locate/bbaexp

Promoter paper

Promoter analysis of the *Catharanthus roseus* geraniol 10-hydroxylase gene involved in terpenoid indole alkaloid biosynthesis☆

Nitima Suttipanta [a,1], Sitakanta Pattanaik [a,b,1], Samir Gunjan [c], Claire H. Xie [a,b], John Littleton [b,c], Ling Yuan [a,b,*]

[a] *Department of Plant and Soil Sciences, University of Kentucky, Lexington, KY 40546, USA*
[b] *The Kentucky Tobacco Research and Development Center, University of Kentucky, Lexington, KY 40546, USA*
[c] *Department of Pharmaceutical Sciences, University of Kentucky, Lexington, KY 40546, USA*

Received 22 August 2006; received in revised form 13 December 2006; accepted 12 January 2007
Available online 30 January 2007

Abstract

Geraniol 10-hydroxylase (G10H) is an important enzyme in the biosynthetic pathway of monoterpenoid alkaloids found in diverse plant species. The *Catharanthus roseus* G10H controls the first committed step in biosynthesis of terpenoid indole alkaloids (TIA). The *C. roseus G10H* promoter sequence was isolated by a PCR-based genome walking method. Sequence analysis revealed that the *G10H* promoter contains several potential eukaryotic regulatory elements involved in regulation of gene expression. The major transcription start site of the promoter was mapped to an adenine 31 bp downstream of the TATA-box. For functional characterization, transcriptional fusions between the *G10H* promoter fragments with 5′ or 3′ deletions and the *GUS* reporter gene were generated and their expressions were analyzed in a tobacco protoplast transient expression assay. Deletion of the promoter down to −318 bp had little effect on GUS activity. However, further deletion of the promoter to position −103 resulted in approximately 5-fold reduction of GUS activity. Gain-of-function experiments revealed the presence of three potential transcriptional enhancers located in regions between −191 and −147, −266 and −188, and −318 and −266, respectively. The *G10H* promoter was capable of conferring stable *GUS* expression in transgenic tobacco plants and *C. roseus* hairy roots. In transgenic tobacco seedlings *GUS* expression was tissue-specific, restricted to leaf and actively growing cells around the root tip, and not detected in the hypocotyls, root cap and older developing areas of the root. The *GUS* expression in both transgenic *C. roseus* hairy roots and tobacco seedlings were responsive to fungal elicitor and methyljasmonate. Compared to other known promoters of TIA pathway genes, the *G10H* promoter contains unique binding sites for several transcription factors, suggesting that the *G10H* promoter may be regulated by a different transcriptional cascade.

Keywords: Promoter; Geraniol 10 hydroxylase gene; Terpenoid indole alkaloid; Catharanthus roseus; Transgenic plants

1. Introduction

Catharanthus roseus (Madagascar periwinkle) is the natural source for about 130 terpenoid indole alkaloids (TIA) including the well-known anticancer agents vinblastine and vincristine. These alkaloids are spindle toxins that have proven effective in chemotherapy treatments for leukemia and Hodgkin's disease (lymph node and spleen cancer). Many TIA are also true lead compounds for drug development. TIA are synthesized from the condensation of tryptamine and secologanin. Tryptamine is derived from tryptophan by the tryptophan decarboxylase (Tdc), whereas secologanin is derived from geranyl pyrophosphate. The condensation of tryptamine and secologanin is catalyzed by strictosidine synthase (Str), and the product strictosidine can be modified by strictosidine glucosidase (Sgd) to produce cathenamine, a precursor to various biologically active alkaloids [1]. Geraniol 10-hydroxylase (G10H) hydroxylates geraniol to form 10-hydroxy-geraniol and thus controls the first committed step in biosynthesis of secologanin and TIA. The *C. roseus* G10H is a cytochrome P450 monooxygenase belonging to the CYP76B superfamily

☆ The sequence data presented in this article has been entered into the EMBL/GenBank database under accession number EF363554.

* Corresponding author. Department of Plant and Soil Sciences, KTRDC, Cooper and University Drives, Lexington, KY 40546, USA. Tel.: +1 859 257 4806; fax: +1 859 323 1077.
E-mail address: lyuan3@uky.edu (L. Yuan).

[1] These authors contributed to this work equally.

0167-4781/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbaexp.2007.01.006

[1]. The gene encoding a cytochrome P450 reductase (Cpr), presumably acting as the electron donor for G10H, has been identified and characterized [2]. The gene expression profiles of the *Cpr* and *G10H* of *C. roseus* are similar and they respond to jasmonic acid (JA) induction with similar kinetics [1].

Responsiveness to hormonal or pathogenic signals, such as JA and fungal elicitors, is a hallmark of TIA pathway enzymes. In *C. roseus* the mRNA levels of several key TIA pathway genes, including *Tdc*, *Str*, *Sgd* and *Cpr*, increase dramatically upon exposure to JA (or its methyl-ester, MeJA) and fungal elicitors [3,4]. A similar response to JA induction by *G10H* has also been reported [1]. The expression of *C. roseus G10H* in response to fungal elicitors has not been reported. A number of JA-responsive transcription factors, including transcriptional activators ORCA2 and ORCA3 [3] as well as repressors ZCT1, ZCT2 and ZCT3 [5], are capable of interacting with the promoters of *Tdc*, *Str* and *Sgd*. These findings present a complex regulatory mechanism of the TIA pathway that involves multiple transcription factors responding to various developmental and environmental signals. The observation that *G10H* is not regulated by the ORCA transcription factors [3] implicates control of a subset of the TIA pathway genes by other JA- or elicitor-responsive transcription factors [6]. The availability and characterization of the *G10H* promoter will facilitate the identification of additional regulators capable of interacting with the promoter and lead to further understanding of the mechanisms underlying TIA biosynthesis.

Here we report the isolation and characterization of the *C. roseus G10H* promoter. Our results suggest that the promoter contains a number of *cis*-elements and transcriptional enhancers. The isolated promoter can control gene expression in transgenic *C. roseus* and tobacco cells and are responsive to induction by JA and fungal elicitors. The presence of several unique transcription factor binding sites in the promoter infers the possible roles of these regulators in TIA biosynthesis.

2. Materials and methods

2.1. Genome walking and isolation of G10H promoter

The *C. roseus* G10H promoter sequence was isolated by a PCR based genome walking procedure using the GenomeWalker Universal Kit (BD Biosciences, USA). Genomic DNA was isolated from *C. roseus* seedlings using DNeasy plant mini kit (Qiagen). Two non-overlapping gene specific primers (G10H1: 5′-gtgtggttggtcgcctaataaatggag-3′ and G10H3: 5′-ggtaaggtaatccatggaatggaagtg-3′) were designed based on the G10H cDNA sequence information available in the GenBank database. An approximately 0.5 kb fragment was isolated from *Eco*RV digested genomic DNA ligated to the GenomeWalker adapter sequence. The primary PCR was carried out using an adapter-specific primer and a gene-specific primer and followed by a second PCR amplification reaction using nested adapter-specific and gene-specific primers. All PCR reactions were amplified using Advantage 2 Polymerase mix (BD Biosciences, USA) as per manufacturer's instruction. The major product from the second PCR was gel purified and cloned into the pGEM-T Easy vector (Promega, USA). A plasmid containing the appropriate size insert was completely sequenced (Fig. 1). Based on this sequence information a forward primer with an *Eco*RI restriction site (G10H-FW: 5′-gcggaattctcgttatagaatttctcgcttca-3′) and a reverse primer with a *Hin*dIII restriction site (G10H-RV: 5′-atgaagcttggaatggaagtgtacaatttg-3′) were designed. To verify the specificity of the sequence obtained by genome walking, the promoter fragment was re-amplified from the genomic DNA using these primers. The fragment was cloned and sequenced to check the integrity. The plasmid containing the G10H promoter was designated as pGEM-G10H.

