Plant Cell Rep (2011) 30:1173–1182
DOI 10.1007/s00299-011-1025-y

ORIGINAL PAPER

Functional analysis of the DAT gene promoter using transient *Catharanthus roseus* and stable *Nicotiana tabacum* transformation systems

Abdullah Makhzoum · Geneviève Petit-Paly · Benoit St. Pierre · Mark A. Bernards

Received: 20 December 2010 / Revised: 18 January 2011 / Accepted: 24 January 2011 / Published online: 10 February 2011
© Springer-Verlag 2011

Abstract The *Catharanthus roseus DAT* gene encodes the enzyme acetyl-CoA:deacetylvindoline-4-*O*-acetyltransferase involved in the last step of the indole alkaloid pathway leading to vindoline. This gene is characterized by specific cell type expression in idioblasts and laticifers. To understand the specific transcriptional regulation mechanism(s) of *DAT*, several *DAT* promoter *GUS* constructs were cloned into pCAMBIA1305.1. Agroinfiltration of different explant types of *C. roseus* resulted in organ-specific accumulation of GUS, albeit at various levels. Heterologous accumulation of GUS in transgenic tobacco revealed both general and non-specific expression with the exception of a stomata-specific expression when 2.3 kb of the *DAT* promoter was coupled with a portion of the *DAT* ORF. These results suggest that in addition to the 2.3 kb upstream of the *DAT* transcriptional start site, additional *cis*-acting elements may be responsible for the specific spatial expression of *DAT* in vivo. Furthermore, hairy roots transformed with *DAT* promoter *GUS* constructs demonstrated *GUS* expression in root tissues (visualized through GUS enzyme activity), even though *DAT* is repressed in non-transformed roots.

Communicated by J. Register.

A. Makhzoum (✉) · M. A. Bernards
Department of Biology and The Biotron,
The University of Western Ontario,
London, ON N6A 5B7, Canada
e-mail: abmakhzoum@hotmail.com

G. Petit-Paly · B. St. Pierre
Biomolécules et Biotechnologies Végétales,
Universite Francois-Rabelais, 37200 Tours, France

Keywords *Catharanthus* · Indole alkaloids · Vindoline · Acetyl-CoA:deacetylvindoline-4-*O*-acetyltransferase · Promoter analysis · Specific cell type

Introduction

Catharanthus roseus is valuable medicinal plant, as it is a source of indole alkaloids, and represents a potential model for molecular and enzymological studies (Svoboda and Blake 1975). While there have been many attempts at genetic engineering and in vitro culturing to enhance indole alkaloid production in this plant, most have been unsuccessful. This is in part due to (1) the compartmentalization of indole alkaloid biosynthesis in planta, and (2) the lack of a suitable, stable genetic transformation protocol for *C. roseus*.

Transient gene expression systems including protoplast (Galun 1981), biolistic bombardment (Ferrer et al. 2000) or *Agrobacterium*-mediated transient gene expression by agroinfiltration (Batra and Kumar 2003) have been used for functional promoter analysis (Hellens et al. 2005) or to produce recombinant proteins (Joh et al. 2005); agroinfiltration has been applied to *C. roseus* gene expression studies using, for example, vacuum infiltration of tobacco leaves for transient expression of strictosidine synthase (*STR*) (Wang and Li 2002). Vacuum infiltration has been used to study the expression of recombinant indole alkaloids enzymes in *C. roseus* leaves (Di Forie et al. 2004).

The *C. roseus DAT* gene encodes the enzyme acetyl-CoA:deacetylvindoline-4-*O*-acetyltransferase involved in the last step of vindoline biosynthesis. Vindoline production requires the participation of at least two cell types (i.e., idioblasts and laticifers) and the intercellular translocation of a pathway intermediate, and is predominantly found in

leaves, stems and young buds (St-Pierre et al. 1998, 1999). Although light is not required for the formation of idioblast and laticifer cell types (Vazquez-Flota et al. 2000), light has been shown to induce *DAT* transcription, along with other indole alkaloid biosynthesis genes such as deacetoxyvindoline 4-hydroxylase (*D4H*) and tryptophan decarboxylase (*TDC*) (Vazquez-Flota and De Luca 1998; Vazquez-Flota et al. 2000).

Earlier studies describing the functional analysis of *C. roseus* indole alkaloid biosynthetic gene promoters have focused on proximal promoters and related motifs and transcription factors and their role in biotic or abiotic stress responsiveness in suspension culture systems. For example, the basal and proximal promoters of *C. roseus STR* and *TDC* genes have been studied. The *STR* promoter was shown to be regulated by auxins, methyl jasmonate (MJ) and fungal elicitors (van der Fits and Memelink 2000). The *TDC* promoter was shown to contain three domains associated with elicitor response, -160 to -99 (the main binding site) and -99 to -87 and -87 to -37 (two cooperative domains) (Ouwerkerk and Memelink 1999). Meanwhile, the *STR* promoter was shown to be the target for a few positive activating factors, including CrBPF-1 (box P-binding factor) (van der Fits et al. 2000) and ORCA2/ORCA3 (Octadecanoid-Responsive Catharanthus AP2/ERF-domain) and negative activating factors, including ZCT1, ZCT2, and ZCT3 (members of the Cys2/His2-type zinc finger protein family from *C. roseus*) (Pauw et al. 2004) and CrGBF (G-box binding factor) (Memelink and Gantet, 2007). Crgbf1 and Crgbf2 factors were shown to be *STR* repressors by interacting with a B-box upstream of an ORCA interaction-binding site (Siberil et al. 2001). Other analysis demonstrated that *CrMYC1*, encoding a BHLH type transcription factor that binds to a G-box on the *STR* basal promoter is induced by MJ and elicitors (Chatel et al. 2003). The geraniol 10-hydroxylase (*GH*) gene promoter contains many domains between (-191 and -147), (-266 and -188), (-318 and -266) that are the target for many activators (Suttipanta et al. 2007). Recently, analysis of a ca. 1.8 kb fragment of the *DAT* promoter, fused with the reporter gene *GUS* was described (Wang et al. 2010). These authors demonstrated transient expression in *C. roseus* suspension cultures and characterized three TGACG motifs involved in MJ signaling.

