UDC 577.21;579.23'315

CLONING AND CHARACTERIZATION OF TRICHOME-SPECIFIC PROMOTER OF *cpr71av1* GENE INVOLVED IN ARTEMISININ BIOSYNTHESIS IN *Artemisia annua* L.

© 2011   Yueyue Wang[1#], Ke Yang[1#], Fuyuan Jing[1#], Meiya Li[1], Ting Deng[1], Runze Huang [1],
Boshi Wang[1], Guofeng Wang[1], Xiaofen Sun[2], Ke-Xuan Tang[1*]

[1]*Plant Biotechnology Research Center, School of Agriculture and Biology, Fudan-SJTU-Nottingham Plant Biotechnology R&D Center, Shanghai Jiao Tong University, Shanghai 200240, China*

[2]*State Key Laboratory of Genetic Engineering, School of Life Sciences, Morgan-Tan International Center for Life Sciences, Fudan-SJTU-Nottingham Plant Biotechnology R&D Center, Fudan University, Shanghai 200433, China*

Artemisinin, a sesquiterpene lactone endoperoxide derived from ***Artemisia annua*** **L. (Asteraceae), is the most effective antimalarial drug. We used two methods: genome walking and thermal asymmetric interlaced polymerase chain reaction, to isolate the unknown 5'-flanking sequence of the *cyp71av1* gene. The subsequent sequence analysis using bioinformatics software revealed that there are several *cis*-acting elements inside the *cyp71av1* promoter. The 5'-rapid amplification of the cDNA ends method was used to determine the transcription start site of the *cyp71av1* gene. We then mapped it at the 18 base upstream of the ATG initiation codon. For simple functional characterization, we built fusion vectors between the 5'-deletion promoter and the *gus* reporter gene. The expression levels of the transferred vectors into *A. annua* L. were analyzed by the transient expression way. The β-glucuronidase assay results indicated that deletion of the region to −1551 bp did not lead to much damage in the GUS activity, whereas further deletion, to −1155 bp, resulted in a 5.5-fold reduction of GUS activity. In stabilized transgenic *A. annua* L. seedlings we observed that GUS expression was restricted to trichomes, which means that the promoter of the *cyp71av1* gene is trichome-specific. Compared with the constitutive CaMV 35S promoter, which can express genes throughout the plant, influence on the trichome system through the trichome-specific expression promoter merely imperils plant growth. In addition, the promoter of the *cyp71av1* gene contains several binding sites for transcription factors, which implies that the *cyp71av1* promoter responds to more than one form of stimulation.**

Keywords: **promoter, genome walking, thermal asymmetric interlaced polymerase chain reaction, 5'-rapid amplification of the cDNA ends*, cyp71av1* **gene,** *cis***-acting elements,** *Artemisia annua* **L.**

Malaria, one of the most serious infectious diseases, results in almost 1 million deaths in 107 countries and regions every year [1]. In the 1970s, researchers from China separated and purified an anti-malaria compound, artemisinin, from *Artemisia annua* L. Artemisinin was recorded in *The International Pharmacopoeia*, published by the World Health Organization, no later than 1995 [2]. With the peroxide lactone group in its structure, artemisinin is effective against anti-chloroquine and cerebral malaria. In contrast to traditional anti-malarial drugs, such as quinine, artemisinin acts quickly and causes minimal adverse side effects. Presently, the curative effects of artemisinin are widely recognized, and the artemisinin family is known for producing the most effective anti-malarial drugs. The World Health Organization recommends artemisinin-based combination therapies (ACTs) as the first choice for medical treatment of malaria [3]. The artemisinin industry is expanding to meet the huge worldwide demand for this drug.

At present, artemisinin comes mainly from harvesting of the above-ground part of *A. annua* L., which contains the compound at very low level (0.01−1% DW). Synthesis of artemisinin by partial chemical methods has previously been reported; however, all published syntheses require complex steps and result in yields of less than 10%. It is more practical and reliable to increase the artemisinin yield through plant genetic and microbial engineering or through a mixed system of the two methods, in which biosynthesis provides precursors for chemical synthesis.

The precursor of artemisinin, isopentenyl pyrophosphate, produced from both mevalonic acid and methylerythritol phosphate pathways, can react with geranyl pyrophosphate to form farnesyl pyrophosphate [4, 5]. The transformation from farnesyl pyrophosphate to amorpha-4,11-diene is catalyzed by amorpha-4,11-diene syn-