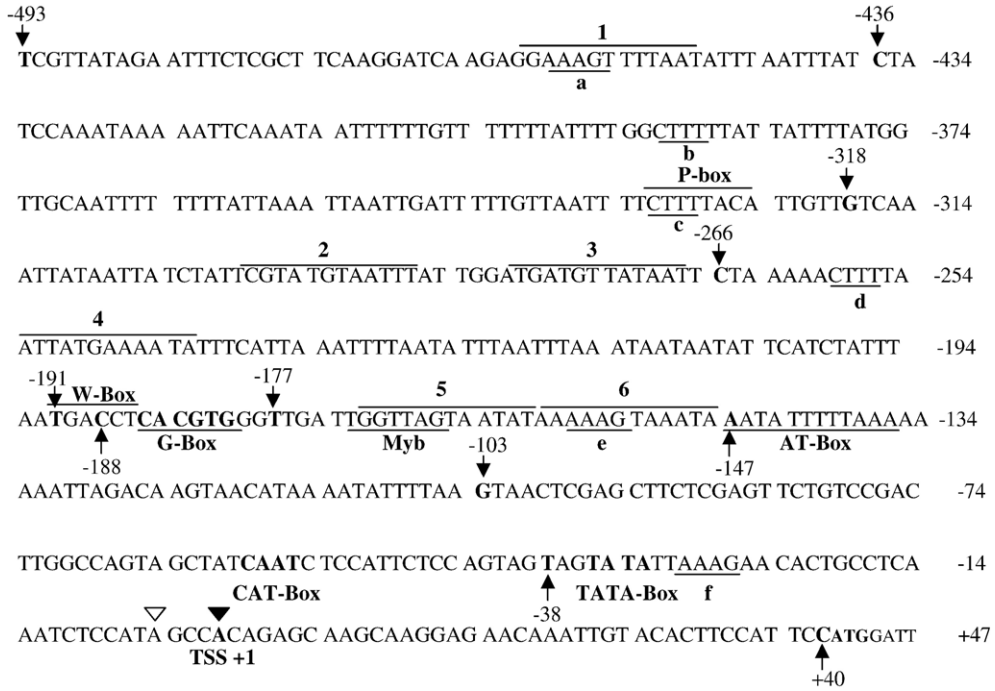

Fig. 1. The DNA sequence of *Catharanthus roseus G10H* promoter. A 533-bp fragment (−493 to +47) with respect to the transcription start site (TSS) is presented. The two possible TSS are indicated with reversed triangles. The major TSS is indicated by the solid triangle and designated as +1. An arrow above or below the sequence indicates the end point of 5′ or 3′ deletion fragments, respectively. The TATA and CAT boxes are shown in bold. Important *cis*-elements are indicated in bold. Potential GT-1 binding sites are indicated by number 1 to 6. Putative Dof transcription factor binding core sequences are underlined and labeled as a through f.

Table 1
Sequence of the oligonucleotides used in this study

Primer no	Primer sequence
G10 (−493)F	5′ GC*GAATTC* TCGTTATAGAATTTCTCGCTTCA 3′
G10 (−436)F	5′ GC*GAATTC* CTATCCAAATAAAAATTCAAATAATT 3′
G10 (−318)F	5′ GC*GAATTC* GTCAAATTATAATTATCTATTCGTA 3′
G10 (−266)F	5′ GC*GAATTC* CTAAAAACTTTTAATTATGAAAATAT 3′
G10 (−191)F	5′ GC*GAATTC* TGACCTCACGTGGGTTGATTG 3′
G10 (−177)F	5′ GC*GAATTC* TTGATTGGTTAGTAATATAAAAAGT 3′
G10 (−103)F	5′ GC*GAATTC* GTAACTCGAGCTTCTCGAGTT 3′
G10 (+40)R	5′ ATG*AAGCTT* GGAATGGAAGTGTACAATTTG 3′
G10 (−38)R	5′ GCT*AAGCTT* ACTACTGGAGAATGGA3′GATTGA 3′
G10 (−147)R	5′ GCG*GTCGAC* TTATTTACTTTTTATATTACTAACC 3′
G10 (−188)R	5′ GCT*AAGCTT* GTCATTAAATAGATGAATATTATTAT 3′
G10 RACE1	5′ GAATTCAAAACTTTTCGAAGACT 3′
G10 RACE2	5′ TCCTGTTTTTGAAGAACTTCTTTC 3′
G10 RACE3	5′ CGTGTTTTTTGGAAAGTTTTGC 3′
GUS-FW	5′ ATGGGTTTACGTCCTGTAGAAACC 3′
GUS-RV	5′ TCATTGTTTGCCTCCCTGCTG 3′

2.2. Determination of the G10H promoter transcription start site by 5′-RACE

The transcription start site (TSS) of the G10H promoter was determined by 5′-rapid amplification of cDNA ends (5′-RACE). Total RNA was isolated from *C. roseus* seedlings using the RNeasy Plant mini kit (Qiagen, USA). RACE was carried out using the 5′-RACE System 2 (Invitrogen, California, USA) following protocols supplied by the manufacturer. A gene specific primer (G10RACE 1; Table 1) was used for the synthesis of first-strand cDNA. For amplification of the dC-tailed first-strand cDNA, the adapter primer from the kit and two nested gene-specific antisense primers (G10RACE 2 and 3; Table 1) were used in subsequent PCR reactions. The 5′-RACE PCR products were cloned into pGEM-T Easy (Promega, USA) and 10 positive clones were sequenced to determine the transcription start site.

2.3. Construction of plasmids for protoplast transient expression assay

A series of G10H promoter fragments of various lengths (Fig. 2) were PCR-amplified from pGEM-G10H with appropriately designed primers (Table 1). PCR amplification was carried out in a 25 μl volume for 30 cycles under the following standard conditions: denaturation (94 °C for 30 s), annealing (55 °C for 30 s), extension (72 °C for 30 s) using Easy-A-high fidelity PCR cloning enzyme (Stratagene, USA). Each PCR-amplified fragment was gel-purified, digested with *Eco*RI and *Hin*dIII and cloned into the corresponding sites of the protoplast expression vector pUCPMAGUS [7]. A total of eight plasmids were constructed, corresponding to coordinates −493 to +40 (pG10H1), −436 to +40 (pG10H2), −318 to +40 (pG10H3), −191 to +40 (pG10H4), −177 to +40 (pG10H5), −103 to +40 (pG10H6), −436 to −38 (pG10H7), and −436 to −188 (pG10H8).

2.4. Construction of plasmids for gain-of-function experiments

For gain-of-function experiments, five *cis*-regions (corresponding to positions −191 to −147, −266 to −188, −266 to −147, −318 to −188 and −318 to −147 relative to the TSS; Fig. 3) were individually inserted upstream of the minimal promoter, pG10H6 (−103 to +40 relative to TSS). The *cis*-regions were individually PCR-amplified from the plasmid pGEMG10H to generate fragments of the following general structure: 5′-*Eco*RI-*cis*-region–*Hin*cII-3′. The G10H minimal promoter was PCR-amplified as a *Hin*cII–*Hin*dIII fragment. The minimal promoter with the respective *cis*-region was then cloned into the corresponding site of the pUCPMAGUS protoplast expression vector. The resulting constructs were designated pG6a, pG6b, pG6c, pG6d and pG6e (Fig. 3). Each promoter fragment was verified by DNA-sequencing.