In the present work, we report on our functional analysis of the *DAT* promoter, and extend the in silico analysis up to 2.4 kb of the 5′ upstream region as well as into the *DAT* open reading frame (ORF). In doing so, we identify numerous specific regulatory motifs, and use a promoter deletion/promoter-ORF fusion approach to explore the spatial expression of *DAT*, observed through the activity of the reported gene GUS, both transiently in *C. roseus* tissue and through stable transformation of tobacco.

Materials and methods

Plant materials

Seeds of *C. roseus* c.v. Pacifica Pink (Ball Ducrettet, Thonon les Bains, France), were surface sterilized by immersion in ethanol for 2 min followed by immersion in 1.25% Na-hypochlorite for 20 min and finally washed three times in sterile distilled water. Seeds were kept moist in the dark for 1 day, sown in Petri dishes containing hormone-free MS culture medium, and incubated at room temperature in the dark for 3 days. Emerged seedlings were grown under a 16 h light/8 h dark cycle for 6 days, and then leaves were excised or infected by *Agrobacterium*. For experiments with *Nicotiana tabacum* c.v. Petit Havana, foliar discs were obtained from in vitro-grown plants from our laboratory (Elbez et al. 2002)

Bacterial strains and plasmid constructs

Agrobacterium tumefaciens LB4404, AGL1 and *A. rhizogenes* 15834 strains were used to obtain shoots, hairy roots and the transient gene expression system using two plant systems: *C. roseus* and tobacco. All transformation experiments were carried out using electroporation. After each cycle of electroporation with *E. coli* or *Agrobacterium*, individual colonies were selected on appropriate antibiotic (kanamycin 50 mg/L) and used to extract the plasmids harboring the different constructs. The binary vector pCAMBIA1305.1, in which the T-DNA region harbors the reporter gene *GUS Plus*™ encoding β-glucuronidase as well as hygromycin phosphotransferase (*hpt*II) (http://www.cambia.org/daisy/cambia/585.html), was used for all constructs. In the standard vector, these two genes are driven by two 35S promoters. For *DAT* promoter constructs, the 35S promoter was excised from the vector and replaced with appropriate *DAT* promoter fragments in line with the reporter and selection genes.

DAT promoter constructs

For *DAT* promoter constructs, genomic clones of *DAT* were inserted in pBluescript II SK+ (pBSIISK+) and used to amplify fragments of different sizes. These clones, including the pBS::λ#6/SalI fragment (13.9 kb), the pBS#6/E#1′-4 (*EcoRI-EcoRI*) fragment (6 kb), the pBS#6/E#1- 1 (*EcoRI–EcoRI*) fragment (2.8 kb), the pBS#6/E,S#2 (*EcoRI–SalI*) fragment (3.6 kb) of pBS::λ#6/SalI, were all cloned in pBSIISK+.

The pCAMBIA 1305.1 vector was used as a control, as well as a template to create all *DAT* promoter constructs (Fig. 1). The pCAMBIA pDAT 812 construct was realized by replacing the forward p35S::*GUS Plus* promoter with an

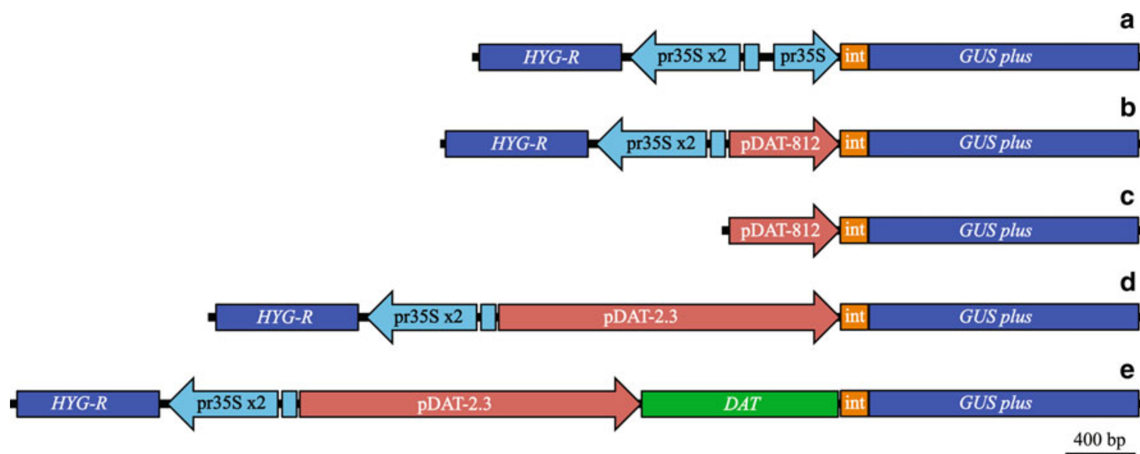

Fig. 1 Schematic representation of DAT gene promoter constructs. The constructs used in this study are derivatives of the pCAMBIA1305.1 vector, which contains both hygromycin resistance (*HYG-R*) and *GUS plus* genes under individual constitutive 35S promoters (pr35S, pr35S x2). **a** Control (p35S) construct. **b** pDAT 812, in which the 35S promoter driving *GUS Plus* has been replaced with an 812 bp fragment of the *DAT* promoter. **c** pDAT 812 Δ35S, represents the same construct as in (**b**), except the 35S(×2) *HYG-R* component of the pCAMBIA1305.1 vector has been deleted. **d** pDAT 2.3, in which the 35S promoter driving *GUS Plus* has been replaced with an 2,431 bp fragment of the *DAT* promoter. **e** pDAT 2.3+*dat* in which the 35S promoter driving *GUS Plus* has been replaced with an 2,431 bp fragment of the *DAT* promoter plus the entire *DAT* ORF

