JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Cloning and Expression of Desoxyhemigossypol-6-*O*-methyltransferase from Cotton (*Gossypium barbadense*)

JINGGAO LIU,[†] CHAUNCEY R. BENEDICT,[†] ROBERT D. STIPANOVIC,*[,‡] CLINT W. MAGILL,[§] AND ALOIS A. BELL[‡]

Department of Biochemistry and Biophysics and Department of Plant Pathology and Microbiology, Texas A&M University, College Station, Texas 77843, and Southern Plains Agricultural Research Center, Agricultural Research Service, U.S. Department of Agriculture, College Station, Texas 77845

Terpenoids play an important role in defense against insects and pathogens in cotton. These terpenoids contain phenolic groups. Metabolites in which the phenolic group has been converted to a methoxy group are less toxic to most insects and pathogens and thus may alter resistance. Here is reported the cloning of a gene from *Gossypium barbadense* that encodes the enzyme that methylates the phenolic group of desoxyhemigossypol (dHG) exclusively at the 6-position, dHG-6-*O*-methyltransferase (dHG-6-OMT). Partial peptide sequences from digests of purified dHG-6-OMT were used to design primers for RT-PCR amplification of cDNA fragments from poly(A) mRNA. Fragments were extended to full length using 5′ and 3′ RACE. The resulting cDNA codes for a 365-residue polypeptide with a calculated molecular weight of 40.6 kDa, in agreement with the molecular mass of purified dHG-6-OMT. When expressed in *Escherichia coli*, the bacterial lysates showed a high specificity for the methylation of desoxyhemigossypol, differentiating the cloned gene from other pathogen-induced methyltransferases.

KEYWORDS: ***Gossypium barbadense*; desoxyhemigossypol; expression in *Escherichia coli*; desoxyhemigossypol-6-methyl ether; host defense; terpene phytoalexins**

INTRODUCTION

Foliar glands in cotton (*Gossypium barbadense*) contain the terpenoids desoxyhemigossypol (dHG), hemigossypol (HG), hemigossypolone (HGQ), and heliocides (HH_2 and HH_3) and the methylated forms of these terpenoids (**Scheme 1**) that are part of the plant's defense system against pathogenic fungi and insects. dHG, HG, desoxyhemigossypol-6-methyl ether (dMHG), and hemigossypol-6-methyl ether (MHG) are induced in cotton in response to infection by *Verticillium dahliae* and *Fusarium oxysporum* f. sp. *vasinfectum* (*1*, *2*). Alchanati et al. (*2*) also showed desoxyhemigossypol-*O*-methyltransferase was induced by *V. dahliae*; this was proposed to account for the dMHG formation. dHG and HG are toxic to the pathogenic fungi, but dMHG and MHG are less toxic (*3*, *4*). HGQ and the heliocides HH_2 and HH_3 have been shown to protect cotton from insects such as *Heliothis virescens* (*5*, *6*). The methylated forms of these terpenoids MHGQ, HB_2, and HB_3 are less toxic toward insects (*7*).

dHG is an intermediate in the biosynthesis of HG, HGQ, HH_2, and HH_3. The formation of dMHG, MHG, MHGQ, HB_2, and HB_3 also proceeds from dHG (**Scheme 1**). The initial step in the formation of the methylated terpenoids is the methylation of the 6-hydroxyl group of dHG catalyzed by *S*-adenosyl-L-methionine:desoxyhemigossypol-6-*O*-methyltransferase (dHG-6-OMT) (*8*). The product of this reaction, dMHG, leads to the synthesis of MHG, MHGQ, HB_2, and HB_3. The modulation of the dHG-6-OMT gene may control the relative amounts of unmethylated and methylated terpenoids and regulate cotton's resistance to pathogenic fungi and herbivorous insects. Herein we report the cloning and expression of dHG-6-OMT from cotton. This is one of relatively few examples where cloning of a target gene has been verified by both sequence analysis and the enzymatic properties of the encoded protein.