Table 1. Primers for TAIL-PCR

Primer	Primer sequence 5'-3'
CYPGSP1	GGAAGGCTTTTTTTGGTGGATTTG
CYPGSP2	CGAGTAGCGAACTTGTAAACGAAC
CYPGSP3	CAAGAGCAATGGAAGTGGTCAGT
AD1	TGWGNAGSANCASAGA
AD2	AGWGNAGWANCAWAGG
AD3	NTCGASTWTSGWGTT
AD4	WGTGNAGWANCANAGA

thase. In 2006, the cytochrome-coding gene, *cyp71av1*, was isolated from specific *A. annua* L. trichome cDNA suppression subtractive hybridization library. CYP71AV1 catalyzes the oxidation reactions of amorpha-4,11-diene, artemisinic alcohol, and artemisinic aldehyde [6]. Transferring *cyp71av1* and *cpr* along with some upstream genes in the artemisinin biosynthesis path into yeast, allowed yeast successfully produce artemisinic acid [7], which further proved that CYP71AV1 actually catalyzes the three oxidation reactions from amorpha-4,11-diene to artemisinic acid. It is still not clear if a similarity, such as hormone responsiveness, exists among the enzymes of the artemisinin biosynthesis pathway. Previous experiments used the real-time polymerase chain reaction (PCR) method to investigate the changes in expression of reporter genes upon stimulation with different hormones. The results suggested that the *cyp71av1* gene was greatly upregulated after being induced with abscisic acid (ABA), while the other genes were not [8]. These observations implied that *cyp71av1* promoter may contain ABA-responsive elements.

Here we report the isolation and characterization of the *A. annua* L. *cyp71av1* gene promoter. Our results show that the promoter contains a wide group of *cis*-act-

ing elements, and is trichome specific. These findings may help to increase the artemisinin output and to investigate the role of this gene in other contexts. In addition, this study also provides a trichome-specific promoter for plant transgenic engineering, which will not imperil normal plant growth by expressing throughout the plant.

EXPERIMENTAL

TAIL-PCR and *cyp71av1* 5′-flanking sequence isolation. The *cyp71av1* 5′-flanking sequence was isolated using thermal asymmetric interlaced polymerase chain reaction (TAIL-PCR). We designed three gene-specific primers and four arbitrary degenerated (AD) primers (table 1). The primary PCR was carried out with CYPGSP1 as the reverse primer and the four ADs as forward primers. The second PCR was carried out using the 50-fold diluted primary PCR products as a template, with CYPGSP2 as the reverse primer and the four ADs as forward primers. The final PCR was carried out using the 20-fold diluted secondary PCR product as the template, with CYPGSP3 as reverse primer, and four ADs as forward primers. The full reaction programs are listed in table 2. The third-step-TAIL-PCR products were gel purified and cloned into the pMD18-T vector ("TaKaRa", Japan) for sequencing.

Further isolation of the flanking sequence by genome walking. Since the 5' flanking sequence isolated by TAIL-PCR was not long enough, we used the GenomeWalker™ Universal Kit ("Clontech", USA) to obtain longer sequences. The available genomic DNA was isolated from *A. annua* L. by the cetyltrimethylammonium bromide (CTAB) method. Three blunt end restriction enzymes, DraI, EcoRV, and StuI, were used to digest the genomic DNA. Each digestion product was purified and ligated to the GenomeWalker adapter sequence. Nested PCR (Advantage II Kit, "Clontech") was used to improve production specificity. The primary PCR was car-

Table 2. PCR programs for TAIL-PCR

Reaction	Thermal condition	No. of cycles
Primary	94°C (3 min)	1
	94°C (30 s), 62°C (1 min), 72°C (2 min 30 s)	4
	94°C (30 s), 25°C (3 min), ramp to 72°C at 0.2°C/s, 72°C (2 min 30 s)	1
	94°C (10 s), 68°C (1 min), 72°C (2 min 30 s); 94°C (10 s) 68°C (1 min), 72°C (2 min 30 s); 94°C (10 s), 44°C (1 min), 72°C (2 min 30 s)	14
	72°C (5 min)	1
Secondary	94°C (3 min)	1
	94°C (10 s), 64°C (1 min), 72°C (2 min); 94°C (10 s), 64°C (1 min), 72°C (2 min); 94°C (10 s), 44°C (1 min), 72°C (2 min);	11
	72°C (5 min)	1
Tertiary	94°C (3 min)	1
	94°C (10 s), 44°C (1 min), 72°C (2 min)	20
	72°C (5 min)	1

Table 3. Primers for the Genome Walker™ Universal Kit

Primer	Primer sequence 5'-3'
AP1	GTAATACGACTCACTATAGGGC
AP2	ACTATAGGGCACGCGTGGT
CYPPR1	CCATCTTCGTTGGTCTTCAGTCTTCACAG
CYPPR2	CAAGCCACGCAAAACACAGAGAAAAT

Table 4. Primers used for amplifying different fragments of the *cyp71v1* promoter

Primer	Primer sequence 5'-3'
DPR(+31)	CTAGCCATGGGAGCAATGGAAGTGGTCAGT
DEL1(−1551)	CGGAATTCTTGGGTTTTGGTTACCGACA
DEL2(−1155)	CGGAATTCTCGTAAAAAGCAGAAGGTCA
DEL3(−727)	CGGAATTCAAAGTCAAGAGTGGACTGCG
DEL4(−138)	CGGAATTCTGACCGATTATGACCAAGGT

ried out using the gene-specific primer CYPPR1 and the adapter-specific primer AP1. In the second PCR the nested gene-specific primer CYPPR2 and the adapter-specific primer AP2 were used (table 3). The targeted sequence was gel purified and cloned into the pMD18-T vector ("TaKaRa") for sequencing.