812 bp *DAT* promoter fragment (plus 49 bp downstream of the TIS) upstream of the *GUS Plus* reporter gene. This fragment was amplified by primers PDAT-2 (gcggaattca agctt-AGGTGTTCTTCCCGACG) and PDAT-8 (cagatctac catgg-TCTCTGTCTCAACCGATATT) and the template pBS#6/E # 1-1. *Hin*dIII and *Bgl*II restriction sites were added to PDAT2 and PDAT8 primers, respectively, to facilitate ligation. The construct pCAMBIA pDAT 812 Δ p35S was built from pCAMBIA pDAT 812 by deletion of the CaMV35S X2 (used to drive hygromycin R expression) with *XhoI* and *SalI*. The pCAMBIA pDAT 2.3 construct was obtained by replacing the forward p35S::*GUS Plus* promoter with a 2,480 bp fragment of the *DAT* promoter (−2,431:+49) upstream of the *GUS Plus* reporter gene. This fragment ≈2.3 kb (−2,431:+49 bp) of the 5′ upstream region of the *DAT* gene was amplified from pBS -#6/E # 1-1 by the primers PDAT8 and M13-Reverse 24-mer after adding *Xba*I and *Bgl*II restriction sites (GG AAACAGCTATGACCATGATTAC). pCAMBIA pDAT 2.3 kb+*dat* was constructed by replacing the forward p35S::*GUS Plus* promoter with a 2,480 bp fragment of the *DAT* promoter (-2,431:+49) ligated to the *DAT* ORF, upstream of the *GUS Plus* reporter gene. The *DAT* coding region was amplified from pBSDAT3 (St-Pierre et al. 1998) by the primers DAT20 (5′ GCGagatctACTCTCC AAAACGTTGATCA 3′) and DAT21 (5′ CCGagatctGT ACCTCCATTAGAAACAAATTGAAGTAGC 3′). These primers integrate a *Bgl*II restriction site (AGATCT), as well as a few nucleotides and delete the *DAT* gene stop codon. All constructs were inserted into pCAMBIA1305.1 by ligation of the corresponding fragments and then transformed into *E. coli*, *Agrobacterium tumefaciens* strains LB4404 and AGL1 and *Agrobacterium rhizogenes* 15834 strain. These constructs were verified by sequencing and digestion after transformation.

Agrobacterium tumefaciens containing the various constructs were used for transient gene expression by vacuum agroinfiltration into *C. roseus* and for stable genetic transformation of tobacco while *Agrobacterium rhizogenes* containing pCAMBIA pDAT 2.3+*dat* and pCAMBIA1305.1 (control) was used to give rise to transformed hairy roots from *C. roseus* leaves. In all cases, approximately 10 plates, with each containing 15 explants (leaves or foliar discs) were subjected to transformation. After transformation, approximately 20–30 shoots/roots were kept on selective medium for several cycles to be sure that they were transgenic. Several small leaves or hairy roots (20–30) for used GUS staining.

Transient gene expression in *Catharanthus roseus*

Agroinfiltration was used to examine *DAT* promoter expression in a transient system (Di Forie et al. 2004). For this, *A. tumefaciens* AGL1, harboring the different *DAT* promoter constructs described above, was infiltrated into various *C. roseus* tissues submerged in *A. tumefaciens* cultures ($OD_{600} = 0.8$) under vacuum (PVK-600 vacu box, Suisse). *DAT* promoter constructs used included pCAMBIA pDAT 812, pCAMBIA pDAT 2.3, pCAMBIA pDAT2.3+*dat* and the plasmid pCAMBIA 1305.1 as control. For each construct, 10 negative pressure cycles (10 min each) were applied. After agroinfiltration, explants

were placed on imbibed filter papers in Petri dishes for 3 days at 25°C ± 1°C with long day light period (16 h:8 h) before GUS assay.

Stable transformation of *Nicotiana tabacum*

Tobacco (*Nicotiana tabacum* cv Petit Havana) leaf disks were infected with a disarmed *A. tumefaciens* LB4404 strain harbouring different *DAT* promoter constructs (including, pDAT 812, pDAT 2.3, pDAT 2.3+*dat* and pCAMBIA1305.1) according to Elbez et al. (2002). Foliar discs derived from young leaves were immersed in a suspension culture ($OD_{600} = 0.1$) for 10 min on a shaker (150 rpm). Leaf disks were co-cultured with *A. tumefaciens* on MS solid medium for 2–3 days, after which the disks were washed twice with sterile distilled water, (15 min/wash with shaking) and once with MS medium supplemented with cefotaxime (500 mg/L). Washed explants were placed on MS solid medium containing NAA (0.1 mg/L), BAP (1 mg/L) and cefotaxime (500 mg/L). After 3 weeks, emerging shoots were transferred to new plates (MS medium supplemented with 50 mg/L hygromycin) for selection. Shoots were subcultured every 3 weeks. Only shoots resistant to antibiotic for at least 3 months (3–4 cycles of selection on hygromycin-containing media) were used for subsequent applications.

Formation of hairy roots

Hairy roots of *C. roseus* were obtained as described by Guillon et al. (2006a, b) by wounding in vitro cultured leaves along the main vein with a scalpel previously dipped in solid colonies of *A. rhizogenes* 15834 harboring either pCAMBIA 1305.1 (as control) or pDAT 2.3+*dat*. Up to seven plates, with each plate containing 15 leaves, were inoculated. Once emerged, hairy roots were selected on MS medium containing NAA (0.1 mg/L), BAP (2 mg/L) and hygromycin (50 mg/L). Only hairy roots resistant to antibiotic for at least 3 months (3–4 subcultures of selection on hygromycin containing media) were used for subsequent applications.

Histochemical GUS assay

Explants of agroinfiltrated *C. roseus*, shoots of stably transformed tobacco and hairy roots of *C. roseus* were used for GUS activity measurements, according to Jefferson (1987). Plants were immersed in staining solution: Na-phosphate buffer (50 mM, pH 7.0), 0.5 mM potassium ferricyanide, 0.5 mM potassium ferrocyanide, 1 % Triton X-100, 1 mg/mL 5-bromo-4-chloro-3-indolylglucuronide, and incubated at 37°C in the dark for 1 day. Later, chlorophyll was removed by immersion of the plant material in 70% ethanol. Stained plant material was viewed by light microscope (Olympus BX51) and photographed with a DP50 camera.

Results

In silico analyses of the DAT gene promoter

For our initial in silico analysis, we examined the sequence spanning ca. 2.4 kb upstream of the translational start site of the *DAT* gene, using publically available databases including PlantCARE (Lescot et al. 2002) and PLACE (Higo et al. 1999). Subsequent analysis of putative regulatory elements within the coding sequence of the *DAT* gene was accomplished using PLACE, PlantCARE Transfac (Wingender et al. 2000), softberry (Shahmuradov et al. 2003) and MIRAGE (Ghosh 2000). Such databases allow the prediction of spatial and temporal acting elements and their putative role in response to biotic and abiotic stresses.