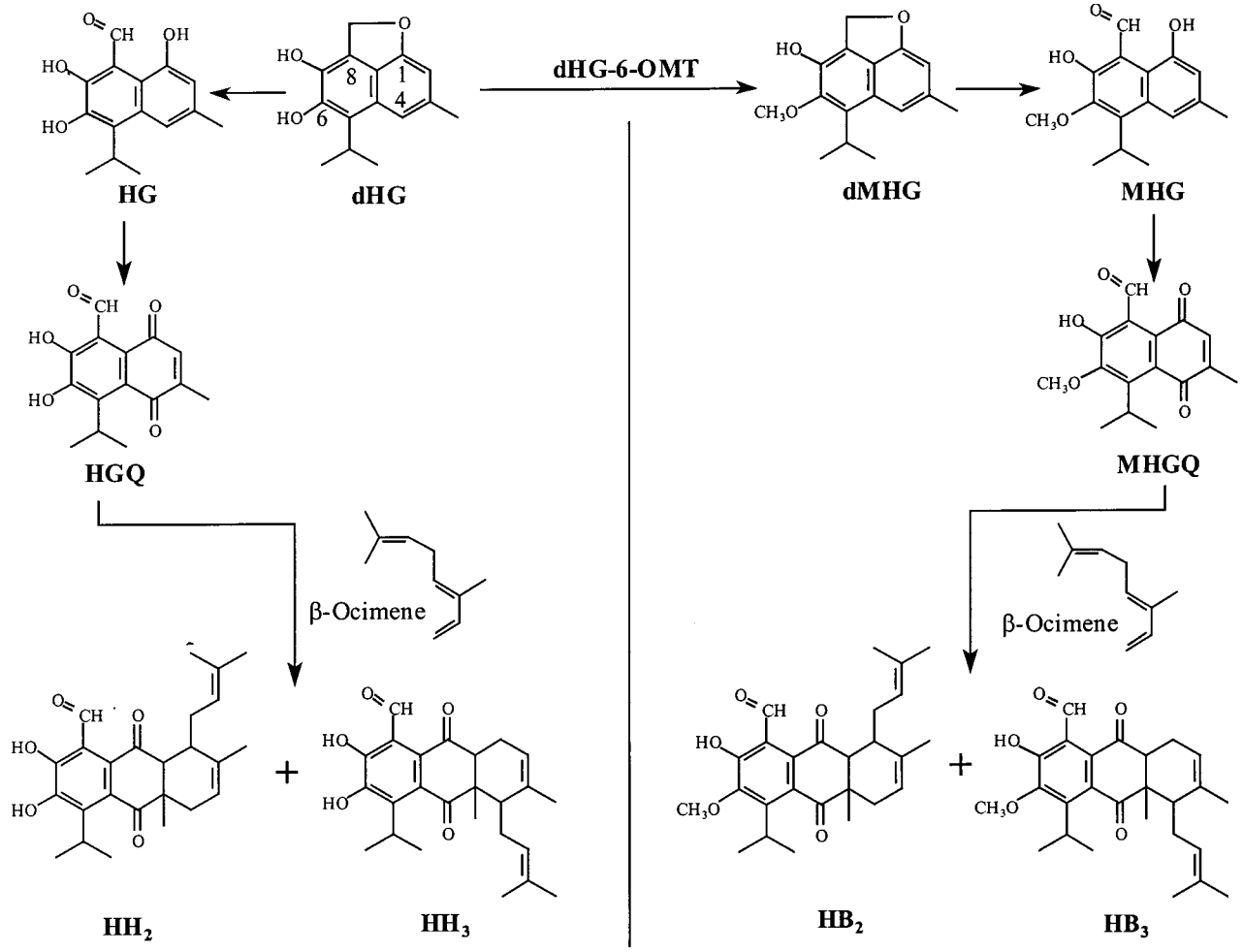
MATERIALS AND METHODS

Chemicals. *S*-Adenosyl-L-[*methyl*-$^3H_3$]methionine was purchased from Amersham (Arlington Heights, IL) at a specific radioactivity of 18.5 Gbq/mmol (500 mCi/mmol). *S*-Adenosyl-L-methionine, reduced glutathione (GSH), insoluble polyvinylpolypyrrolidone (PVP), and tris-(hydroxymethyl)aminomethane (Tris) were purchased from Sigma (St. Louis, MO). 4-Methylcatechol, 2,3-dihydroxynaphthalene, and caffeic acid were purchased from Aldrich (Milwaukee, WI). Ten percent Ready gels and SDS-PAGE standards (low range) were purchased from Bio-Rad (Hercules, CA). Coomassie plus protein assay reagent was purchased from Pierce Chemical (Rockford, IL). dMHG was isolated

* Author to whom correspondence should be addressed [e-mail rdstip@cpru.usda.gov; telephone (979) 260-9232; fax (979) 260-9319].
[†] Department of Biochemistry and Biophysics.
[‡] U.S. Department of Agriculture.
[§] Department of Plant Pathology and Microbiology.

10.1021/jf011701y CCC: $22.00
Published on Web 04/12/2002

Scheme 1



from cotton stems infected with *V. dahliae* according to the procedure of Stipanovic et al. (*9*). dHG is highly labile and was prepared by demethylation of dMHG according to the procedure of Stipanovic et al. (*10*). The dHG prepared by this method had an mp of 145−147 °C after crystallization from ether−hexane solution.

Plant and Inoculum Preparation. *V. dahliae* defoliating strain V-76 was grown, and the plants were inoculated as described previously (*8*).

dHG-6-OMT Purification and Assay. dHG-6-OMT was purified and the enzyme assayed as previously described (*8*). The protein concentration in the different extracts was determined by using the method of Bradford (*11*). SDS-PAGE of the purified protein was carried out according to the procedure of Laemmli (*12*) in 10% separation gel and 5% stacking gel. Proteins were stained on the gels using Coomassie blue. The SDS-PAGE gel was stained with Coomassie blue, and a band at 42.1 kDa was cut out. This gel slice containing ~7 μg of pure dHG-6-OMT was sent to the Protein Chemistry Core Laboratory at Baylor College of Medicine for amino acid sequence analysis. The dHG-6-OMT was digested with Lys-C in gel, and the resulting peptide mixture was separated on a narrow-bore C18 reversed phase column (Pharmacia SMART system). Selected peptide peaks were microsequenced with an Applied Biosystems 477A automated sequencer.

RNA Extraction and mRNA Purification. Crude enzyme extracts were prepared from cotton stems as previously described (*8*) at four different inoculations, and the dHG-6-OMT and *o*-diphenol-OMT activities in each extract were determined. The tissue having the highest dHG-6-OMT versus *o*-diphenol-OMT activity ratio was used for RNA extraction. Thus, 1.4 g of cotton stems infected with *V. dahliae* was ground in liquid N_2, and total cellular RNA was isolated using an RNeasy plant total RNA kit as directed by the manufacturer (Qiagen), except for inclusion of 1% PVPP-40 in the lysis buffer (RLC). Absorbance at 260 and 280 nm was used to determine the purity and concentration of RNA. A typical yield was 102 μg. Polyadenylated mRNA was isolated by two cycles of chromatography on oligo(dT) cellulose columns (type 7, Pharmacia, Uppsala, Sweden) according to the procedure of Avis and Leder (*13*).

Table 1. Sequences of Lys-C Peptides of Purified dHG-6-OMT from Cotton

peptide	sequence[a]
PL27	ATMDPIITK
PL53[b]	VLHDWNXEG
PL67	GAFEGLDSLVDVGGGTGTTA
PL73	LFLAAGFSSFK
PL85	YVGGNMFEAFPTGDAILLK
PL89	LIIIDMVVRENXXV
PL99	PWSFLGTWFQNDDHTPFATAYGK

[a] X denotes undetermined amino acids. [b] The sequenced fraction consisted of at least two peptides. Only the prominent sequence is listed.

Synthesis of First Strand cDNA and PCR Amplification. Pools of oligonucleotide primers encoding portions of two dHG-6-OMT Lys-C peptides (**Table 1**), PL99 and PL85, were synthesized with the number of different species (degeneracy) in each pool minimized as described by Klein et al. (*14*). In general, by avoiding amino acids with four or six codons when possible, especially at the 3′ terminus (none in the last two bases), and taking advantage of the ability of IMP to pair with A, C, or U, low-complexity primers can generally be synthesized (*15*−*17*). Homology to other OMTs revealed by BLAST searches suggested that peptide PL99 lies upstream of peptide PL85. Therefore, a sense primer RTSP (29-mer, 16-fold degeneracy) was made corresponding to the PL99 fragment, and an antisense primer RTAP