Determination of the *cyp71av1* gene transcription start site by 5'-RACE. 5'-Rapid Amplification of cDNA Ends (5'-RACE) was used to determine the transcription start site (TSS) of the *cyp71av1* gene. Total RNA was extracted from *A. annua* L. seedlings using the RNA plant full RNA Isolation Kit ("Tiangen", Beijing, China). 5'-RACE was carried out using the SMART™ RACE cDNA Amplification Kit ("Clontech"). The Universal Primer A Mix (long primer: 5'-CTAATACGACT-CACTATAGGGCAAGCAGTGGTATCAA-CGCA-GAGT-3' and short primer: 5'-CTAATACGACTCAC-TA-TAGGGC-3') and two gene-specific primers (GSP1: 5'-CTG-GTCCTTGATCCCTTTCCCAAA-TG-3' and GSP2: 5'-TCGGA-GATGACACCACGA-TTGTTG-3') were used for subsequent PCR amplification (Advantage II Kit, "Clontech"). The 5'-RACE products were gel purified and ligated into the pMD18-T vector for sequencing.

Fusion of the 5'-deletion promoter and the *gus* reporter gene. Four 5'-deletion promoters (−1551 nt to +31 nt, −1155 nt to +31 nt, −727 nt to +31 nt, and −138 nt to +31 nt) were joined with the *gus* reporter gene. Five forward primers with EcoRI digestion sites (DEL1−DEL4) and one reverse primer DPR with an NcoI digestion site were designed (table 4). Fragments of different lengths from the *cyp71av1* promoter were PCR amplified and digested by EcoRI and NcoI. These fragments were gel purified and inserted into pCAMBIA1301 vectors to generate pCAMBIA *CYP (DEL1−DEL4)*. Each vector was verified by sequencing. A freeze-thawing method was used to transfer the newly built pCAMBIA *CYP* vectors

into *Agrobacterium tumefaciens* EHA105 for further transformation into plant cells.

Transient transformation of Artemisia leaves. After the introduction of constructed expression vectors pCAMBIA CYP (*DEL1-DEL4*) into competent *A. tumefaciens* EHA105, positive clones were selected by PCR and kept for transformation. The clones were cultivated in liquid LB medium with kanamycin, rifampicin and streptomycin (triple antibiotic combination) till their OD_{600} reached 0.8, then 200 ul of bacteria liquid was transferred into 200 ml of new liquid LB medium (triple antibiotic combination) till their OD_{600} reached 2.0. After that, bacteria was centrifuged at 4000 rpm and suspended in 400 ml liquid MS medium (pH 5.8, 200 µM AS). Artemisia leaves 30 days of age were immersed into the bacteria liquid under vacuum condition at 0.007 MPa for 20 min. Solid MS_0 (200 µM AS) was used for infected leaves co-culture. The co-culture was set inside dark condition for 5 days.

Stable transformation of Artemisia leaves. After the introduction of constructed expression vectors pCAMBIA CYP−*DEL1* into competent *A. tumefaciens* EHA105, positive clones were selected by PCR and kept for transformation. According to the methods of Vergauwe [9], germinated seedlings were collected when they had grown to 5 cm in length; the leaves were cut into 0.5 cm diameter segments and used as the explants in the cocultivation medium (1/2 MS + 100 µM AS). The 1/2 MS suspension, with the activated *A. tumefaciens* − engineered bacteria, was dripped down until it satisfactorily covered the explants, which were then cultured at 28°C in the dark for 3 days. After kanamycin selection in the selective shooting medium MS_1 (MS_0 + 0.5 mg/l 6-BA + 0.05 mg/l NAA + 50 mg/l kanamycin + 500 mg/l carbenicillin) with a photoperiod of 16 h light/8 h dark at 25°C, the kanamycin-resistant plantlets were subcultured 2 or 3 times every two weeks.
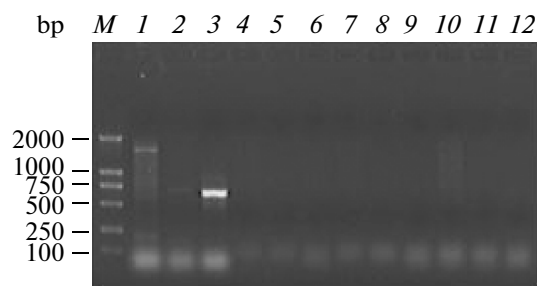
Fig. 1. Isolation of the 5'-flanking region of the *cyp71av1* gene. *M* – DL2000 ("Takara", Japan). *1–3* – Products of the primary, secondary, and tertiary TAIL-PCR using the AD1 primer; *4–6* – products of the primary, secondary, and tertiary TAIL-PCR using the AD2 primer; *7–9* – products of the primary, secondary, and tertiary TAIL-PCR using the AD3 primer; *10–12* – products of the primary, secondary, and tertiary TAIL-PCR using the AD4 primer.
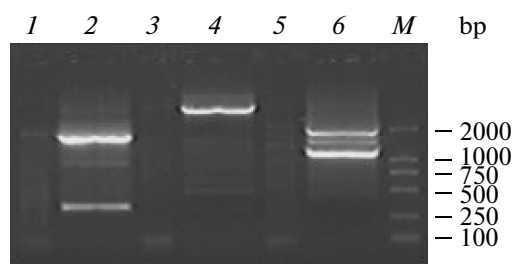