For the 5′ upstream region, within the first ca. 1.8 kb, we found many of the motifs described by Wang et al. (2010), and extend this analysis by an additional 600 bp. From the PlantCARE and PLACE databases, common basal promoter elements such as the TATA-box and CAAT-box were readily identified within the 2.4 kb 5′ UTR region analyzed (Fig. 2). In addition, many regulatory motifs associated with plant responses to biotic and abiotic stresses, as well as regulators of spatial specific expression, and expression enhancers were identified. For example, numerous light (G-box, I-box, AE-box, AT1-motif, GT1-motif, ATC-motif, ATCC-motif, ATCT-motif, MRE, LAMP-element, chs-CMA1a, chs-CMA2a, TCT-motif) (Arguello-Astorga and Herrera-Estrella 1998), abscissic acid (ABRE, TCE), MJ (CGTCA, TGACG), and ethylene (ERE) motifs, as well as motifs associated with response to fungal elicitors (Box-W1, ELI-box3), gibberellins (P-box, TATC-box) and other elements involved in response to stress (e.g., temperature, wounding and to drought) were identified (Fig. 2).

Analysis of the coding region of *DAT* using PLACE, Plant CARE, Transfac, Softberry and MIRAGE identified putative motifs involved in stomata specific expression, such as DOFs, in a few positions (Table 1). Other motifs such as WRKY, C1, SEFs and (TGA1a and TGA1b) (Katagiri et al. 1989) were identified in the *DAT* coding region.

Analysis of DAT gene promoter-driven GUS expression

All five DAT gene constructs, including the control (p35S), pDAT 812 (in which the 35S promoter driving *GUS Plus*

-2432 GAATTCACCAAATGGATTAAAATAATATTTCA

-2400 CCATGTGTAAAGTCTCCTAAAAACAAAAAAAAAAAAATCTCTTTCAAGAAAGTAGTTAAAGCACCTTTCCAAAAAAGAAAAGTACTTAAAGAAAAATGAT

HSE

-2300 TAAATTTCAATTTCGTCCTTCAAAATATAGAACTATCTCAATCAAGGACTTTGATTTAAAAAGTTTTCAATTTCATCCTTCAAAATGGAAAAAATTTGTG

-2200 TCCATCAAGTCCTTAACCTGGCTTCTATTAAAGTTTTGACAATCAAACTCCATTCAATCAATTTCATGAGTTGAGAAGTCTTGGTCTTTTTGTCCACAAA

-2100 AGAACCGGTAAAAAAAATAAAAATAAAAGGTTAGAAGACTAACACTTTTGGGATTAAATAAAAAGTCCCTGGCAATGATATCAAAATTTTACACGTGCTA

G-box (light and ABA)

-2000 TCGTAGGCATGTAAAATTAATTCTAATATTCTTCCTAAAAATATACTTAGAGCAAGCAAGTGTCAATCCCACAAAGACGCAATTTAATCCTTTAAGATAA

ATCT-motif (light)

-1900 TGTACTAATCACAAAGTAATGGTTTTAAGGGGTTTGAGATTCATTAAAATAAATAACTAACCAAAATGAACTAAAGTTAACTAAGAAATCAAATTAAACG

-1800 CTTAAAAGAATCAAGGGAAAAAACGTTAGCAATGCACATGCAATTGATTATATTTTTATGGAATCTATCATTCACTACTAGAAAATACCCTTATAATGAC

CGTCA-motif (Me-JA)

-1700 GCAGAAATGTGTCACCATACAACTTTATGGTGACGCTTTGCCTCAGGTGACGCTTCCGAACAGCCTAAAGTCAAGTAGTAATAGACCCAGTAACGCTTTA

CGTCA-motif (Me-JA)

-1600 AAGCGTTACCATACACGATCTATGTATACACTTTAAAGGTCATTAAAAACATATATCTACATTTTTTAGAGTACCAACTAACATTCTAATGGTAAAAAGA

-1500 TTTAAACGTTTTCAAAATTATTATTAC GGTTAAAAGTACCAGCGTAGATGTATGATGATGCTCTAAAAGCGTCACTATAGTGTTAATATGCTTAGAATTG

Box I (light) MBSI (MYB binding site)

GT1-motif (light; GGTTAA)

-1400 CTACTGTAGATTACCTATTACTACGCTTTTAGACTTAGCACTAACATATATATCATTATGCTATTTTATGTCATAAATGCTCAATGCATTAATATATTGC

Box 4 (light)

-1300 TGCGCTATTCAAGTGTCACTATAGGTCACTTCTAAAAAAAATTTGAAAGCAATGTTTATAGTATTTATATATTTCTATACCACCTAAACAGCATTAAAAA

HSE

-1200 AATAAATAATATAGAATAGAATATAAATAAGAGACGTTTAAACTGATATATAAATTAATACTCTAACATCAACAAAAATTATTACGATATCTAAATCCAA

Box 4 (light) Box 4 (light)

-1100 ATAATAAATCACATGACAAAAATAAACAACCACATCCATTTTACTTCCTAGTCTCCCAAAAACTCCATAATCAATCACAAAATCCAAGACATTACTATA

-1000 TAAAAAATCAAGGAATTGTTCAAACTATCAAATTCCAAGTGTTTGTATTGTCCTCTCAAACTCATCATAAGCAAAAAGCTAAGCATCCAACCATAAAACA

-900 TTACCCTCTTCACCTAGAAACTTTCATGAACGGATATTGCTTACAAAATGTGTCTTCAATTAGAACCTTCCAATTCAGAAACATTATAGGTGTATCTTCC

HSE HSE

-800 CACTTATAATATTAGACATAAAAATAAGAAATATATAAATGTTATGAATTTGATGAACGCGTCATAATGGAATATTTAGAATAATAATCCAATATAAGGT

CGTCA-motif (Me-JA)

-700 AGAATAATAAAGTGAAGAGAAACGAAGTAGAGAGAAAGTAAAACTAAAAAATCAAATCTGGATCATATTATTAATATTTT TCGTAAGACATATTTATAGA

Box 4 (light)

-600 TTTAAGATACAATAGATTACAAAACCATAAAAGAAATGTAATTTGTCATTAGAAGGACAAAATCCAATTAGAATTGAACTTTGAGAGTCCTAGTTCCCTT

WUN-motif(wounding) WUN-motif

-500 TAAATAATATATTTGAACATCTATTATGTGTACATTAATTCAATAACATTCCCTTTTTGTATTTATTATGCTAATTTCTTGCTGTGAATGTTCCTTTTTT

P-box:CCTTTT)gibberellins(A-box TAACATTCCCT(ABA) wun-motif ACATTCCCTT(wounding 2 sites)