Table 2. Primers Used in PCR Amplifications, First-Strand cDNA Synthesis, and Sequencing

primer	DNA sequence[a]	position in dHG-6-OMT cDNA
sense		
RTSP	GGIACITGGTTYCARAATGATGAYCAYAC	450–478
L3′-GSP1	TCAGCTAAGATGAATATGGGCAATG	33–57
L3CP	GCACGCCAAAAGCTAGACAATG	300–321
L7CP	TCATGGAGTGCACTGTG	697–713
L11CP	CTCCTATTGTGGGTTTGAC	1096–1114
LEP[b]	GACTGAGAATTCATGAATATGGGCAATGCTA (*Eco*RI)	42–60
antisense		
RTAP	TCICCIGTIGGRAAIGCTTCRAACATRTT	777–805
L5′-GSP1	CTCCAACATATTTCAAGT	757–774
L5′-GSP2	TGCAAGCCAGCAACAATATTAG	724–745
L5′-GSP3	GATCAAACACAGTGCACTCCATGA	697–720
LXP[b]	CATGGACTCGAGTCAGGGGTAAACCTCAATG (*Xho*I)	1121–1139

[a] Y = C + T; R = A + G. [b] Position refers to the bases after the restriction site.

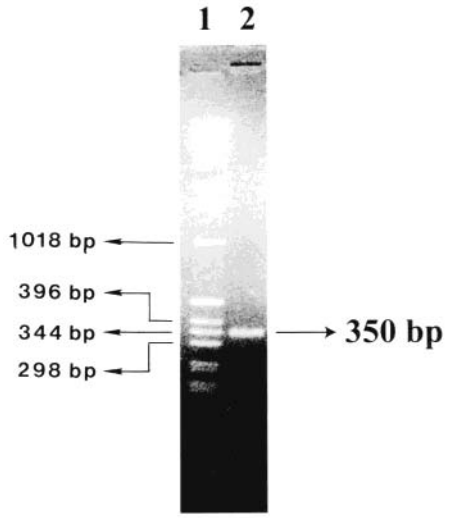


Figure 1. Amplification products of RT-PCR reaction directed by primers corresponding to dHG-6-OMT Lys-C digest peptides PL99 and PL85 (lane 2). Lane 1 represents 1 kb ladder from GIBCO-BRL.

(29-mer, 8-fold degeneracy) was made corresponding to the PL85 fragment (**Table 2**).

A 6 ng sample of poly(A) mRNA was used as template for synthesis of first-strand cDNA. Random hexamers were annealed to the RNA and extended with Moloney murine leukemia-virus reverse transcriptase, first at room temperature (10 min) and then at 42 °C (15 min) under conditions recommended by the manufacturer (Perkin-Elmer Applied Biosystems). The whole reaction mixture (20 μL) was then used as the PCR template. PCRs were performed in 10 mM Tris-HCl (pH 8.3), 50 mM KCl using 2.5 units of Amplitaq DNA polymerase, 4.1 mM $MgCl_2$, 0.2 mM each dNTP, and 0.15 μM each primer corresponding to peptides PL99 and PL85 in a total volume of 0.1 mL. After holding samples at 95 °C for 105 s, a total of 40 cycles of denaturing at 95 °C for 15 s, annealing at 59 °C for 15 s, and extension at 72 °C for 30 s were carried out. After the final cycle, the reaction was extended for 7 min at 72 °C. The amplified products were analyzed by 1.5% agarose gel electrophoresis (**Figure 1**), which showed a single band at ~350 bp. The PCR product was cloned into the pNoTA vector, a component of the PCR cloning kit available from 5 Prime → 3 Prime Inc. The size of the insert in plasmid DNA extracted from selected white colonies was tested by electrophoresis on 1.3% agarose gels following digestion of plasmid DNA with *Bam*HI. Two colonies that contained an insert of ~330 bp were grown overnight in 2 mL of 2× YT medium, and pure plasmid DNA was extracted using a Qiaprep spin plasmid kit. Sequencing was performed with an ABI Prism BigDye terminator cycle sequencing ready reaction kit using (−21) M13 primer on an ABI Prism 377.

5′ RACE for Amplification of the 5′ End of cDNA. RACE (Rapid Amplification of cDNA Ends) was used to obtain the 5′ end of the dHG-6-OMT using components and protocols provided by GIBCO-BRL (Gaithersburg, MD) except 30 ng of poly(A) mRNA was substituted for total RNA as template. Three nested antisense primers, L5′-GSP1, L5′-GSP2, and L5′-GSP3 (**Table 2**), were designed using the sequence information of the 326 bp dHG-6-OMT cDNA obtained above. First-strand cDNA synthesis was primed by L5′-GSP1 and PCR amplification of the dC-tailed cDNA was primed by a nested dHG-6-OMT specific primer, L5′-GSP2 and the 5′ RACE abridged anchor primer. A second round of nested amplification was carried out using the dHG-6-OMT specific primer L5′-GSP3 and the abridged universal amplification primer. The amplified 720 bp PCR product was cloned and sequenced as described above.

3′ RACE for Amplification of the Whole cDNA. A dHG-6-OMT specific primer, L3′-GSP1 (**Table 2**), was designed using the sequence information of the above cloned 720 bp cDNA fragment starting within the leader sequence to clone the entire cDNA using 3′ RACE. First-strand cDNA was synthesized using 30 ng of poly(A) mRNA as template and an adapter primer (supplied by GIBCO-BRL) as the primer. Amplification of the cDNA was accomplished utilizing the primer L3′-GSP1 and the abridged universal amplification primer (supplied by GIBCO-BRL). Pfu turbo DNA polymerase (Stratagene, La Jolla, CA) was used in place of Taq DNA polymerase to ensure high fidelity. The amplified 1500 bp PCR product was cloned and sequenced as described above. Among eight randomly chosen clones, the four largest clones had identical sequences and were designated pdHG6OMT-NoTA. In addition to the M13 universal primers, primers L3cp, L7cp, L11cp, and L5′-GSP3 (**Table 2**) were also used for the sequencing reactions.

Expression of Recombinant dHG-6-OMT in *Escherichia coli*. A pair of PCR primers, LEP and LXP (**Table 2**), was designed to provide an *Eco*RI site immediately upstream of the start codon of dHG-6-OMT cDNA and an *Xho*I site immediately downstream of the stop codon. Amplification using pdHG6OMT-NoTA plasmid DNA as template was performed in 20 mM Tris-HCl (pH 8.8), 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100, 0.1 mg/mL BSA (nuclease-free), 200 μM dNTP, 0.4 μM for both primers, and 2.5 units of Pfu turbo DNA polymerase with a final volume of 100 μL. The reaction involved denaturing at 94 °C for 50 s, annealing at 55 °C for 1 min, and extension at 72 °C for 1.5 min for a total of 25 cycles.