Fig. 2. Products of the primary and secondary PCR using genomic DNA digested with DraI (*1*, *2*); StuI (*3*, *4*); EcoRV (*5*, *6*); *M* – DL2000.
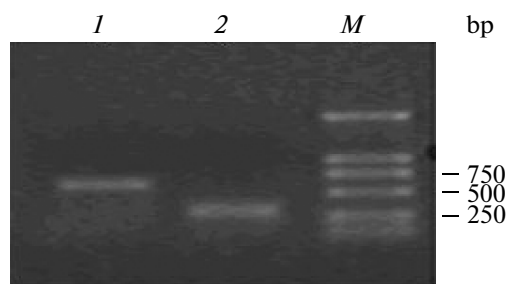

Fig. 3. PCR with GSP1 (*1*) and with GSP2 (*2*). *M* – DL2000.

They were then transferred into the rooting medium MS_2 (1/2 MS_0 + 125 mg/l carbenicillin) in order to obtain the kanamycin-resistant regenerated plantlets.

β-Glucuronidase assay. Using the method of Jefferson [10], the leaves were immersed into the GUS-strain solution and then incubated at 37°C. The reaction was stopped after 2 days; the leaves were incubated with 70% ethanol at 37°C till the uncolored parts of leaves became pale. Five random GUS-positive leaves were picked up from each plantlet, with a different length of

the 5'-deletion promoter, as material for the β-glucuronidase assay. The procedures were done following the Cote's method [11].

RESULTS AND DISCUSSION

Clone of cyp71av1 5'-flanking sequence

Four ADs were respectively amplified through the three step PCR with three gene-specific primers. Only AD1 (table 1) amplified the specific strap of 650 bp in gel electrophoresis (fig. 1). Sequencing revealed that the fragment had an 8 bp overlapping sequence with the product of *cyp71av1* transcription. A forward primer designed according to the 650 bp fragment with a reverse primer CYPGSP1 (table 1), was able to amplify fragments of the expected size, from which the AD1 production was confirmed to be specific and corresponds to the *cyp71av1* 5'-flanking region.

Using the GenomeWalker™ Universal Kit, the DNA digested by three restriction enzymes and ligated with the GenomeWalker Adapter was used as a template for two step nested PCR amplification. The genomic DNA walking electrophoresis results showed that the products from the second PCR varied in length from 300 bp to over 2 kb (fig. 2). The sequencing results demonstrated that fragments of approximately 1.6 kb potentially were on the *cyp71av1* 5'-flanking region. The TAIL-PCR-amplified sequence, together with the 1.6 kb fragment, suggests that an approximately 2 kb upstream sequence of the *cyp71av1* gene was obtained.

Determination of TSS

5'-RACE was used to determine the TSS of the *cyp71av1* gene. GSP1 and GSP2 were used for PCR amplification with the Universal Primer A Mix after first strand of cDNA was synthesized. Electrophoresis results (fig. 3) revealed that a fragment of approximately 600 bp was amplified with GSP1, and a fragment of approximately 300 bp was amplified with GSP2. Since there are approximately 300 bp between GSP1 and GSP2, this variation in the size of PCR products is expected, and suggests that two fragments may be specific. The 600 bp fragment was ligated into the pMD18-T vector, and five individual clones were picked up for sequencing. Sequence comparison suggested that the TSS of *cyp71av1* is located at the 18th bp upstream before the initiation codon ATG. The 18th base is A, which conforms to the commonly held view that A is the base, located at the transcription initiation site.