-400 CTTTATGAGTTTAACAAAGATTTAGATAATTATGAGAGATTTTAAAGTATTTTTTCAATTTTTAACATTGAAATAATTTAGAAATGAAATATAATTATTA

p-box (gibberellins) as-2-box (light) ERE (ethylene) double motifs ERE (ethylene)

-300 GTAATTCTCACACTTTGAAGAGAATATTTTTATTGATTTATGTAAAATTTATTTCCAATACTTTCTCAACTAATATTCATTACATGACTACACATCTACT

CHS-CMA2a (light) ERE (ethylene)

- 200 TTATCTTTGACTTTTAGTTATTATATATATATATATATATATTTAATACT CTTCAAAAGAAAAAACAACTGATTTATGTCATAGGAACTAAAACCAAAAC

MBS(MYB Drought-inducibility) GCN4-motif (endosperm)

-100 CATTCTGGCCACAAAAAAAATGAAAGTATCATAAGAACGTTGGGTGGGATGGGCAACTGGATAGCCTATATATAAGTGTTGTGTGTTGGGGTCTGAATATA

HSE (Heat shock element) MBS (MYB Drought-inducibility)

+1 TAACAGCAAGCAAAATG

Fig. 2 In silico analysis of the 5′ upstream region (ca. 2.4 kb) of the *DAT* gene. Regulatory motifs (*shaded*) were identified using PlantCARE (Lescot et al. 2002) and PLACE (Higo et al. 1999). The transcriptional start site (+1) is based on Wang et al. (2010)

Table 1 Regulatory elements and transcription factor binding sites in the coding region of the *Catharanthus roseus DAT* gene

Motifs	Organism	Position	Description
C1_CS	*Zea mays*	137 and 705	Myb domain; involved in flavonoid biosynthesis
ARE (TGGTTT)	*Zea mays*	157 and 595	*cis*-acting regulatory element essential for anaerobic induction
CCAAT-box (CAACGG)	*Hordeum vulgare*	153 and 723	MYBHv1 binding site
HSE (AGAAAATTCG)	*Brassica oleracea*	215	*cis*-acting element involved in heat stress response
CGTCA-motif (CGTCA)	*Hordeum vulgare*	297 and 619	*cis*-acting element involved in the MeJA-responsiveness
NtBBF1_RS	*Nicotiana tabacum*	335	Dof protein targets the ACT T TA sequence in domain B is of rolB necessary for tissue-specific auxin-regulated expression
Atmyb2-RS	*Arabidopsis thaliana*	386	is a homolog of MYB, was induced by dehydration and binds to a consensus MYB recognition sequence (TAACTG)
MBS (TAACTG)	*Arabidopsis thaliana*	406	MYB binding site involved in drought-inducibility
Dofs_CS	*Zea mays*	457, 745, 1007, 1010, 1048 and 1123	DNA binding with one finger (Dof): family of transcription factors involved in many processes in plants like specific regulation in specific cell types, response to light, Gibberellins, auxins and defence attacks
TCA-element GAGAAGAATA	*Brassica oleracea*	564	*cis*-acting element involved in salicylic acid response
ABRE (TACGTG)	*Arabidopsis thaliana*	576	*cis*-acting element involved in the abscisic acid response
G-box (CCACGTGG)	*Arabidopsis thaliana*	592	Binding site for transcription factors with bZlP motifs
Hex motif (CCACGTCA)	*Triticum aestivum*	599	The HBP-1 Family of Wheat BasicLeucine Zipper Proteins which Interact this Hexamer motif of Plant Histone Genes
W-box	*Arabidopsis thaliana*	738	Site of binding of WRKYs which are transcription factors that are unique to plants, including pathogen defense, senescence and trichome development
BPC1-STK-site-1	*Arabidopsis thaliana*	972	BASIC PENTACYSTEINE1 (BPC1) is a regulator of the homeotic Arabidopsis thaliana gene SEEDSTICK (STK), which controls ovule identity
phyA3-undefined	*Oryza sativa*	975	Regulatory element involved in the repression of phytochrome A
NICE-motif-1	*Sesbania rostrata*	986	nodule-infected cell expression (NICE) element,Is and able to confer nodule-specific expression pattern on this normally root-enhanced promoter in the leghemoglobin glb3 (Srglb3) *Sesbania rostrata*

Identification of putative regulatory elements was achieved by in silico analysis of the full *DAT* open reading frame using PlantCARE, PLACE, Softberry, Transfac and MIRAGE databases

has been replaced with an 812 bp fragment of the *DAT* promoter), pDAT 812 Δ35S (same as pDAT 812 except the 35S::Hygromycin component of the pCAMBIA1305.1 vector has been deleted), pDAT 2.3 (in which the 35S promoter driving *GUS Plus* has been replaced with an 2,431 bp fragment of the *DAT* promoter) and pDAT 2.3+*dat* (in which the 35S promoter driving *GUS Plus* has been replaced with an 2,431 bp fragment of the *DAT* promoter plus the entire *DAT* ORF) (Fig. 1a–e) were verified by re-isolation from *E. coli*, digestion with appropriate restriction enzymes and gel analysis (data not shown). Isolated fragments were also sequenced to verify their identity (data not shown).

Transient expression in seedlings of *C. roseus* after agroinfiltration

Agroinfiltration was used to transiently express the different pDAT::*GUS* constructs in *C. roseus* in order to localize specific *DAT* expression (using GUS enzyme activity as a proxy for gene expression) and potentially identify *cis*-acting regulatory sequences. Agroinfiltration of whole seedlings with the p35S *GUS Plus* construct revealed strong GUS activity levels throughout the plantlet (Fig. 3a). By contrast, whole plantlets agroinfiltrated with pDAT 812 *GUS Plus* (Fig. 3b), pDAT 812 Δ35S *GUS Plus* (data not shown) and pDAT 2.3 *GUS Plus* (Fig. 3c)

constructs showed weaker and more organ specific GUS activity. That is, GUS accumulated in cotyledons and the hypocotyl region (Fig. 3b, c), but not in roots. Inclusion of the *DAT* ORF in *GUS* constructs (pDAT2.3+*dat GUS Plus*) greatly diminished overall GUS accumulation throughout agroinfiltrated plantlets (Fig. 3d). In contrast to stably transformed tobacco leaves (see below), higher magnification of leaves of *C. roseus* transiently expressing pDAT2.3+*dat GUS Plus* constructs did not reveal any apparent tissue specificity (Fig. 3e); indeed, GUS activity was effectively non-detectable in these tissues.