The PCR product was digested with *Eco*RI and *Xho*I. The recovered dHG-6-OMT coding region was cloned into the identical sites in pET28a(+) (Novagen) to make the final product pdHG6OMT-ET28. This plasmid was transformed into nonexpression host DH5α *E. coli* cells. Positive clones were screened with mini prep restriction analysis and PCR. Plasmids recovered from the positive clones were then transformed into *E. coli* expression host BL21(DE3)pLysS. Single colonies were used to inoculate 10 mL of LB containing kanamycin (30 μg/mL) and chloramphenicol (34 μg/mL) at 37 °C for 4 h. The culture was added to 500 mL of fresh medium containing antibiotics. After 4 h (OD_{600} = 0.7), IPTG was added to a final concentration of 1 mM to induce in vivo protein expression. After 3 h of growth at 25 °C, cells were harvested by centrifugation at 5000*g* for 5 min. The pellet was resuspended in cold 50 mM Tris-HCl (pH 8.0), and the cells were frozen at −80 °C after collection by centrifugation. Aliquots of cells were lysed by quickly thawing the frozen cells at 4 °C. The cell lysate was suspended in one-eighth the original culture volume of 50 mM Tris-HCl (pH 8.0) containing 5 mM GSH, and passed through a 27 gauge needle several times. The soluble enzyme preparation was obtained by centrifugation of this crude cell lysate at 100000*g* for 20 min at 4 °C. These preparations were used directly for protein concentration determination, SDS-PAGE, western blotting, and enzyme activity assay.

For western blotting, the samples were separated on SDS-PAGE and then transferred to a nitrocellulose membrane. The expressed

proteins containing T7.Tag fusion proteins were detected on the membrane using a T7.Tag AP LuminBlot Kit according to the manufacturer's instructions (Novagen).

The insert DNA in the plasmid pdHG6OMT-ET28 recovered from cells of BL21(DE3)pLysS displaying dHG-6-OMT activity was sequenced to verify fidelity of the PCR amplification step in the pdHG6OMT-ET28 construction procedure. In addition to T7 promoter and terminator sequences, L3CP and L5′-GSP1 (**Table 2**) were also used as sequencing primers.

RESULTS AND DISCUSSION

Cloning and Characterization of dHG-6-OMT cDNA Clone. Numerous *O*-methyltransferases occur in plants, and many of them are involved in secondary metabolism (*18*). In addition, both caffeic acid 3-*O*-methyltransferase (COMT) and caffeoyl CoA 3-*O*-methyltransferase (CCOMT) are typically induced in stressed plants (*19*, *20*). An *O*-methyltransferase presumed to be involved in lignin synthesis has previously been shown to be induced in cotton following inoculation with *V. dahliae* (*21*). Thus, both of the possibilities that dHG is methylated by a nonspecific methylase or that another methyltransferase might be cloned were important considerations in the effort to clone dHG-6-OMT. By first isolating, purifying, and demonstrating substrate specificity of an enzyme with dHG-OMT activity in cotton steles induced by inoculation with *V. dahliae* (*8*), it was possible to specifically clone the target gene. Sequence information derived from fragments of the purified protein permitted the design of gene-specific PCR primers. The protein was initially blotted to a PVDF membrane and subjected to N-terminal sequence analysis directly on the membrane. The N terminus turned out to be blocked (data not shown). Therefore, sequencing was performed on peptides generated by proteolytic cleavage. The protein was digested with Lys-C, and the resulting mixture of peptides was purified by a narrow-bore HPLC. Selected fractions were subjected to sequence analysis. Internal amino acid sequences from seven peptide fragments of 9−23 amino acids have been determined (**Table 1**). One peptide fraction, however, was heterogeneous and contained at least two subsequences. Only the more prominent sequence is listed in **Table 1**. The highest identities of the seven peptides ranged from 54 to 83% with other OMTs.

The partial amino acid sequences of dHG-6-OMT were used to design primers for the RT-PCR amplification of a cDNA fragment from poly(A) mRNA. To increase amplification specificity, the complexity of the degenerate primer pools was minimized either by incorporation of dIMP at ambiguous positions or by providing the more stable (*16*) or the more probable (*22*) bases except in the last two positions. The resulting primer pair, RTSP (sense, 29-mer, 16-fold degeneracy) corresponding to peptide PL85 and RTAP (antisense, 29-mer, 8-fold degeneracy) corresponding to peptide PL99 (**Table 2**), directed the synthesis of a single 350 bp amplification product using a random-hexamer primed first-strand cDNA as a template (**Figure 1**). Further consideration for improving amplification specificity was given to the choice of mRNA source for the first-strand cDNA synthesis. First, poly(A) mRNA was used in place of total RNA. Next, several extracts of cotton stele tissues were screened for *o*-diphenol-OMT versus dHG-6-OMT activities. The activity ratio ranged from 1.2 to 5.6. The extracts that had the ratio of 1.2 were chosen as mRNA source to lower the probability of amplifying the *o*-diphenol-OMT sequence. Because both dHG-6-OMT peptides (**Table 1**) and the 20 amino acid N-terminal peptide of the *o*-diphenol-OMT showed high homology to the same OMTs from other plants, it was necessary that any putative clones we obtained be checked for their authenticity.

The PCR product was cloned into a pNoTA vector, and the inserts in two randomly chosen white colonies were sequenced. Both inserts were 326 bp long and identical to each other except for ∼20 bp overhanging bases at the ends. The sequences share up to 57% identity at the amino acid level with OMTs from other plants. Although the primers used for PCR truncated the codons at the ends of the insert, the remaining codons are in complete agreement with the corresponding peptide sequences. When translated, 60 nucleotides in the middle of the inserts matched exactly with dHG-6-OMT peptide fragment PL68 (20 amino acids). This establishes that the inserts code for part of dHG-6-OMT.