Analysis of cyp71av1 promoter sequence

The PLANTCARE (http://bioinformatics.psb. ugent.be/webtools/plantcare/html) [12] and the PLACE Web Signal Scan (http://www.dna.affrc.go.jp/ PLACE/signalscan.html) [13] software were used to analyze the *cyp71av1* promoter in order to search for *cis*-act-

```
−1972  TCAAGCTCATAAGCTCATTTGTAGTTTATTTAATCAAATTGTGACTTTCTATTGTCGAAT
−1912  AGGATGACTCATTTTCCTAGTGCCAAAATTGAGTTCTTTCTATTGGTGTAACTGAAACCT
−1852  TATGAAAGATATAATTTCACATTTTGAATGTTTTCTTAAATCATATATCGAAGATATGGT
−1792  CGTATAATACATATGTTTTGTATTGCCAATTGGAAATGATCACGTTTTTATTTGTTTCTT
−1732  AAATGTTATAGCTCTATGCATCTTTCATCTATTTTCTTTACAAGTGGTCTAATGCATAAT
−1672  ATATGTAGTTAGAAAATTACTATATATATAATGTAAGGACGGATCTATGGAAGCGTGCCA
−1612  CACAATTAAATAATACTTGGGGCCTGTATTGGGTTAAGTGAGTCCATTAGATAACATCGA

−1552  TTTCGGGCCGGCCTGCCTTTTGGGTTTTGGTTACCGACACAGACTTATGTACATAGCTAA
                                              LTRE
−1492  AAGTAAAAGTAAACAATAATAATAGGAGGGACCTTTGTGGATGGTTACTGTGCTCTACGC
−1432  TGACTTATCTACATAGCTATCTATGAATTAACTATGAATGGCAGAAAACCATGTATTTTC
−1372  TTATTGATCTTAGCACAACTAAAAGTAAACACTACTATTTACAAAAATAAGCGCAACAAA
−1312  ATACACGAAAATTAAAGCTAACAGACATAAACAATAATTTGTACTTTAAGATAGCCACTA
−1252  CAACGCCTCTACTACAATAACAGAAACGACACAAGCTTCACAGTACGAGGCACAACAAAC
                               TGA-element              ABRE
−1192  ATAGTGTGCAACCTTCCAAGTATATCAACCTCAAATGGTTGAATCTCAACATTGTTCGTA
−1132  AAAAGCAGAAGGTCACTAATCAGCTTACAATCTTCGAAAACATAAACTGCATCAGCATTG
−1072  TTCGAAAAAAACCAACGCCCCGCCGCAACGCAGCGGACCGGCGCCCAACTCCTTAAGTCA
−1012  AAATATGGGTACGTTTATAAGAGCATCCACAGTGGGTCAAGTTGTTAAAAGCTTTTCACT
 −952  ATTCAAAAAGTCATTTTAACGGGTTTTTTGCATTATATCAAAAGCTTTTTGAAAAGCTT
 −892  TTCTTCAAAGACTTGGTGAACCTCAAGTCTTTCCTCCAAGCTTTAGTCACTAAATGTATT
 −832  ATGCATTAAATGCATTGAATGCATCAAGTCTTTAAGACTTGATGCATTGTAGATGCTCTA
 −772  AGACTAAGAATGTTTAGGTAGTTTCTCTTCCCAAAGTGGGCATAAAAAAGTAAAACTCAT
 −712  AAAAAAGTCAAGAGTGGACTGCGCGCTGCTACTCAAAAAAGTAGGCTAGGCACAATGGTG
 −652  ATCACAGACGATCATGCCACATCAGCTTTCTCTCTTCAATGAGGTGACGAGTATAGTAGC
                                                    TGACG-motif
 −592  GACCTTCTTATTTTTCCTTTATTTATTTATTTATTCACTTTAACCCATGCATGCATGCAT
 −532  GACATCTGCTTATTTAGTTAGTTATACGTGTGCATCAAAACAAAGTGTCACCACAACAAA
                                 ABRE
 −472  GTTTCATTGAAGAAGCCATCAGTGTGGGGACCTTCTTGCATGCTACCATCTATCTGCGCAT
 −412  CAAAACAAAGTTTCAGCTTATGTGCATCTATCTGTGCTACCAATCTTGCATGGCATGCTA
 −352  CCATCAGTGTGCATCAAAACAAAGTTTCACCAAATTTTCTCTGTGTTTTGCGTGGCTTGC
 −292  ATGCTACGTACCATCTTGCATGCTACCATCTATCTGTGAAGACTGAAGACCAACGAAGAT
          ACE
 −232  GGTACACGTACGGTAACAAATTTGTTTGTTTGTGTGTGTATTAAAGTCGGTCGACCAACT
          G-box
 −172  TGTGTATTTAGGCGACCGACGACCGATCAAGGACTAAGCCCCACGTTGGGCATGACCGAT
                                                 G-box
 −112  TATGACCAAGGTCGACCGATGGGTGGTGACCGACCACTGTGCCCAGCCTTATATTCAATT
                                                          CAAT-box
  −52  CCTATATCAATCCCATACAACAATATAAACAATCTCATCATTTGCGACCATAAAGAGTAT
             TATA-box
   +9  ACTAAAAGCAATGGCACTCTCACTGACCACTTCCATTGCTCTTGCAACGATCCTTTTGTT
  +69  CGTTTACAAGTTCGCTACTCGTTCCAAATCCACCAAAAAAGCCTTCCTGAGCCATGGCG
```