Expression in stably transformed *Nicotiana tabacum*

Due to the difficulty of obtaining stably transformed *C. roseus* whole plants, a tobacco heterologous expression system was used for pDAT *GUS* promoter construct analysis. Tobacco leaves from plants transformed with p35S *GUS Plus* showed strong GUS activity (Fig. 3f). Similarly, transformation with the pDAT812 *GUS Plus* (Fig. 3g) and pDAT2.3 *GUS Plus* (Fig. 3h) constructs resulted in GUS activity in most cell types throughout leaves, albeit with less intensity than observed for the p35S *GUS Plus* construct. By contrast, transformation with the pDAT2.3+*dat GUS Plus* construct resulted in strong, specific GUS activity in stomatal guard cells and vascular tissues cell types (Fig. 3i).

Expression in *C. roseus* hairy roots

The construct pDAT2.3+*dat GUS Plus* and the positive control p35S *GUS Plus* were transformed into *Agrobacterium rhizogenes*, strain15834. Although *DAT* expression was restricted to specific cells types found in leaves and young buds (St-Pierre et al. 1999), it has also been expressed in *C. roseus* hairy roots under the regulation of the 35S promoter (Magnotta et al. 2007). Herein, we succeeded in expressing the *GUS* reporter gene under the control of the *DAT* promoter, in a *C. roseus* hairy root system. First, we expressed *GUS* in *C. roseus* hairy roots derived from *A. rhizogenes* transformed with our p35S::*GUS* construct, and found relatively strong GUS activity in all tissues, but especially the vascular tissue and meristem regions (Fig. 3j). By contrast, expression of the reporter gene (as revealed by GUS activity) in *C. roseus* hairy roots derived from *A. rhizogenes* transformed with our pDAT 2.3+*dat GUS Plus* construct was restricted to the vasculature (Fig. 3k). As with the transient expression in *C. roseus* plantlets and stable expression in *N. tabacum*, GUS activity was weak in the tissue transformed with the pDAT 2.3+*dat GUS Plus* construct.

Discussion

The expression of genes is normally thought to be regulated by the untranslated sequence upstream of the ORF;

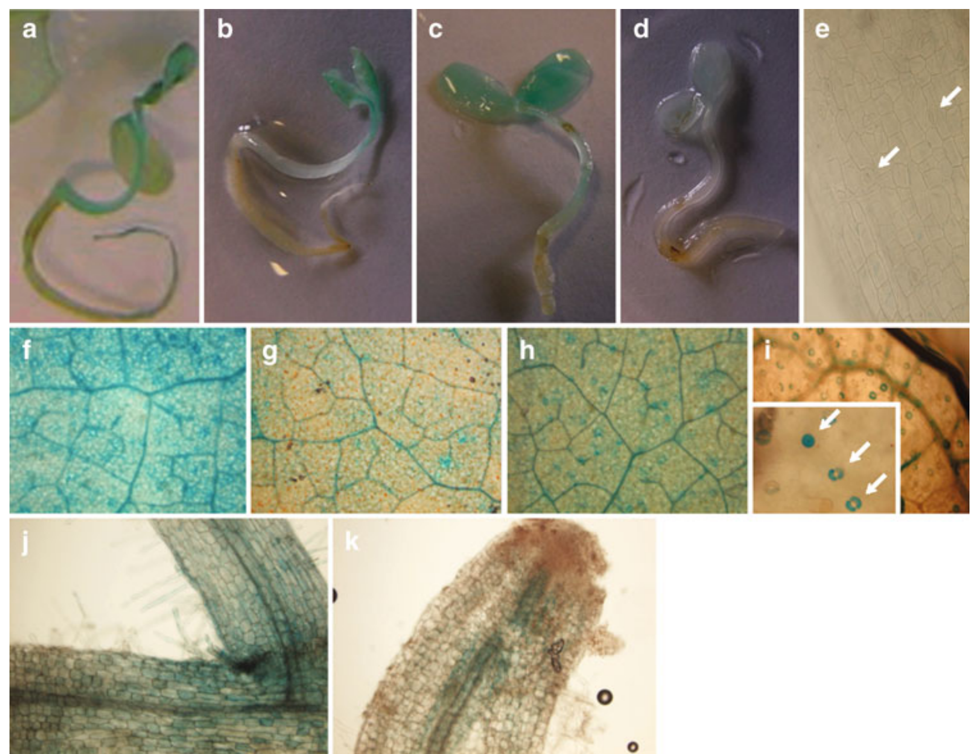

Fig. 3 β-Glucuronidase (GUS) activity in *Catharanthus roseus* and *Nicotiana tabacum*: **a–e** Agroinfiltrated seedlings of *Catharanthus roseus* harbouring different pDAT constructs: **a** p35S, **b** pDAT 812, **c** pDAT 2.3, **d**, **e** pDAT 2.3+*dat*. **f–i** Leaves of *Nicotiana tabacum* stably transformed with **f** p35S, **g** pDAT 812, **h** pDAT 2.3, **i** pDAT 2.3+*dat.* **j**, **k** Hairy roots of *Catharanthus roseus* harbouring different pDAT constructs: **j** p35S, **k** pDAT 2.3+*dat*. *Arrows* in (**e**) and (**i**, inset) point to guard cells. **f–k** ×100, **e**, **i** (*inset*), ×200

however, gene expression can also be affected by numerous other elements, which can be located in the upstream (i.e., promoter) region, coding region (including intron and exon elements), as well as in the 5′UTR and 3′UTR (see, for example, Curie and McCormick 1997; Buzeli et al. 2002; Ito et al. 2003; Menossi et al. 2003; Fiume et al. 2004; Gowik et al. 2004; Reddy and Reddy 2004; Ng et al. 2006). The *DAT* gene has been shown experimentally to be responsive to light and MJ (St-Pierre et al. 1998; Hernandez-Dominguez et al. 2004; St-Pierre and De Luca 1995; St-Pierre et al. 1998; Vazquez-Flota and De Luca 1998). Consistent with this, Wang et al. (2010) characterized three TGACG motifs involved in MJ signaling in the first 1.8 kb of the *DAT* promoter, while our in silico analysis of the *DAT* promoter region also identified numerous motifs known to be light responsive elements. We also found many other motifs known to be involved in the acclimation and/or response to biotic and abiotic stresses. These factors participate in the regulation of many processes such as the synthesis of storage proteins, regulation of the light response of carbohydrate metabolism genes, several defense mechanisms and gibberellin and auxin responses (Yanagisawa and Schmidt 1999; Plesch et al. 2001). However, none of these imparted any specific tissue expression in either transient or stable gene transformation systems. Indeed, it was not until the *DAT* ORF was incorporated into the reporter gene construct that tissue specificity (i.e., guard cell specific expression) was observed. At this stage, it is difficult to pinpoint exactly which elements within the ORF are responsible for the explicit guard cell-specific expression observed, but DOF (DNA binding with one finger) elements at 457, 745, 1007, 1010, 1048 and 1123 bp within the ORF are likely involved.