Having established with near certainty that a portion of the target gene had in fact been cloned, sequences inside the clone were next used with 5′ RACE to obtain the beginning of the dHG-6-OMT mRNA. First-strand cDNA synthesis was initiated with primer L5′-GSP1 (**Table 2**) for subsequent amplification with a nested primer (L5′-GSP2) (**Table 2**). Because the amplified product showed a broad background of bands in addition to a major band at ∼730 bp in an agarose gel, a second round of amplification was carried out using a second nested primer, L5′-GSP3 (**Table 2**), with the 100-fold-diluted product of the first round of amplification as template. This gave a 720 bp cDNA; this cDNA was cut out from the gel, cloned into pNoTA, and sequenced. Sequence analysis confirmed the identity of the cDNA as encoding for the 5′ portion of dHG-6-OMT including a 41 bp 5′-untranslated region. In the region where the sequence overlapped with the previous 326 bp cDNA, complete sequence identity was observed. This sequence also provided an exact match for three dHG-6-OMT peptides, PL27, PL23, and PL67 (**Table 1**).

Using a primer starting within the leader sequence, primer L3′-GSP (**Table 2**), a 1500 bp cDNA was obtained by following the 3′ RACE protocol. The cDNA contained an open reading frame that codes for a 365 residue polypeptide with a calculated molecular mass of 40576 and p*I* of 5.23 (**Figure 2**). The calculated molecular mass corresponds very well with the reported subunit molecular mass of 42.1 kDa from SDS-PAGE for dHG-6-OMT (*5*).

The dHG-6-OMT cDNA contains a 5′-leader of 41 nucleotides and a 299-nucleotide 3′-untranslated region including a poly(A) tail. The sequences surrounding the first two ATG codons at positions 42 and 48 both match reasonably with the consensus for the plant initiation sequences (*23*). Furthermore, sequence alignment of dHG-6-OMT cDNA with other OMTs' sequences indicated that both ATG codons can be aligned into initiation positions of other OMTs. However, because most plant genes initiate at the first ATG (*24*), the ATG at position 42 was used as the initiation site for the cDNA's open reading frame. The richness in TA content in the leader sequence (73%) also pinpoints this ATG as the most probable functional initiation site (*24*). The conserved poly(A) addition signal AATAAT (*25*) was located in tandem at 108 nt upstream of the mRNA poly(A) tail (positions 1351−1362), and sequences upstream and downstream of the second AATAAT are in favorable context (*25*).

Protein Sequence Comparison and Analysis. Exact match of all seven peptide sequences (**Table 2**) in the cDNA-translated protein confirm again that the cDNA indeed codes for dHG-6-OMT. In a search of the GenBank database, the best match for the deduced amino acid sequence of dHG-6-OMT was an

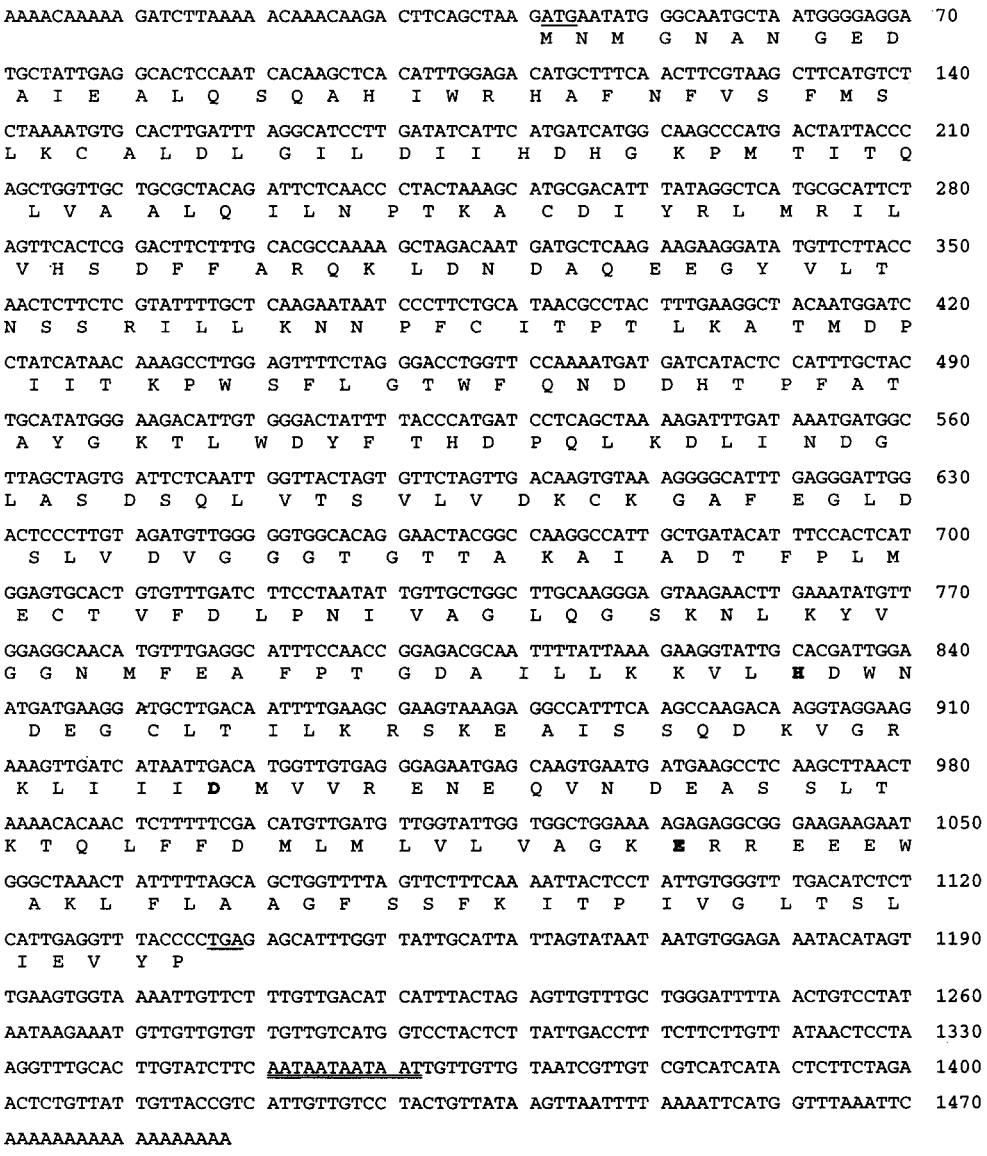


Figure 2. Nucleotide and deduced amino acid sequences of cotton dHG-6-OMT cDNA. The start and stop codons are underlined, and polyadenylation signals are marked by double underlines. The three putative active site residues involved in the deprotonation of the hydroxy group are in bold type.