Fig. 4. Putative *cis*-elements on the *cyp71av1* promoter region. Numbers indicate the positions relatively to the transcription start site +1, which is in boldface. The translation initiation codon ATG is in bold. The putative TATA box, CAAT box, and other important putative *cis*-elements are boxed and in bold with a name label below. The WRKY-domain TGAC is only in bold. The underlined sequences are the possible motifs for interactions with the ABRE motif from −503 to −508 nt.

ing elements of this promoter. A TATA-box-like sequence (ATATAA) was found starting at position −25 nt to −30 nt, and CAAT-box-like sequence is present 23 bp upstream of the TATA-box-like sequence, which matches the traits of a prokaryotic promoter [14]. The results support the

idea that the TSS is located 18 bp upstream of the initiation codon ATG.

The G-box (−126 nt to −131 nt: CACGTT; −233 nt to −228 nt: CACGTA), ACE (−281 nt to −290 nt: GCTACGTACC), ABRE (−503 nt to −508 nt: TACGTG;
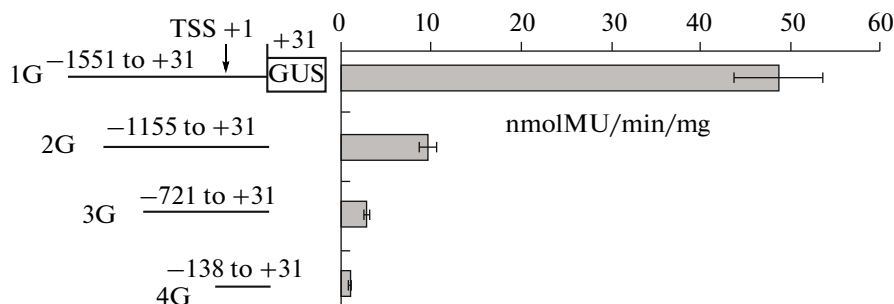
Fig. 5. The β-glucuronidase assay results. 1G–4G, corresponding to DEL1–DEL4, show the GUS activity record obtained at 0 min up to five points.

−1202 nt to −1211 nt: AGTACGAGGC), TGACG motif (−604 nt to −608 nt: TGACG), LTRE (−1465 nt to −1469 nt: CCGAC), TGA-element (−1123 nt to −1128 nt: AACGAC) and several W-box sites, which bind to the WRKY transcription factors, were also found inside the *cyp71av1* promoter (fig. 4).

The G-box is a hexameric and highly conserved motif in many promoters of plant genes [15]. It was first characterized as essential *cis*-element involved in the expression regulation of light-responsive genes, which can interact specifically with a family of bZIP proteins [16]. In *Arabidopsis*, bZIP proteins, such as the G-box binding factors, are called GBFs. The specificity of GBFs binding to a certain G-box depends on the DNA sequence close to this G-box element. In *Arabidopsis*, GBFs bound to a G-box might respond to a signal indirectly, as the signal is initially perceived by a factor bound to an adjacent motif such as an I-box. The G-box has been reported to respond to stimuli such as light, oxygen-free conditions, and several plant hormones [17]. The ACE/ACE motif, an ACGT-containing element, responds to UV stimulation. The ACGT sequence is common to many plant regulatory elements; many *cis*-acting elements containing this sequence contribute to the response to light stimulation [18], because many ACGT-binding proteins in the bZIP family are light induced. The existence of both elements − the G-box and the ACE/ACE−may be helpful in explaining the fact that the artemisinin output of *Artemisia* varies greatly and depends on daytime.

ABRE binds with ABA-responsive element binding protein 1 (AREB1). The core sequence of ABRE is part of more than two ABRE motifs in the promoter. Overexpression of AREB1 brings enhanced drought tolerance and makes the gene greatly up-regulated [19]. Another study suggested that the ABRE motif is functional when a CE3 motif is present at an approximate distance of 50 bp. In other instances, ACGT-containing ABREs and other elements are clustered and appear to function in a complex. Hobo and Asada demonstrated that a common sequence structure CGTG(G/T)C is considered important for both ABRE and CE3 function [20]. In our study, there was a cluster of elements formed by the ABRE (−503 nt to −508 nt) and three possible CGTG-motifs.

According to the location, it is possible that the sequence AGTGTG is the motif that interacts with the upstream ABRE-motif.