The *DAT* construct pDAT 2.3+*dat GUS Plus* imparted stomatal guard cell specific expression of *GUS* (viewed by proxy through GUS enzyme activity) in stably transformed tobacco leaves and predominantly vascular expression in stably transformed *C. roseus* hairy roots. By contrast, little tissue specificity of *DAT* expression was observed in constructs that did not contain the *DAT* ORF. The *DAT* gene is normally expressed in idioblast cells in *C. roseus* (St-Pierre et al. 1998, 1999), however, neither tobacco leaves, nor *C. rosues* roots contain idioblasts. Nevertheless, a similar stomatal guard cell specific expression was reported for other promoters associated with idioblast-specific gene expression [e.g., TGG1 (2.5 kb) and *Brassica napus* Mir1 Bn1 (2.9 kb)] expressed in *Arabidopsis thaliana* (Thangstad et al. (2004)). We speculate that the stomatal expression of *DAT* was governed in part by elements in the *DAT* ORF. Specific gene expression driven by elements within the ORF of the corresponding gene in other related cell types have been described for *A. thaliana.* For example, the *OASA1* promoter drives reporter gene expression in glandular trichomes in *Arabidopsis*, provided domains between -266 and -66 *and* +112 and +375 are present. When expressed in a heterologous system (e.g., tobacco, peppermint) expression remained restricted to non-glandular trichomes as long as the ORF domain was present (Gutierrez-Alcala et al. 2005).

In previous work, the *DAT* coding region was expressed in a hairy root system, albeit under the control of a constitutive (35S) promoter (Magnotta et al. 2007). In the present work, we demonstrated the expression of the *DAT* coding region (through GUS enzyme activity) under the control of its own promoter in both *C. roseus* hairy roots and tobacco leaves. From this, it was clear that the coding region of the *DAT* gene, in conjunction with elements in the 5′ upstream region, play a role in its cell/tissue specific expression. However, in order to fully elucidate which elements or motifs (e.g., *cis* acting elements and their trans-acting factors) are responsible for the spatiotemporal expression of the *DAT* gene under different environmental and physiological conditions, additional, shorter constructs, derived from the pCAMBIA pDAT 2 3+*dat* construct must be created and tested. Additionally, the role of DOFs located within the *DAT* coding region need to be elucidated by gel shift analysis, footprinting and chromatin immunoprecipitation.

Acknowledgments The authors gratefully acknowledge financial support from the Department of Plant Physiology, Tours University (BSt-P) and the Natural Sciences and Engineering Research Council (NSERC) of Canada (MAB). Additional financial support to ABM, from the Ministry of Higher Education in Syria is gratefully acknowledged.

References

Arguello-Astorga G, Herrera-Estrella L (1998) Evolution of light-regulated plant promoters. Ann Rev Plant Physiol Plant Mol Biol 49:525–555

Batra S, Kumar S (2003) Agrobacterium-mediated transient GUS gene expression in buffel grass (*Cenchrus ciliaris* L.). J App Genet 44:449–458

Buzeli RA, Cascardo JC, Rodrigues LA, Andrade MO, Almeida RS, Loureiro ME, Otoni WC, Fontes EP (2002) Tissue-specific regulation of BiP genes: a *cis*-acting regulatory domain is required for BiP promoter activity in plant meristems. Plant Mol Biol 50:757–771

Chatel G, Montiel G, Pre M, Memelink J, Thiersault M, Saint-Pierre B, Doireau P, Gantet P (2003) 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 54:2587–2588

Curie C, McCormick S (1997) A strong inhibitor of gene expression in the 5′ untranslated region of the pollen-specific LAT59 gene to tomato. Plant Cell 9:2025–2036

Di Forie S, Hopmann V, Fischer R, Schillberg S (2004) Transient gene expression of recombinant terpenoid indole alkaloid enzymes in *Catharanthus roseus* leaves. Plant Mol Biol Report 22:15–22

Elbez M, Kevers C, Hamdi S, Rideau M, Petit-Paly G (2002) Les protéines de pathogenèse PR-10 des végétaux. Acta Bot Gallica 149:415–444

Ferrer E, Linares C, Gonzalez JM (2000) Efficient transient expression of the b-glucuronidase reporter gene in garlic (*Allium sativum* L.). Agronomie 20:869–874

Fiume E, Christou P, Giani S, Breviario D (2004) Introns are key regulatory elements of rice tubulin expression. Planta 218:693–703

Galun E (1981) Plant protoplasts as physiological tools. Ann Rev Plant Physiol 32:237–266

Ghosh D (2000) Object-oriented transcription factors database (ooTFD). Nucl Acids Res 1(28):308–310

Gowik U, Burscheidt J, Akyildiz M, Schlue U, Koczor M, Streubel M, Westhoff P (2004) *cis*-Regulatory elements for mesophyll-specific gene expression in the C4 plant *Flaveria trinervia*, the promoter of the C4 phosphoenolpyruvate carboxylase gene. The Plant Cell 16:1077–1090

Guillon S, Tremouillaux-Guiller J, Pati PK, Rideau M, Gantet P (2006a) Hairy root research: recent scenario and exciting prospects. Curr Opin Plant Biol 9:341–346 (style)

Guillon S, Tremouillaux-Guiller J, Pati PK, Rideau M, Gantet P (2006b) Harnessing the potential of hairy roots: dawn of a new era. Trends Biotechnol 24:403–409