OMT from almond flower tissue that reflected a 57% sequence identity (GenBank accession number AJ223151) over the entire protein. All of the best matches were OMTs from plant sources including (iso)flavonoid-OMTs, alkaloid-OMTs, caffeic acid-OMTs, catechol-OMTs, and *o*-diphenol-OMTs. For example, 46% identity was found with an isoflavonoid-OMT from garden pea (GenBank accession number T06786), 37% identity with an alkaloid-OMT from Japanese goldthread (GenBank accession number BAB08004), 34% identity with a catechol-OMT from barley (GenBank accession number S52015), 35% identity with a caffeic acid-OMT from Monterey pine (GenBank accession number AAD24001), and 28% identity with an *o*-diphenol-OMT from pepper (GenBank accession number T12259). The three possible SAM binding motifs, LDSLVDVGGGTG (amino acids 196−207), PTGDAILLKKV (amino acids 252−262), and AISSQDKV (amino acids 281−288), were also found in the same spatial arrangement (*26*).

By aligning 27 plant *O*-methyltransferases, Ibrahim (*27*) identified five highly conserved regions (I−V) comprising 36 amino acid residues. Except for an unusually short caffeoyl CoA-OMT from parsley, the overall identities of the individual OMTs with the consensus in these regions ranged from 61 to 100%. In comparison, dHG-6-OMT had 63.9% identity with consensus. In individual regions, dHG-6-OMT amino acid positions 199−207 share 100% identity with consensus region I (LVDVGGGTG vs LVDVGGGXG), amino acids 222−230 (CTVFDLPNI) show 44.4% identity with region II (GINFDL-PHV), amino acids 241−248 share 62.5% identity with region III (EHVGGDMF vs KYVGGNMF), residues 285−293 show 33.3% identity with region IV (NGKVI vs QDKV(GRKL)I), and four of six amino acids in positions 328−333 match consensus region V (GGKERT vs AGKERR). Regions I and IV correspond to the first and third SAM binding motifs identified by Li et al. (*26*). Examination of the X-ray crystal structures of the two plant OMTs (*28*) and a catechol OMT (*29*) indicate that regions I−III and the second SAM binding motif identified by Li et al. (*26*) are involved in SAM binding, region IV in metal binding, and region V in catalysis. The low homology of dHG-6-OMT to the consensus in region IV is well within expectation because dHG-6-OMT does not require a metal ion for its enzyme activity (*8*). Three residues, His 264, Asp 296, and Glu 331, are probably involved in the deprotonation of the 6-hydroxy group of dHG-6-OMT for the subsequent methylation by SAM because they match a pattern

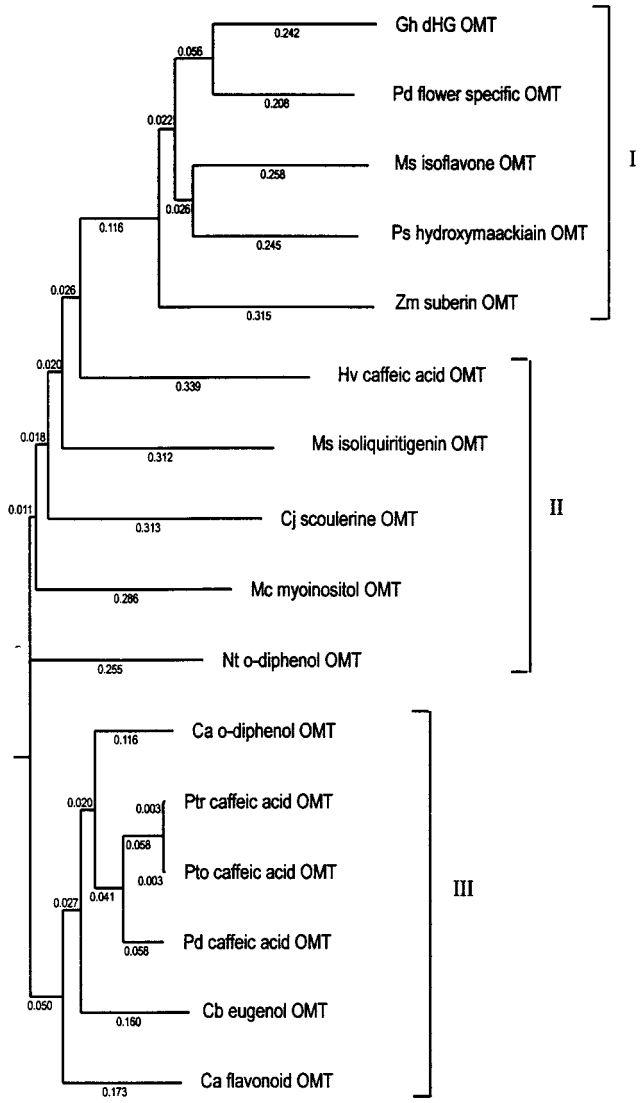


Figure 3. Phylogenetic analysis of plant OMT amino acid sequences using the Clustal W alignment program from MacVector. Genus names and accession numbers (in parentheses) are listed in descending order as follows: *Gossypium hirsutum* (present paper), *Prunus dulcis* (AJ223151), *Medicago sativa* (T09707), *Pisum sativum* (T06786), *Zea mays* (P47917), *Hordeum vulgare* (AAC18643), *Medicago sativa* (T09617), *Coptis japonica* (BAA06192), *Mesembryanthemum crystallinum* (P45986), *Nicotiana tabacum* (JQ2344), *Capsicum annuum* (T12259), *Populus tremuloids* (Q00763), *Populus tomentosa* (Aaf63200), *Prunus dulcis* (Q43609), *Clarkia breweri* (AAC01533), and *Chrysosplenium americanum* (AAA80579).

highly conserved among the plant OMTs (*28*). The three residues corresponding to Phe 185, Met 189, and Met 329 of chalcone OMT are highly conserved in a variety of plant OMTs and are thought to be involved in substrate binding (*28*). However, in dHG-6-OMT, only the third residue was conserved, indicating the unique substrate binding requirement of this enzyme.