2.5. Protoplast isolation and electroporation

Isolation of protoplasts from tobacco cell suspension cultures (*Nicotiana tabacum* L. cv. Xanthi 'Brad') and electroporation of tobacco protoplasts with plasmid DNA containing the promoter fragment fused to a β-glucuronidase (GUS) gene were performed as described previously [8]. A Gene Pulser II

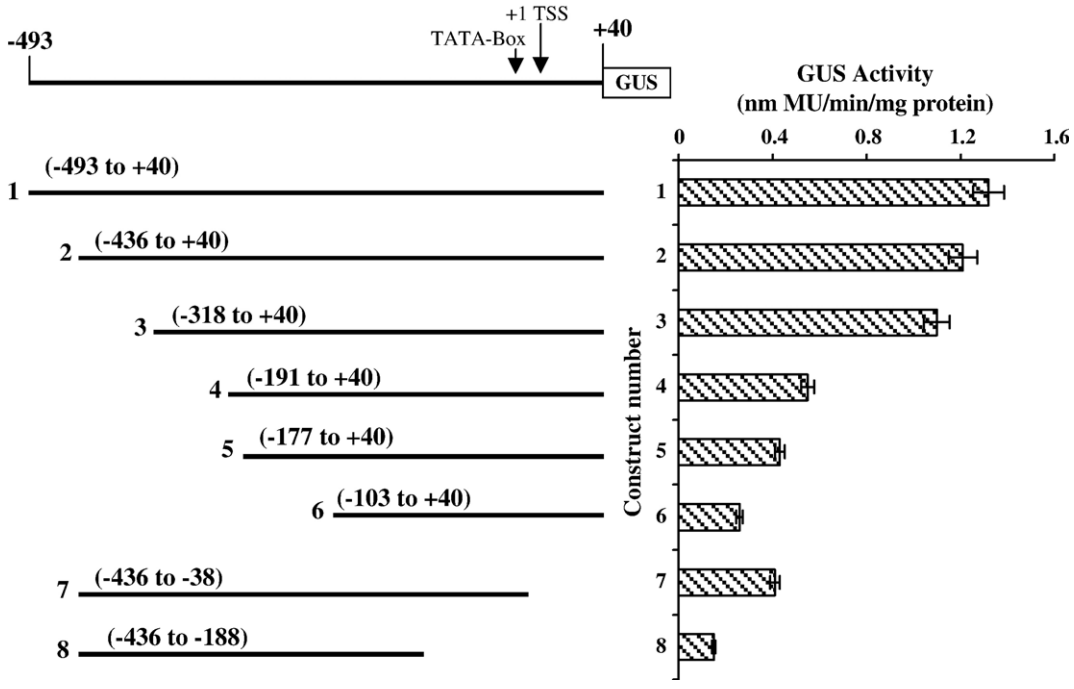

Fig. 2. Deletion analysis of the *G10H* promoter-GUS constructs. The left panel shows a schematic map of the promoter deletion constructs. The 5′- and 3′-end coordinates of the deletion fragments are given in parentheses. At the top, the relative position of the TATA box, transcription start site (TSS, +1) and the G10H promoter coordinates are indicated. The right panel shows the corresponding GUS activities of each construct. The GUS activities were determined in the protoplast transient expression assay. Each construct was assayed at least three times in three independent experiments. The relative GUS activity (nm MU/min/mg protein) with standard deviation is presented.

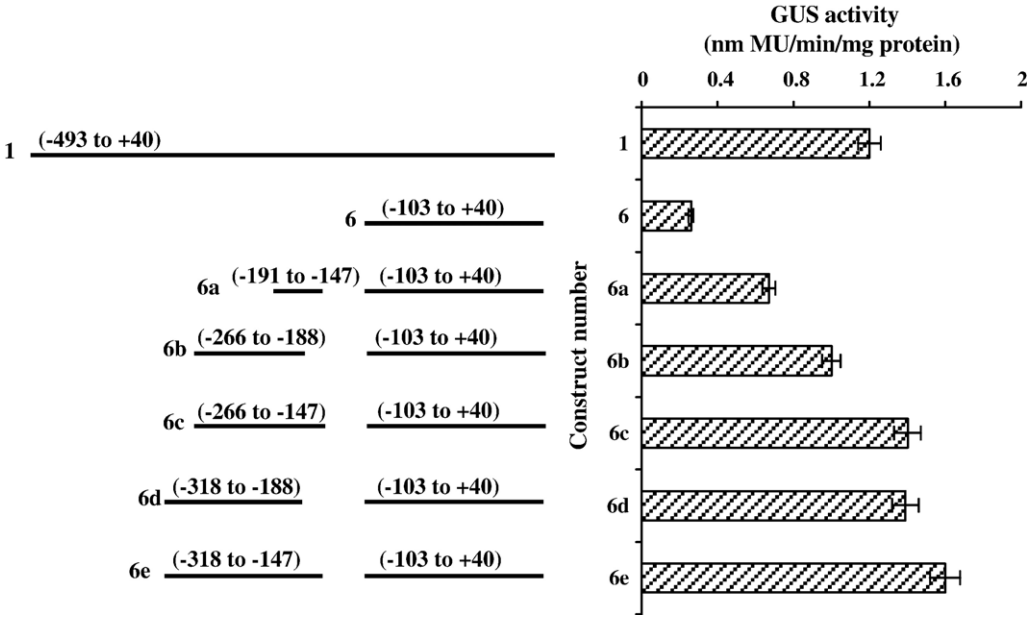

Fig. 3. Gain-of-function analysis of the *G10H* promoter. The left panel shows a schematic map of the constructs. Construct 1 represents the full-length promoter and construct 6 represents the minimal promoter (equivalent to construct 6 in Fig. 2). In constructs 6a to 6e, 5′-fragments of selected length are fused to the minimal promoter. The right panel indicates the corresponding GUS activities of each construct. The GUS activities were determined using the protoplast transient expression assay. Each construct was assayed at least three times in three independent experiments. The relative GUS activity (nm MU/min/mg protein) with standard deviation is presented.

apparatus (BioRad, USA) with the Capacitance Extender II (Model 165–2107) was used for electroporation. An aliquot of 750 µl containing 2×10^6 protoplasts in a 0.4 cm electroporation cuvette was electroporated with 10 µg of plasmid DNA (200 V and 950 µF capacitance for 50–60 ms). A plasmid containing the fire-fly luciferase coding sequence under the control of CaMV35S promoter and rbcS terminator was co-electroporated with each promoter construct as an internal control. After 20–22 h, protoplasts were harvested for GUS and luciferase activity assays. All constructs were tested in at least three independent experiments. The GUS activity was normalized against luciferase activity. A promoter-less *GUS* plasmid was used as a negative control in this transient expression assay.

2.6. β-Glucuronidase and Luciferase assays

Fluorometric assays to measure GUS activity in transfected protoplasts or plant tissue and histochemical staining to localize the distribution of GUS activity in plants were performed according to protocols described in Jefferson et al. [9]. Total protein concentration in plant extracts was determined using the Bio-Rad protein assay (Bio-Rad, CA, USA) with bovine serum albumins as standard. Luciferase activity in transfected protoplasts was measured using a Luciferase assay system (Promega, CA, USA) and a luminometer (Model No. TD2020; Turner Designs, CA, USA).