Gutierrez-Alcala G, Calo L, Gros F, Caissard JC, Gotor C, Romero LC (2005) A versatile promoter for the expression of proteins in glandular and non-glandular trichomes from a variety of plants. J Exp Bot 56:2487–2494

Hellens RP, Allan AC, Friel EN, Bolitho K, Grafton K, Templeton MD, Karunairetnam S, Gleave AP, Laing WA (2005) Transient expression vectors for functional genomics, quantification of promoter activity and RNA silencing in plants. Plant Methods 1:13

Hernandez-Dominguez E, Campos-Tamayo F, Vazquez-Flota F (2004) Vindoline synthesis in in vitro shoot cultures of *Catharanthus roseus*. Biotechnol Lett 26:671–674

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

Ito T, Sakai H, Meyerowitz EM (2003) Whorl-specific expression of the SUPERMAN gene of *Arabidopsis* is mediated by *cis* elements in the transcribed region. Curr Biol 13(17):1524–1530

Joh LD, Wroblewski T, Ewing NN, VanderGheynst JS (2005) High-level transient expression of recombinant protein in lettuce. Biotechnol Bioeng 91:861–871

Jefferson R (1987) Assaying chimeric genes in plants: the GUS gene fusion system. Plant Mol Bio Rep 5:387–405

Katagiri F, Lam E, Chua NH (1989) Two tobacco DNA-binding proteins with homology to the nuclear factor CREB. Nature 340:727–730

Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouze P, Rombauts S (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucl Acids Res 30:325–327

Magnotta M, Murata J, Chen J, De Luca V (2007) Expression of deacetylvindoline-4-*O*-acetyltransferase in *Catharanthus roseus* hairy roots. Phytochemistry 68:1922–1931

Memelink J, Gantet P (2007) Transcription factors involved in terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. Phytochem Rev 6:353–362

Menossi M, Rabaneda F, Puigdomènech P, Martínez-Izquierdo JA (2003) Analysis of regulatory elements of the promoter and the 3′ untranslated region of the maize *Hrgp* gene coding for a cell wall protein. Plant Cell Rep 21:916–923

Ng DW, Chandrasekharan MB, Hall TC (2006) Ordered histone modifications are associated with transcriptional poising and activation of the phaseolin promoter. The Plant Cell 18:119–132

Ouwerkerk PBF, Memelink J (1999) A G-box element from the *Catharanthus roseus* strictosidine synthase (*Str*) gene promoter confers seed-specific expression in transgenic tobacco plants. Mol Gen Genet 261:635–643

Pauw B, Hilliou FAO, Martin VS, Chatel G, de Wolf CJF, Champion A, Pre M, van Duijn B, Kijne JW, van der Fits L, Memelink J (2004) Zinc finger proteins act as transcriptional repressors of alkaloid biosynthesis genes in *Catharanthus roseus*. J Biol Chem 279:52940–52948

Plesch G, Ehrhardt T, Mueller-Roeber B (2001) Involvement of TAAAG elements suggests a role for Dof transcription factors in guard cell-specific gene expression. Plant J 28:455–464

Reddy VS, Reddy AS (2004) Developmental and cell-specific expression of ZWICHEL is regulated by the intron and exon sequences of its gene. Plant Mol Biol 54:273–293

Shahmuradov IA, Gammerman AJ, Hancock JM, Bramley PM, Solovyev VV (2003) PlantProm: a database of plant promoter sequences. Nucl Acids Res 31:114–117

Siberil Y, Benhamron S, Memelink J, Giglioli-Guivarc'h N, Thiersault M, Boisson B, Doireau P, Gantet P (2001) *Catharanthus roseus* G-box binding factors 1 and 2 act as repressors of strictosidine synthase gene expression in cell cultures. Plant Mol Biol 45:477–488

St-Pierre B, De Luca V (1995) A Cytochrome P-450 Monooxygenase Catalyzes the First Step in the Conversion of Tabersonine to Vindoline in *Catharanthus roseus*. Plant Physiol 109:131–139

St-Pierre B, Laflamme P, Alarco A, Vincenzo MD, Luca E (1998) The terminal Oacetyltransferase involved in vindoline biosynthesis defines a new class of proteins responsible for coenzyme A dependent acyl transfer. Plant J 14:703–713

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

Suttipanta N, Pattanaik S, Gunjan S, Xie CH, Littleton J, Yuan L (2007) Promoter analysis of the *Catharanthus roseus* geraniol 10-hydroxylase gene involved in terpenoid indole alkaloid biosynthesis. Biochim Biophys Acta 1769:139–148

Svoboda GH, Blake DA (1975) The phytochemistry and pharmacology of *Catharanthus roseus* (L.) G. Don. In: Taylor WI, Farnsworth NR (eds) The *Catharanthus* alkaloids: botany, chemistry, pharmacology and clinical uses

Thangstad OP, Gilde B, Chadchawan S, Seem M, Husebye H, Bradley D, Bones AM (2004) Cell specific, cross-species expression of myrosinases in *Brassica napus*, *Arabidopsis thaliana* and *Nicotiana tabacum*. Plant Mol Biol 54:597–611

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

Vazquez-Flota FA, De Luca V (1998) Developmental and light regulation of desacetoxyvindoline 4-hydroxylase in *Catharanthus roseus* (L.) G. Don. Evidence of a multilevel regulatory mechanism. Plant Physiol 117:1351–1361

Vazquez-Flota FA, St-Pierre B, De Luca V (2000) Light activation of vindoline biosynthesis does not require cytomorphogenesis in *Catharanthus roseus* seedlings. Phytochem 55:531–536

Wang M, Li QR (2002) Transient expression of strictosidine synthase in tobacco leaves by vacuum infiltration. Sheng wu hua xue yu sheng wu wu li xue bao Acta biochim biophys Sin 34:703–706

Wang Q, Yuan F, Pan Q, Li M, Wang G, Zhao J, Tang K (2010) Isolation and functional analysis of the *Catharanthus roseus* deacetylvindoline-4-*O*-acetyltransferase gene promoter. Plant Cell Rep 29:185–192

Wingender E, Chen X, Hehl R, Karas H, Liebich I, Matys V, Meinhardt T, Prüß M, Reuter I, Schacherer F (2000) TRANSFAC: an integrated system for gene expression regulation. Nucl Acids Res 28:316–319

Yanagisawa S, Schmidt RJ (1999) Diversity and similarity among recognition sequences of Dof transcription factors. Plant J 17:209–214