A phylogenetic analysis (**Figure 3**) of 16 different plant OMTs with diverse substrate specificities indicates that these OMTs can be separated into three clusters. Cluster III OMTs are mainly enzymes that are constitutively expressed in lignin biosynthesis and show little divergence from a common ancestral plant OMT. On the other hand, cluster I OMTs, which include dHG-6-OMT, are involved in the biosynthesis of phytoalexins and show the most divergence from the common ancestor. The separation as a cluster suggests that an ancestral form may have

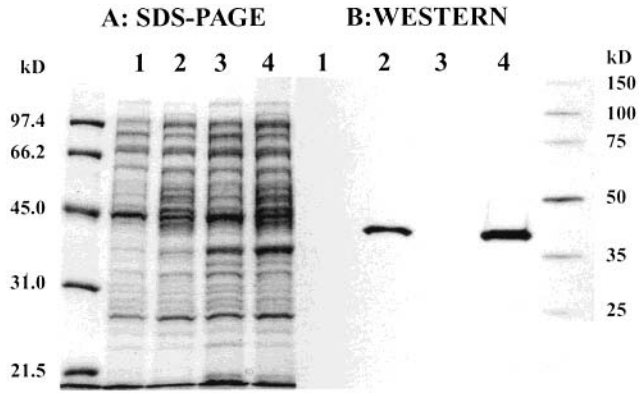


Figure 4. (A) SDS-PAGE (10%) and (B) Western blot of protein extracts from *E. coli* BL21(DE3)pLysS transformed by pdHG6OMT-ET28: (lanes 1 and 3) 3.8 μg of protein each from uninduced soluble and crude cell lysate, respectively; (lanes 2 and 4) 3.8 μg of protein each from induced soluble and crude cell lysate, respectively; (far-left lane) low-range SDS-PAGE molecular weight standards; (far-right lane) Perfect Protein Western markers. SDS-PAGE was stained with Coomassie blue, and western blot was detected with T7.Tag AP LumiBlot kit.

acquired a prototype phytoalexin producing ability, but the degree of separation between the lineages that developed suggest that each species evolved rapidly and independently to meet different environmental challenges. Cluster II members fall between clusters I and III in terms of divergence from the ancestral OMT. One characteristic of the cluster II OMTs is that they all appear to have evolved independently from an ancestral OMT. This may reflect the diverse functionality of these enzymes. For example, Ms isoliquiritigenin OMT is involved in the biosynthesis of 4,4′-dihydroxy-2′-methoxychalcone, the most potent of the nod-gene-inducing flavonoid derivatives released from alfalfa roots (*30*), Cj scoulerine OMT is involved in the biosynthesis of berberine, an alkaloid antibacterial agent (*31*), whereas Mc myoinositol OMT is induced by osmotic stress and involved in the biosynthesis of the cyclic sugar alcohol pinitol that may help the plant tolerate high saline conditions (*32*).

Expression of dHG-6-OMT in *E. coli*. To provide unequivocal proof for the identification of the cloned cotton dHG-6-OMT cDNA, the sequence was expressed in *E. coli*. The region flanking the coding DNA sequence was modified in a PCR reaction to incorporate unique restriction sites and to remove any noncoding plant DNA. The coding region was then subcloned into the *E. coli* expression vector pET28a(+). This vector encodes a His-Tag followed by a T7.Tag leader sequence. *E. coli* strain BL(DE3)pLysS was transformed with the above plasmid, and protein production was induced in cultures grown at 25 °C with 1 mM IPTG. Cultures grown at 37 °C resulted in loss of enzyme activity. The expressed His-Tag−dHG-6-OMT fusion protein failed to bind to a His-Tag sequence-specific affinity column in both native and denaturing conditions, and therefore purification of fusion protein was not attempted. His-Tag peptides can fail to bind to a Ni^{2+} column if they are buried within the protein or if they are folded such that binding is prohibited. The protein profiles in the crude and soluble preparation of *E. coli* lysate were determined by SDS-PAGE (**Figure 4A**). An induced single fusion protein band was detected by western blot using the T7.Tag antibody (**Figure 4B**).

E. coli lysates, both crude and soluble preparations, were tested for dHG-6-OMT activity. No activities were found in uninduced samples. Induced samples showed specific activities

Table 3. Activity of dHG-6-OMT Expressed in *E. coli* BL21(DE3)pLysS/pdHG6OMT-ET28 with Different Cosubstrates

substrate	specific activity (pkat/mg of protein)
dHG	31.76
4-methylcatechol	0.10
2,3-dihydroxynaphthalene	0.11
hemigossypol	0.05
2,7-dihydroxycadalene	0.05
caffeic acid	0.05

of 31.75 and 28.76 pkat/mg of protein for the crude and soluble preparations, respectively. This is ~9 times the dHG-6-OMT activity found in the crude enzyme extract from cotton stele tissue inoculated with *V. dahliae* (*8*). The crude *E. coli* lysate showed no enzymatic activity with 4-methylcatechol, 2,3-dihydroxynaphthalene, hemigossypol, 2,7-dihydroxycadalene, or caffeic acid (**Table 3**). This result is consistent with the enzyme characteristics displayed by the native dHG-6-OMT enzyme in cotton and verifies that the target enzyme was cloned rather than other *V. dahliae*-induced OMTs, including *o*-diphenol-OMT (assayed using 4-methylcatechol as substrate) or *o*-dinaphthol-OMT (assayed using 2,3-dihydroxynaphthalene as substrate).

Although sequence homology can be very useful in cloning similar genes from the same or different organisms, the large number of methyltransferases present in plants, many of which are involved in host defense responses, make cloning of specific genes a problem. Others have shown that changing only a few amino acids can alter the substrate specificity of methyltransferases (*18*). To achieve our goal of cloning the gene that encodes dHG-6-OMT, we were forced to sequence fragments of the purified enzyme, clone the gene, and express it in *E. coli*. Because dHG-6-OMT utilizes a unique substrate, we were able to unequivocally demonstrate authenticity of the cloned gene. This is one of the rare examples where cloning of a specific member of a large family of genes has been confirmed by sequence analysis and substrate specificity of the expressed protein.

Our work provides the basis to utilize antisense constructs to block methylation of dHG, thus altering biosynthesis of the methylated terpenoids. Preserving the more potent terpenoids by down-regulation of dHG-6-OMT could significantly increase resistance of *G. hirsutum* to important insect and fungal pests.

ABBREVIATIONS USED

dHG, desoxyhemigossypol; dHG-6-OMT, desoxyhemigossypol-6-*O*-methyltransferase; HG, hemigossypol; dMHG, desoxyhemigossypol-6-methyl ether; MHG, hemigossypol-6-methyl ether; HGQ, hemigossypolone; MHGQ, hemigossypolone-6-methyl ether; HB2, heliocide B2; HB3, heliocide B3; GSH, reduced glutathione; PVP, insoluble polyvinylpolypyrrolidone; COMT, caffeic acid 3-*O*-methyltransferase; CCOMT, caffeoyl CoA 3-*O*-methyltransferase.