2.7. Construction of plant expression vectors for transformation of tobacco and C. roseus

The G10H promoter fragment isolated from pG10H1GUS (−493 to +40 relative to the TSS) was cloned into the plant expression vector pKYLX71-GUS [10] at *Eco*RI and *Hin*dIII sites. The resulting plant expression vector was designated pKG10HGUS. The construct was introduced into *Agrobacterium tumefaciens* strain C58C1:pGV3850 by freeze–thaw. Tobacco plants (*Nicotiana tabacum* cv Samsun NN) were transformed with the *Agrobacterium* containing pKG10HGUS as described previously [11]. Twelve independent plant lines were generated for this construct. Regenerated, kanamycin-resistant tobacco plants were grown under greenhouse conditions.

Transformation of *C. roseus* with *Agrobacterium rhizogenes* for production of transgenic hairy root cultures was conducted as previously described [12]. The plasmid pKG10HGUS was mobilized into *Agrobacterium rhizogenes* R1000 by freeze–thaw method. After removal of the roots, 1-week-old *in vitro* germinated *C. roseus* seedlings were immersed in the bacterial suspension for 30 min and then transferred onto solid MS basal medium [13] for co-cultivation. Seedlings infected with *A. rhizogenes* R1000 without the binary vector were used as control. After 48 h the explants were transferred to half-strength MS basal medium containing cefotaxime (400 mg/L). Hairy roots developed at the cut surface of explants in 3–4 weeks.

2.8. Analysis of transgenic tobacco plants: isolation of RNA and RT-PCR analysis

Total cellular RNA from leaves and roots of transgenic tobacco seedlings expressing the pKG10GUS construct was isolated using the RNeasy plant mini kit (Qiagen, Chatsworth, USA). Total RNA (4 µg) was treated with RNase-free DNase (Sigma, USA) and used for synthesis of first-strand cDNA 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 (Table 1) to detect the full-length *GUS* transcript. 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 30s, 55 °C for 30 s, 72 °C for 2 min). As a negative control, the 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.

2.9. Jasmonate induction, fungal elicitor treatment and GUS assay

The transgenic tobacco seedlings and hairy roots induced from hypocotyl segments of *C. roseus* were used for elicitor treatments. Methyl jasmonate (MeJA; Bedoukian Research, USA) was diluted with DMSO and added to liquid MS medium at a final concentration of 100 µM. Only DMSO was added to the control treatment. The seedlings or hairy roots were incubated for 8–20 h in the medium and then either stained using GUS histochemical staining buffer containing 1 mg/ml X-Gluc (5-bromo-4-chloro-3-indolyl-beta-D-glucuronide; Gold Biotech, USA) or assayed for GUS activity using MUG (4-methylumbelliferyl-β-D-glucuronide;

Sigma, USA). For fungal elicitor treatments, transgenic seedlings or transformed hairy roots were incubated for 12–24 h in 0.5–1.5% yeast extract in water (Difco). Control (water only) and treatment samples were then stained with X-Gluc or used for GUS activity assay. The pictures of the GUS-stained roots were captured using an Olympus BX51 microscope with Olympus DP12 digital camera. The color intensities of the roots were quantified using the ImageJ software.

3. Results and discussion

3.1. Structure and sequence analysis of the G10H promoter

An approximately 0.5 kb *G10H* promoter fragment was isolated by PCR-based genome walking (Fig. 1). Sequence analysis using PLACE (www.dna.affrc.go.jp/PLACE) [14] and the NSITE-PL program (www.softberry.com) reveals that the G10H promoter contains several consensus eukaryotic regulatory domains. The TATA-box is present 31-bp upstream of the major transcription start site (TSS). The TATA proximal sequences up to −103 of the G10H promoter are relatively GC-rich while the distal regions are AT-rich (Fig. 1). A CAT-box-like sequence (CAAT) is present 19-bp upstream of the TATA-box sequence. In addition to the TATA-box and CAT-box, several other potential regulatory elements with roles in regulation of gene expression are present in G10H promoter. These include a prolamin box (P-box) like sequence (CTTTTACA) at position −331 to −324, a G-box (CACGTG) at position −185 to −180, a Myb-type transcription factor recognition sequence (TGGTTAG) at position −172 to −166, an AT-box (AATATTTTTAAA) at position −147 to −136 and a W-box (TGACCT) at position −191 to −186 relative to the TSS (Fig. 1). In addition, the G10H promoter contains six GT-1 transcription factor recognition sequences and six Dof (DNA binding with one finger) transcription factor recognition core sequences (AAAG and CTTT).

The P-box, or '−300 elements', is found in the promoters of many cereal seed storage protein genes. This *cis*-element is identified on the basis of both its highly conserved nucleotide sequence (TGTAAAG) and position (−300 region) relative to translation start codon of the prolamin genes. The P-box has been shown to interact with endosperm-specific Dof transcription factors and may be involved in activation of storage protein gene expression during endosperm development [15]. The G-box (CACGTG) is a highly conserved DNA sequence that has been identified in the promoter regions of many plant genes regulated by a variety of environmental and physiological factors. G-box or G-box-like sequences are reported to be present in the promoters of other TIA biosynthetic pathway genes. G-boxes have also been shown to effect gene expression in response to light, abscisic acid and methyl jasmonate [16]. The W-boxes are present in the promoters of many plant defense-related genes. The WRKY transcription factors, which are involved in the regulation of various plant physiological processes including pathogen defense and trichome development, bind specifically to the W-box elements [17]. The AT-box has been found in the 5′ upstream region of many plant genes regulated by light, and the nuclear factor AT-1 binds specifically to this AT rich sequence [18]. GT elements were first identified in the promoter region of the photo-regulated pea ribulose bisphosphate carboxylase/oxygenase (Rubisco) small sub-unit as box II. The GT-elements have since been identified in the promoter regions of many plant genes including some that are not regulated by light [19]. The GT-1 transcription factors recognize the consensus sequence (G/A)NNN(G/A)GT(A/T)AA(A/T)(A/T) [20]. The promoter regions of several genes involved in TIA biosynthesis in *C. roseus* contain the consensus GT-elements that bind the nuclear factor GT-1 [21,22]. Mutagenesis of the GT-elements in the *C. roseus Tdc* promoter significantly lowers its response to UV light [20].

Dof proteins are plant specific transcription factors which bind to DNA with an AAAG core sequence or its reverse complement sequence CTTT. These transcription factors have been isolated from many monocot and dicot plants and play critical roles as transcriptional activators or repressors in plant growth and development [23]. The abundance of Dof binding sites in the *G10H* promoter is interesting. Several genes involved in *C. roseus* TIA biosynthesis are regulated by multiple zinc finger proteins. Three two-fingered transcription factors have been demonstrated to bind both *Tdc* and *Str* promoters and act as transcriptional repressors [5]. These zinc finger repressors function in junctions with multiple activators to provide combinatorial control of environmental responses and timing of gene expression. To date no Dof protein has been directly implicated in *C. roseus* TIA biosynthesis. While *G10H* differs from *Tdc*, *Str* and *Sgd* in the response to ORCA transcription factors, it is not clear whether the two-fingered factors affect the *G10H* expression. Because G10H is likely to be regulated by different elicitor-responsive transcription factors than those controlling *Tdc*, *Str* and *Sgd*, it will be worthwhile to explore the possibility that Dof proteins play a significant role in regulation of *G10H* expression.

Comparison of the known promoter sequences of *C. roseus* TIA biosynthetic pathway genes, such as *Tdc* and *Str*, reveals that some of the *cis*-elements described above are unique to the *G10H* promoter. These elements include W-box, MYB-transcription factor recognition sequence, AT-box and P-box-like sequence (TGTAAAG/CTTTTACA). The presence of such novel *cis*-elements suggests that the *G10H* promoter may be regulated by a different transcriptional cascade.