The TGACG-motif and the TGA elements play important roles in the methyl jasmonate (MeJA) responses [21]. Previous researchers demonstrated that the TGACG-motif is a binding site for bZIP *trans*-activating factors. Inserting this region into the constitutive CaMV 35S promoter gives it MeJA-responsive ability. Mutagenesis of these motifs abolishes MeJA-responsive expression. The TGA-motif also responds to SA stimulation.

LTRE is a low temperature responsive *cis*-acting element, which regulates cold-, drought-, and also ABA-related gene expression [22].

WRKY proteins are a superfamily of transcription factors. They are involved in the regulation of various functions in plants, from pathogen defense to trichome development, and all the WRKY proteins have highly conserved binding domains [23]. The presence of the WRKY core domain (TGAC) in the *cyp71av1* promoter is interesting. We know that the *cyp71av1* gene codes for P450 and that the product of this gene actually catalyzes a three step oxidation reaction. The WRKY domain inside the *cyp71av1* promoter may imply that the *wrky* gene may regulate the artemisinin biosynthesis pathway (fig. 4).

β-Glucuronidase assay

The charts (fig. 5) show the results of the β-glucuronidase assay of transient transgenic *Artemisia*. Each of the *cyp71av1* 5'-deletion promoters (DEL0–DEL4) of different length contains four repeats. For functional characterization of the *cyp71av1* promoter, fragments of different lengths of the 5'-sequence were deleted. 1G contains most of the detected *cis*-acting elements. 2G's TGA-elements and the ABRE located from −1252 nt to −1192 nt were deleted. 3G's *cis*-acting elements were almost the same as those of 2G's. The *cis*-acting elements of 4G only have the CAAT-box and the TATA-box. According to the quantitative fluorometric GUS assay results, the expression level of the 1G promoter fragments was 48.7493 nmol MU/min/mg protein. Since there were no obvious *cis*-acting elements from −2048 nt to −1551 nt, we considered that the 1G fragment has full-
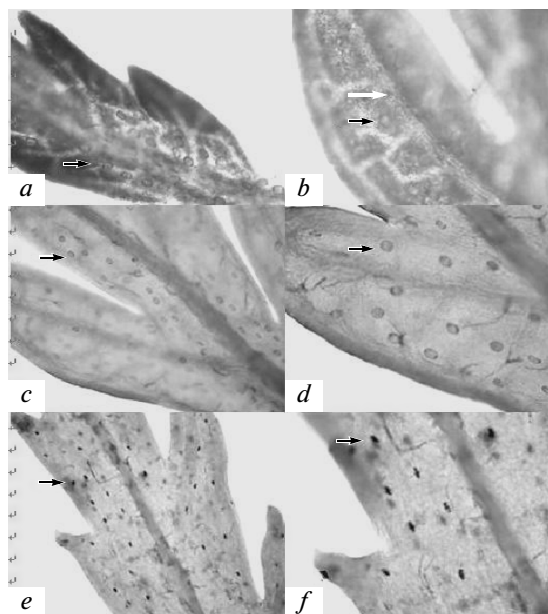
Fig. 6. Observation of GUS expression in stable transgenic *Artemisia* leaves under the microscope. *a* and *b* − control group − wild type *Artemisia* without dying; *c* and *d* − wild type *Artemisia*; *e* and *f* − *Artemisia* infected with pCAMBIA CYP-DEL1 EHA105 (black arrow: trichome, white arrow: stomata).

length promoter activity. Deletion of the region from −1125 nt to +31 nt, in the absence of the TGA-elements and one ABRE, remarkably reduced promoter strength: more than 80% of the promoter activity was lost. Deletion of the region from position −727, the 3G fragment, did not lead to additional promoter activity loss, however, compared to the 2G deletion, the strength of the promoter fragments was lost by more than 60%. The *cyp71av1* promoter sequence doesn't contain any obvious *cis*-acting elements between −1155 and −727 what contradicts to the GUS-assay results. Since the *cyp71av1* is an inducible promoter, it is worthwhile to use a yeast single-hybrid system to test whether possible protein responses to stimuli of interest interact with this DNA sequence domain. The 4G minimal promoter fragments, which have only two necessary *cis*-regulation elements, did not show much activity. This may result from the interference of the sequence 3D-structures from −727 nt to −138 nt. Though the 590 bp sequence contains several stimuli responsive elements, such as the ACE, G-box, ABRE, and Box E, their contribution to promoter strength may be less than their interference with it. Also it could be caused by the loss of interaction with the TGA-elements and the outer ABRE. Our deletion analysis indicates that all of the *cis*-elements are necessary for the *cyp71av1* promoter to achieve its maximal activity. Any deletion will definitely cause a reduction in promoter strength. In conclusion, the TGA-elements and the outer ABRE are critical for maximal promoter strength.

The *Artemisia* leaf-dyeing results (fig. 6) showed that only the *Artemisia* trichome cells in this picture were dyed deep blue and the other tissues were not. This implies that the *cyp71av1* gene is expressed only in glandular tri-

chome cells. These findings are sufficient to support the conclusion that the *cyp71av1* promoter is trichome specific.