3.2. Transcription start site of the G10H promoter

The TSS of the *G10H* promoter was determined by 5′-RACE using total RNA isolated from *C. roseus* seedlings followed by cloning and sequencing of amplified PCR products. More than ten 5′-RACE products were sequenced and the mRNA 5′ end points suggested two possible TSS (Fig. 1). One starts at an adenine (A) base located 31-nucleotides downstream of the TATA box and the other one starts at an adenine 4 bp upstream. Joshi [24] has analyzed over 75 published genomic DNA sequences from several higher plant species and found that an adenine flanked by pyrimidine bases is present at the TSS in the majority of cases. The adenine 31 nucleotides downstream of the TATA box conforms with this consensus and is therefore considered to be the major TSS and designated as position +1 (Fig. 1).

3.3. 5′and 3′-end deletion analysis

For functional characterization of the *G10H* promoter, the 533 bp promoter fragment was subjected to progressive 5′- and 3′-end deletion analysis. Eight promoter fragments were generated by PCR amplification and cloned upstream of the *GUS* reporter gene to create transcriptional fusions. A schematic map of deletion constructs and the corresponding GUS activities of each construct are shown in Fig. 2. Constructs containing the promoter fragments fused to the *GUS* gene were introduced into tobacco protoplasts for transient expression assays. The expression level of the full-length promoter (construct 1; Fig. 2) was 1.32 nmol MU/min/mg protein, and this value was considered to represent full (100%) promoter activity for comparison with the deletion constructs. Deletion of the *G10* promoter from the 5′-end down to −318 had small effects on the promoter activity (constructs 2 and 3). However, further deletion down to −191 (construct 4) markedly reduced the promoter strength; less than 40% of the GUS activity remained in comparison to the full-length promoter. Further deletion of 14 bp (construct 5) resulted in an additional 10% loss in activity. Deletion to position −103 (construct 6) decreased the promoter activity to approximately 20% of that of the full-length promoter. This promoter region (−103 to +40) with minimal activity contains the TATA-box but not other known regulatory elements and is thus considered as the minimal promoter. Sequence analysis, discussed above, indicates that the promoter region between −318 bp and −147 bp contain a number of potential *cis*-elements. These *cis*-regulatory elements are well-characterized and have been shown to be involved in the regulation of promoter activities in many plants. Our deletion analysis indicates that these *cis*-elements are critical for maximal activity of the *G10H* promoter; deletions of these elements lead to significant reductions in promoter strength. The deletion analysis also indicates that the TATA box is important for maximal promoter activity; the 3′-end deletion constructs 7 and 8 that lack the TATA box exhibit considerably lower GUS activities compared to that of the full-length promoter (Fig. 2).

3.4. Gain-of-function experiments

While loss-of-function analyses such as deletion studies can be informative especially in determining the minimum requirement for promoter functions, gain-of-function experiments often further define the regulatory functions of promoter sub-regions. In order to further characterize the *cis*-elements that influence *G10H* promoter activity, the region containing the majority of *cis*-elements (−318 to −147) was divided into four sub-regions (−191 to −147, −266 to −188, −266 to −147, −318 to −188, and −318 to −147, respectively). These sub-regions were fused with the minimal promoter (construct 6; Fig. 2) to generate constructs 6a–6e (Fig. 3). These constructs were analyzed using a tobacco protoplast transient expression assay. As shown in Fig. 3, GUS activity of construct 6a was 2.5-fold higher than the minimal promoter. Construct 6b (−266 to −188) has a comparable level of GUS activity as the full-length *G10H* promoter. Constructs 6c, 6d and 6e showed higher activities than the full-length *G10H* promoter. The highest GUS activity from these constructs (construct 6e) is approximately 20% above that of the full-length promoter. Taken together these data implicate the presence of three potential transcription-enhancing *cis*-regulatory sequences. The first enhancing sequence is present in the 45-bp region between −191 and −147 as it significantly enhances transactivation of the minimal promoter (6a; Fig. 3). The second, and perhaps even stronger enhancing sequence is present in the 78-bp region between −266 and −188 (6b; Fig. 3). Combining these two enhancing sequences increases GUS activity higher than that observed for the full-length promoter (6c; Fig. 3). The third enhancing sequence is present in the 52-bp region between −318 and −266. Construct 6d, which contains the second and third enhancing sequences, has a similar enhancing effect to construct 6c which contains the first and second enhancing sequences (6d; Fig. 3). The combination of all three enhancing sequences results in the highest improvement in transactivation (6e; Fig. 3). Transcriptional activity of many plant and animal promoters is often the result of combinatorial or synergistic interactions of different *cis*-elements with cognate transacting nuclear factors [25]. Because isolated enhancing sequences lead to higher transcriptional activity than that of the full-length promoter, our results also implicate the presence of negative regulatory elements within the *G10H* promoter. The region between −318 to −188 is AT-rich and contains several putative GT-1 binding sites. The presence of potential GT-1 binding sites in the enhancer region of the promoter suggests that this nuclear factor may play a role in the *G10H* promoter function.

3.5. G10H promoter-controlled GUS expression in transgenic tobacco plants and C. roseus hairy roots

Twelve independent tobacco transgenic lines were developed for the construct pKG10HGUS. Seedlings of transgenic lines (R1 progeny) showing a segregation ratio (Kan^R:Kan^S) of 3:1 for the kanamycin (*Kan*) marker gene were selected for further analysis. RT-PCR analysis was conducted using total RNA isolated from leaves and roots of 4-week-old seedlings to amplify the full-length *GUS* transcript. In all selected transgenic lines the expected 1.8 kb *GUS* product was detected in both leaves and roots (data not shown). Histochemical staining of whole seedlings indicated *GUS* expression was localized to leaves and roots (Fig. 4A–E). No GUS staining was observed in hypocotyls. In roots, *GUS* expression was restricted to actively growing cells around the apical meristem and appeared to decrease at older developmental stages (Fig. 4C–E). No GUS accumulation was observed in the root cap, root hair and older developing regions of the roots (Fig. 4D). These results show that in addition to controlling transient gene expression in tobacco protoplasts, the *G10H* promoter is capable of conferring stable gene expression in transgenic tobacco plants.

The biosynthetic pathway of monoterpenoid indole alkaloids (MIA) in *C. roseus* has been studied extensively in the past two decades. Experiments with germinating seedlings have demonstrated that alkaloid biosynthesis and accumulation are developmentally regulated and the genes involved in the MIA biosynthesis are preferentially expressed in actively growing