REFERENCES

1. WHO. 2005. *World Malaria Report.*

2. Tianwei Liu, Lingbo Qu, Bingren Xiang. 2003. The progress of artemisinin anti-malaria drugs. *Chin. J. Med. Guide*. **5**, 399−401.

3. WHO. 2006. *Guidelines for the treatment of malaria.* WHO, Geneva.

4. Liu C.Z., Zhao Y., Wang Y.C. 2006. Artemisinin: current state and perspectives for biotechnological production of an antimalarial drug. *Appl. Microbiol. Biotechnol.* **71**, 11−20.

5. Bertea C.M., Freije J.R., van der Woude H., et al. 2005. Identification of intermediates and enzymes involved in the early steps of artemisinin biosynthesis in *Artemisia annua. Planta Med.* **71**, 40−47.

6. Teoh K.H., Polichuk D.R., Reed D.W., et al. 2006. *Artemisia annua* L. (Asteraceae) trichome-specific cDNAs reveal CYP71AV1, a cytochrome P450 with a key role in the biosynthesis of the antimalarial sesquit-

erpene lactone artemisinin. *FEBS Lett.* **580**, 1411–1416.

7. Ro D.K., Paradise E.M., Ouellet M., et al. 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*. **440**, 940–943.

8. Jing F.Y., Zhang L., Li M.Y., et al. 2009. Abscisic acid (ABA) treatment increases artemisinin content in *Artemisia annua* by enhancing the expression of genes in artemisinin biosynthetic pathway. *Biologia*. **64**, 319–323.

9. Vergauwe A., Cammaert R., Vandenberghe D., et al. 1996. *Agrobacterium tumefaciens*-mediated transformation of *Artemisia annua* L. and regeneration of transgenic plants. *Plant Cell Rep*. **15**, 929–933.

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

11. Cote C., Rutledge R.G. 2003. An improved MUG fluorescent assay for the determination of GUS activity within transgenic tissue of woody plants. *Plant Cell Rep*. **21**, 619–624.

12. Lescot M., Déhais P., Moreau Y., et al. 2002. Plant-CARE: 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.

13. Higo K., Ugawa Y., Iwamoto M., Korenaga T. 1999. SIGNAL SCAN: A computer program that scans DNA sequences for eukaryotic transcriptional elements. *CABIOS*. **7**, 203–206.

14. Ke J., Choi J.K., Smith M., et al. 1997. Structure of the *CAC1* gene and *in situ* characterization of its expression (the *Arabidopsis thaliana* gene coding for the biotin-containing subunit of the plastidic acetyl-coenzyme a carboxylase). *Plant Physiol*. **113**, 357–365.

15. Menkens A.E., Schindler U., Cashmore A.R. 1995. The G-box: a ubiquitous regulatory DNA element in plants bound by the GBF family of bZIP proteins. *Trends Biochem. Sci*. **20**, 506–510.

16. Guiltinan M.J., Marcotte W.R.J., Quatrano R.S. 1990. A plant leucine zipper protein that recognizes an abscisic acid response element. *Science*. **250**, 267–271.

17. McKendree W.L.J., Ferl R.J. 1992. Functional elements of the Arabidopsis Adh promoter include the G-box. *Plant Mol. Biol*. **19**, 859–862.

18. Block A., Dangl J.L., Hahlbrock K., et al. 1990. Functional borders, genetic fine structure, and distance requirements of *cis* elements mediating light responsiveness of the parsley chalcone synthase promoter. *Proc. Natl. Acad. Sci. USA*. **87**, 5387–5391.

19. Fujita Y., Fujita M., Satoh R., et al. 2005. AREB1 is a transcription activator of novel ABRE-dependent ABA signaling that enhances drought stress tolerance in *Arabidopsis*. *Plant Cell*. **17**, 3470–3488.

20. Hobo T., Asada M., Kowyama Y., Hattori T. 1999. ACGT-containing abscisic acid response element (ABRE) and coupling element 3 (CE3) are functionally equivalent. *Plant J*. **19**, 679–689.

21. Mason H.S., DeWald D.B., Mullet J.E. 1993. Identification of a methyl jasmonate-responsive domain in the soybean *vspB* promoter. *Plant Cell*. **5**, 241–251.

22. Baker S.S., Wilhelm K.S., Thomashow M.F. 1994. The 5'-region of *Arabidopsis thaliana cor15a* has *cis*-acting elements that confer cold-, drought- and ABA-regulated gene expression. *Plant Mol. Biol*. **24**, 701–713.

23. Urao T., Yamaguchi-Shinozaki K., Urao S., Shinozaki K. 1993. An *Arabidopsis myb* homolog is induced by dehydration stress and its gene product binds to the conserved MYB recognition sequence. *Plant Cell*. **5**, 1529–1539.