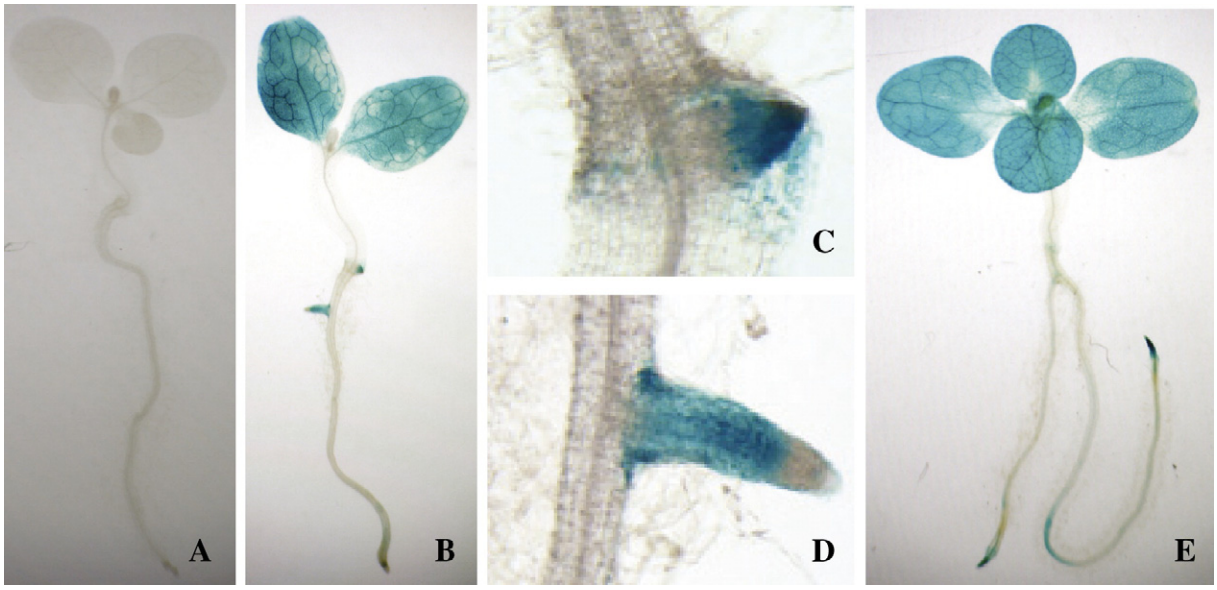

Fig. 4. Histochemical localization of GUS activity in transgenic tobacco seedlings: 2 weeks old, control (A) and transgenic (B) seedlings were stained for GUS activity. (C–D) Close-up view of developing roots showing GUS localization in root tip. (E) A 4-week-old seedling showing GUS staining in leaves and root tips.

tissues such as young leaves and root tips [26]. RNA blot analysis, *in situ* RNA hybridization and immunocytochemistry have been used to locate the tissue/cell-type involved in MIA biosynthesis [26,27]. Northern blot analysis showed that the transcriptional expression of *G10H* is higher in roots and actively growing aerial organs (young leaf, internode) as compared to mature organs (mature/old leaf, fruit) [27]. *In situ* RNA hybridization and immunocytochemistry experiments have revealed that mRNAs of genes, including *G10H*, known to be involved in biosynthesis of the terpenoid part of MIA accumulate in plant vascular cells. In contrast, genes involved in the earlier steps of the MIA pathway, such as *Tdc* and *Str*, are expressed exclusively in epidermal cells of young leaves. In roots expression of *Tdc* and *Str* is restricted to protoderm and cortical cells around the apical meristem. These mRNA do not appear to accumulate in root caps and steles [26,27]. Recently more sophisticated techniques including laser-capture microdissection and carborundum abrasion have been applied to localize MIA metabolites and expression of the biosynthetic pathway genes at the cellular level [28]. These experiments indicate that *G10H* transcripts are expressed in epidermal and laticifer cells as well as in vascular cells of leaves. Our results of *G10H* controlled *GUS* expression in transgenic tobacco corroborate with the organ/tissue specific localization *G10H* transcripts in *Catharanthus* tissue described above. In transgenic tobacco roots, the *GUS* expression pattern is identical to the gene expression patterns observed for *Tdc* and *Str* in *C. roseus* roots [26].

In this study *C. roseus* hairy roots were also developed from hypocotyls of young seedlings by *A. rhizogenes*-mediated transformation of the pKG10HGUS vector. Histochemical staining of *C. roseus* seedlings showed intense GUS staining only in hairy roots (Fig. 5A–B). GUS expression was not observed in hairy roots developed from *C. roseus* seedlings inoculated with the wild-type *A. rhizogenes* control (data not shown).

3.6. Responses of G10H promoter to methyljasmonate and fungal elicitor induction in transgenic tobacco seedlings and C. roseus hairy roots

Expression of several *C. roseus* TIA pathway genes are responsive to induction by JA or its volatile derivative methyljasmonate (MeJA) [1]. *G10H* mRNA is undetectable in Northern blot analysis of *C. roseus* suspension cells without JA treatment but accumulates to high levels following JA induction. A similar JA induction profile has been observed for a cytochrome P450 reductase (*Cpr*) in the same cells. Other genes involved in the *C. roseus* TIA pathway, including *Tdc*, *Str* and *Sgd*, also respond to JA induction however with different kinetics [1]. To investigate the effect of JA on *G10H* promoter activity, transgenic tobacco seedlings and *C. roseus* hairy roots expressing pKG10HGUS were treated with 100 μM MeJA for 8–20 h. While untreated control roots (in the presence of 0.1% DMSO) showed basal GUS activity, JA treatment resulted in greater than 8-fold increase of GUS activity (Fig. 6). GUS activity of MeJA treated tobacco seedlings was also higher than the untreated ones (Fig. 7). These data agree well with the JA-responsive expression of *C. roseus G10H* and suggests that the *G10H* promoter contains elements that are recognized by JA-responsive factors. G-boxes and CAT-boxes in plant gene promoters have been implicated as JA responsive elements [29,30]. The GCC-box (AGCCGCC) may play a role in JA-mediated gene activation [31]. However, a GCC-box has not been identified in the *C. roseus G10H* promoter. It has also been suggested that palindrome-forming inverted repeats in the promoter are recognized by JA-responsive factors [29]; however, no significant inverted repeat has been detected in the *G10H* promoter.

Fungal elicitors induce responses from the *C. roseus* TIA pathway genes. In addition to *Tdc* and *Str*, the cytochrome P450 reductase gene, *Cpr*, also responds to fungal elicitor induction

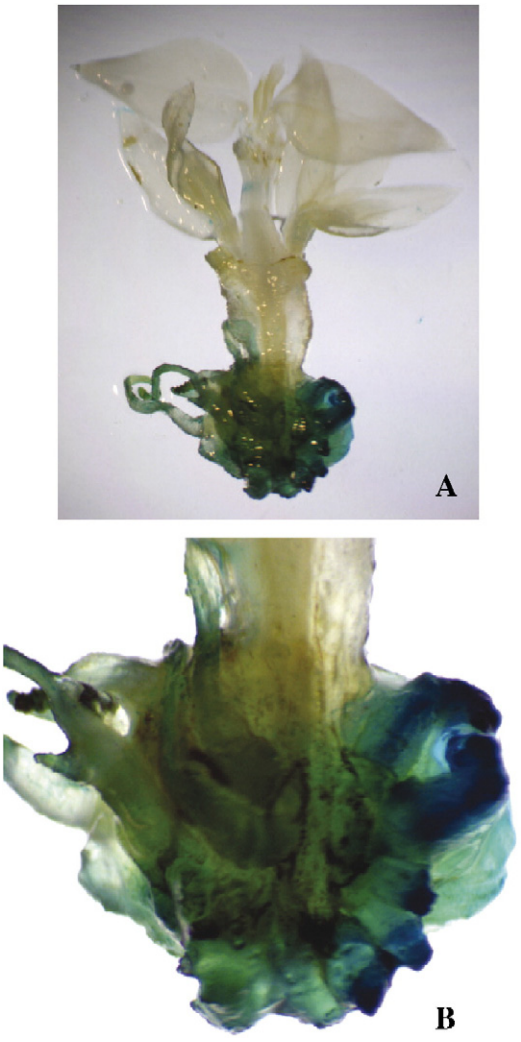

Fig. 5. Histochemical localization of GUS activity in transgenic *C. roseus* hairy roots. (A) A 4-week-old *A. rhizogenes*-inoculated seedling showing GUS activity in transgenic hairy roots. (B) A close-up view of young developing roots showing GUS localization in root tip.

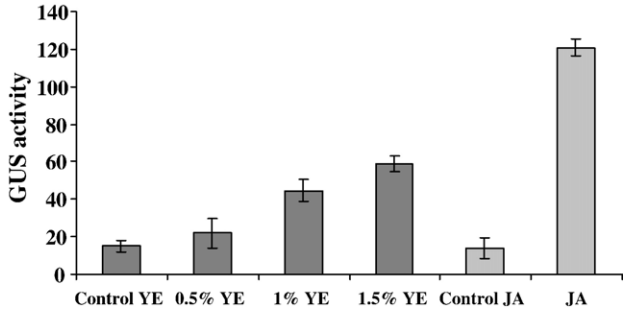

Fig. 6. Fungal elicitor and methyljasmonate (MeJA) induction of the *G10H* promoter activity in *C. roseus* hairy roots. For fungal elicitor induction, 0.5–1.5% of yeast extract (YE) was added to transgenic *C. roseus* hairy roots in water and incubated for 12 h at room temperature. Roots incubated in water served as control (Control YE). Following incubation, roots were stained for GUS activity. For JA induction, the transgenic *C. roseus* roots, with or without 100 μM JA treatment for 8 h, were stained for GUS activity. GUS activity is presented as the average color densities of the stained root as quantified by the ImageJ software in an Olympus BX51 microscope with Olympus DP12 digital camera.

[2]. G10H is a cytochrome P450 enzyme and requires Cpr for its function; therefore, we tested the effect of fungal elicitors on the *G10H* promoter activity to investigate the possibility of coordinated regulation of *G10H* and *Cpr* by the same induction signals. After incubation with yeast extract, a significant increase in GUS activity was observed in both transformed *C. roseus* hairy roots (Fig. 6) and transgenic tobacco seedlings (Fig. 7) compared to the controls. Under our experimental conditions the increase of GUS activity correlated with increasing concentrations of yeast extract; GUS activities increased approximately 0.5-, 2- and 3-fold when 0.5%, 1% and 1.5% of yeast extract, respectively, were used to induce the *C. roseus* hairy root cultures (Fig. 6). Presently it is difficult to pinpoint the promoter elements that are responsive to the induction. Although it has been hypothesized that the G-box is associated with the elicitor-responsive element, there is no evidence to suggest it is involved in the response of the *Str* promoter [22]. Fungal elicitor activates the tobacco chitinase gene *CHN48* via a *cis*-element containing W-boxes [32]. In the *C. roseus G10H* promoter the W-box is present adjacent to the G-box element (Fig. 1).

In this paper we report the isolation and characterization of the *C. roseus G10H* promoter. The major TSS maps to an adenine 40 bp upstream of the *G10H* start codon and 31 bp downstream of the TATA box. Many *cis*-elements are found between the -300 position and the TATA box. In other *C. roseus* TIA pathway promoters, the majority of *cis*-elements are located within the ~300-bp regions upstream of the TATA-box [22]. The *G10H* promoter responds to induction by both JA and fungal elicitors. A model to explain the effects of these elicitors on the *Tdc* and *Str* promoters suggests that a putative membrane receptor responding to fungal elicitors signals the increase of JA levels which, in turn, activates the ORCA transcription factors that bind to and regulate the promoters [5,33]. The fungal elicitor also induces the production of repressors such as ZCT proteins which repress the promoter

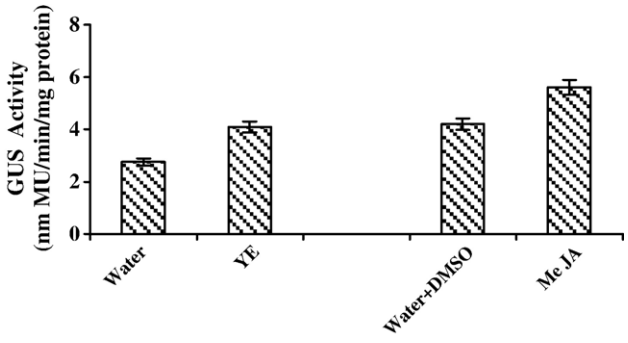

Fig. 7. Effects of fungal elicitor (yeast extract) and methyljasmonate (MeJA) on GUS expression in transgenic tobacco seedlings. For fungal elicitor induction, transgenic tobacco seedlings were incubated in 1.5% yeast extract in water for 24 h at room temperature. Seedlings incubated in water served as control (Control YE). Following incubation, seedlings were assayed for GUS activity. For JA induction, transgenic tobacco seedlings, were treated for 20 h with or without 100 μM JA and assayed for GUS activity.

activities. The simultaneous induction of both activators and repressors by JA and fungal elicitors provide a refined, combinatorial mechanism regulating the TIA pathway. However, ORCA proteins do not activate the expression of *G10H*, therefore it is unlikely that the same transcriptional cascade regulates the *G10H* promoter [3]. The *G10H* promoter contains single WRKY and MYB binding site as well as multiple Dof and GT-1 binding sites (Fig. 1). A number of JA-responsive transcription factors, including those from the AP2-, bHLH- and two-fingered families, have been shown to be involved in regulation of the *C. roseus* TIA pathway [3,5,34]. The *Cpr* promoter is also responsive to the GT-1 transcription factor [21]. At this time, however, the participation of WRKY, MYB, Dof regulatory proteins is not well documented. The possible involvement of these transcription factors supports the notion that the *G10H* regulation utilizes a different transcriptional cascade than those of *Tdc* and *Str*. This prospect hints at additional levels of complexity in the regulation of the TIA pathway and may serve as a guide in our continuing efforts to identify regulatory factors involved biosynthesis of valuable *C. roseus* secondary metabolites.

Acknowledgements

We would like to thank K. Shen for critical reading of the manuscript and thank Drs. J. Kurepa and J. Smalle for technical advice and assistance in microscopic photography. This work is supported by a grant (to LY) from the Kentucky Tobacco Research and Development Center, University of Kentucky.

References

[1] G. Collu, N. Unver, A.M. Peltenburg-Looman, R. van der Heijden, R. Verpoorte, J. Memelink, Geraniol 10-hydroxylase, a cytochrome P450 enzyme involved in terpenoid indole alkaloid biosynthesis, FEBS Lett. 508 (2) (2001) 215–220.

[2] A.H. Meijer, M.I. Lopes Cardoso, J.T. Voskuilen, A. de Waal, R. Verpoorte, J.H. Hoge, Isolation and characterization of a cDNA clone from *Catharanthus roseus* encoding NADPH: cytochrome P-450 reductase, an enzyme essential for reactions catalysed by cytochrome P-450 monooxygenases in plants, Plant J. 4 (1) (1993) 47–60.

[3] L. van der Fits, J. Memelink, ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism, Science 289 (5477) (2000) 295–297.

[4] J. Memelink, R. Verpoorte, J.W. Kijne, ORCAnization of jasmonate-responsive gene expression in alkaloid metabolism, Trends Plant Sci. 6 (5) (2001) 212–219.

[5] B. Pauw, F.A. Hilliou, V.S. Martin, G. Chatel, C.J. de Wolf, A. Champion, M. Pre, B. van Duijn, J.W. Kijne, L. van der Fits, J. Memelink, Zinc finger proteins act as transcriptional repressors of alkaloid biosynthesis genes in *Catharanthus roseus*, J. Biol. Chem. 279 (51) (2004) 52940–52948.

[6] R. Verpoorte, J. Memelink, Engineering secondary metabolite production in plants, Curr. Opin. Biotechnol. 13 (2) (2002) 181–187.

[7] N. Dey, I.B. Maiti, Structure and promoter/leader deletion analysis of mirabilis mosaic virus (MMV) full-length transcript promoter in transgenic plants, Plant Mol. Biol. 40 (5) (1999) 771–782.

[8] S. Pattanaik, C.H. Xie, Q. Kong, K.A. Shen, L. Yuan, Directed evolution of plant basic helix–loop–helix transcription factors for the improvement of transactivational properties, Biochim. Biophys. Acta N, Gene Struct. Expr. 1759 (6) (2006) 308–318.

[9] R.A. Jefferson, T.A. Kavanagh, M.W. Bevan, GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants, EMBO J. 6 (13) (1987) 3901–3907.

[10] C.L. Schardl, A.D. Byrd, G. Benzion, M.A. Altschuler, D.F. Hildebrand, A.G. Hunt, Design and construction of a versatile system for the expression of foreign genes in plants, Gene 61 (1) (1987) 1–11.

[11] S. Pattanaik, N. Dey, S. Bhattacharyya, I.B. Maiti, 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 (3) (2004) 427–438.

[12] P.S. Choi, Y.D. Kim, K.M. Choi, H.J. Chung, D.W. Choi, J.R. Liu, Plant regeneration from hairy-root cultures transformed by infection with *Agrobacterium rhizogenes* in *Catharanthus roseus*, Plant Cell Rep. 22 (11) (2004) 828–831.

[13] T. Murashige, F. Skoog, A revised medium for rapid growth and bioassays with tobacco tissue cultures, Physiol. Plant. 15 (1962) 473–497.

[14] K. Higo, Y. Ugawa, M. Iwamoto, T. Korenaga, Plant cis-acting regulatory DNA elements (PLACE) database: 1999, Nucleic Acids Res. 27 (1) (1999) 297–300.

[15] J. Vicente-Carbajosa, S.P. Moose, R.L. Parsons, R.J. Schmidt, A maize zinc-finger protein binds the prolamin box in zein gene promoters and interacts with the basic leucine zipper transcriptional activator Opaque2, Proc. Natl. Acad. Sci. U. S. A. 94 (14) (1997) 7685–7690.

[16] F. Ishige, M. Takaichi, R. Foster, N.H. Chua, K. Oeda, A G-box motif (GCCACGTGCC) tetramer confers high-level constitutive expression in dicot and monocot plants, Plant J. 18 (4) (1999) 443–448.

[17] T. Eulgem, P.J. Rushton, S. Robatzek, I.E. Somssich, The WRKY superfamily of plant transcription factors, Trends Plant Sci. 5 (5) (2000) 199–206.

[18] N. Datta, A.R. Cashmore, Binding of a pea nuclear-protein to promoters of certain photoregulated genes is modulated by phosphorylation, Plant Cell 1 (11) (1989) 1069–1077.

[19] D.X. Zhou, Regulatory mechanism of plant gene transcription by GT-elements and GT-factors, Trends Plant Sci. 4 (6) (1999) 210–214.

[20] P.B. Ouwerkerk, T.O. Trimborn, F. Hilliou, J. Memelink, Nuclear factors GT-1 and 3AF1 interact with multiple sequences within the promoter of the Tdc gene from Madagascar periwinkle: GT-1 is involved in UV light-induced expression, Mol. Gen. Genet. 261 (4–5) (1999) 610–622.

[21] M.I. Cardoso, A.H. Meijer, S. Rueb, J.A. Machado, J. Memelink, J.H. Hoge, A promoter region that controls basal and elicitor-inducible expression levels of the NADPH:cytochrome P450 reductase gene (Cpr) from *Catharanthus roseus* binds nuclear factor GT-1, Mol. Gen. Genet. 256 (6) (1997) 674–681.

[22] G. Pasquali, A.S. Erven, P.B. Ouwerkerk, F.L. Menke, J. Memelink, The promoter of the strictosidine synthase gene from periwinkle confers elicitor-inducible expression in transgenic tobacco and binds nuclear factors GT-1 and GBF, Plant Mol. Biol. 39 (6) (1999) 1299–1310.

[23] S. Yanagisawa, Dof domain proteins: plant-specific transcription factors associated with diverse phenomena unique to plants, Plant Cell Physiol. 45 (4) (2004) 386–391.

[24] C.P. Joshi, An inspection of the domain between putative TATA box and translation start site in 79 plant genes, Nucleic Acids Res. 15 (16) (1987) 6643–6653.

[25] P.N. Benfey, N.H. Chua, The cauliflower mosaic virus-35s promoter—Combinatorial regulation of transcription in plants, Science 250 (4983) (1990) 959–966.

[26] B. St-Pierre, F.A. Vazquez-Flota, V. De Luca, Multicellular compartmentation of *Catharanthus roseus* alkaloid biosynthesis predicts intercellular translocation of a pathway intermediate, Plant Cell 11 (5) (1999) 887–900.

[27] V. Burlat, A. Oudin, M. Courtois, M. Rideau, B. St-Pierre, Co-expression of three MEP pathway genes and geraniol 10-hydroxylase in internal phloem parenchyma of *Catharanthus roseus* implicates multicellular translocation of intermediates during the biosynthesis of monoterpene indole alkaloids and isoprenoid-derived primary metabolites, Plant J. 38 (1) (2004) 131–141.

[28] J. Murata, V.D. Luca, Localization of tabersonine 16-hydroxylase and 16-OH tabersonine-16-*O*-methyltransferase to leaf epidermal cells defines them as a major site of precursor biosynthesis in the vindoline pathway in *Catharanthus roseus*, Plant J. 44 (4) (2005) 581–594.

[29] F. Guerineau, M. Benjdia, D.X. Zhou, A jasmonate-responsive element within the *A. thaliana* vsp1 promoter, J. Exp. Bot. 54 (385) (2003) 1153–1162.
[30] A.E. Menkens, U. Schindler, A.R. Cashmore, The G-box: a ubiquitous regulatory DNA element in plants bound by the GBF family of bZIP proteins, Trends Biochem. Sci. 20 (12) (1995) 506–510.
[31] R.L. Brown, K. Kazan, K.C. McGrath, D.J. Maclean, J.M. Manners, A role for the GCC-box in jasmonate-mediated activation of the PDF1.2 gene of Arabidopsis, Plant Physiol. 132 (2) (2003) 1020–1032.
[32] S. Yamamoto, T. Nakano, K. Suzuki, H. Shinshi, Elicitor-induced activation of transcription via W box-related cis-acting elements from a basic chitinase gene by WRKY transcription factors in tobacco, Biochim. Biophys. Acta N, Gene Struct. Expr. 1679 (3) (2004) 279–287.
[33] B. Pauw, B. van Duijn, J.W. Kijne, J. Memelink, Activation of the oxidative burst by yeast elicitor in *Catharanthus roseus* cells occurs independently of the activation of genes involved in alkaloid biosynthesis, Plant Mol. Biol. 55 (6) (2004) 797–805.
[34] G. Chatel, G. Montiel, M. Pre, J. Memelink, M. Thiersault, B. Saint-Pierre, P. Doireau, P. Gantet, CrMYC1, a *Catharanthus roseus* elicitor- and jasmonate-responsive bHLH transcription factor that binds the G-box element of the strictosidine synthase gene promoter, J. Exp. Bot. 54 (392) (2003) 2587–2588.