LITERATURE CITED

(1) Bell, A. A. Formation of gossypol in infected or chemically irritated tissues of *Gossypium* spp. *Phytopathology* **1967**, *57*, 759−764.

(2) Alchanati, I.; Benedict, C. R.; Stipanovic, R. D. The enzymatic conversion of desoxyhemigossypol to desoxy methyl hemigossypol in cotton stems: dHG-*O*-methyltransferase. In *Proceedings of the Biochemical Cotton Workshop*; Jividen, G., Benedict, R. C., Eds.; Cotton Incorporated: Raleigh, NC, 1994; pp 35−39.

(3) Mace, M. E.; Stipanovic, R. D.; Bell, A. A. Toxicity and role of terpenoid phytoalexins in Verticillium wilt resistance in cotton. *Physiol. Plant Pathol.* **1985**, *26*, 209−218.

(4) Zhang, J.; Mace, M. E.; Stipanovic, R. D.; Bell, A. A. Production and fungitoxicity of the terpenoid phytoalexins in cotton. *J. Phytopathol.* **1993**, *139*, 247−252.

(5) Hedin, P. A.; Parrot, W. L.; Jenkins, J. N. Relationship of glands, cotton square terpenoid aldehydes and other allelochemicals to larvae growth of *Heliothis virescens (*Lepidoptera: Noctui). *J. Econ. Entomol.* **1992**, *85*, 359−364.

(6) Jenkins, J. N. Host resistance to insects in cotton. In *Proceedings of the World Cotton Research Conference−1, Challenging the Future*; Constable, G. A., Forrester, N. W., Eds.; CSIRO: Melbourne, Australia, 1995; pp 359−372.

(7) Stipanovic, R. D.; Bell, A. A.; Lukefahr, M. J. Natural insecticides from cotton. In *Host Plant Resistance to Pests*; ACS Symposium Series 62; Hedin, P. A., Ed.; American Chemical Society: Washington, DC, 1977; pp 197−214.

(8) Liu, J.; Benedict, C. R.; Stipanovic, R. D.; Bell, A. A. Purification and characterization of *S*-adenosyl-L-methionine:dexosyhemigossypol-6-O-methyltransferase from cotton plants. An enzyme capable of methylating the defense terpenoids of cotton. *Plant Physiol.* **1999**, *121*, 1017−1024.

(9) Stipanovic, R. D.; Bell, A. A.; Mace, M. E.; Howell, C. R. Antimicrobial terpenoids of *Gossypium*: 6-methoxygossypol and 6,6′-dimethoxygossypol. *Phytochemistry* **1975**, *14*, 1077−1081.

(10) Stipanovic, R. D.; Mace, M. E.; Bell, A. A.; Beier, R. C. The role of free radicals in the decomposition of the phytoalexin desoxyhemigossypol. *J. Chem. Soc., Perkin Trans. 1* **1992**, 3189−3192.

(11) Bradford, N. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1977**, *72*, 248−254.

(12) Laemmli, U. K. Cleavage of structural proteins during assembly of the head of bacteriophage T4. *Nature* **1970**, *227*, 680−685.

(13) Avis, H.; Leder, P. Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. *Proc. Natl. Acad. Sci. U.S.A.* **1972**, *69*, 1408−1412.

(14) Klein, R. R.; Crafts-Brandner, S. J.; Salvucci, M. E. Cloning of the sucrose-phosphate synthase gene from spinach. *Planta* **1993**, *190*, 498−510.

(15) Aarts, J. M. M. J. G.; Hontelez, J. G. J.; Fischer, P.; Verkerk, R.; Van Kammen, A.; Zabel, P. Acid phosphatase-1, a tightly linked molecular marker for root-knot nematode resistance in tomato: from protein to gene, using PCR and degenerate primers containing deoxyinosine. *Plant Mol. Biol.* **1991**, *16*, 647−661.

(16) Martin, F. H.; Castro, M. M.; Aboul-ela, F.; Tinoco, I. Base paring involving deoxyinosine: implications for probe design. *Nucleic Acids Res.* **1985**, *13*, 8927−8938.

(17) Ohtsuka, E.; Matsuki, S.; Ikehara, M.; Takahashi, Y.; Matusbara, K. An alternative approach to deoxyoligonucleotides as hybridization probes by insertion of deoxyinosine at ambiguous codon positions. *J. Biol. Chem.* **1985**, *260*, 2605−2608.

(18) Wang, J.; Pichersky, E. Identification of specific residues involved in substrate discrimination in two plant *O*-methyltransferases. *Arch. Biochem. Biophys.* **1999**, *368*, 172−180.

(19) Grimmig, B.; Kneusel, R. E.; Junghanns, K. T.; Matern, U. Expression of bifunctional caffeoyl-CoA 3-*O*-methyltransferase in stress compensation and lignification. *Plant Biol.* **1999**, *1*, 299−310.

(20) Ni, W.; Fahrendorf, T.; Balance, G. M.; Lamb, C. J.; Dixon, R. A. Stress responses in alfalfa (*Medicago sativa* L.). XX. Transcriptional activation of phenylpropanoid pathway genes in elicitor-induced cell suspension cultures. *Plant Mol. Biol.* **1996**, *30*, 427−438.

(21) Cui, Y.; Bell, A. A.; Joost, O.; Magill, C. W. Expression of potential defense response genes in cotton. *Physiol. Mol. Plant Pathol.* **2000**, *56*, 25−31.

(22) Murray, E. E.; Lotzer, J.; Eberle, M. Codon usage in plant genes. *Nucleic Acids Res.* **1989**, *17*, 447−498.

(23) Lutcke, H. A.; Chow, K. C.; Mickel, F. S.; Moss, K. A.; Kern, H. F.; Scheele, G. A. Selection of AUG initiation codons differs in plants and animals. *EMBO J.* **1987**, *6*, 43−48.
(24) Joshi, C. P. An inspection of the domain between putative TATA box and translation start site in 79 plant genes. *Nucleic Acids Res.* **1987**, *15*, 6643−6653.
(25) Joshi, C. P. Putative polyadenylation signals in nuclear genes of higher plants: a compilation and analysis. *Nucleic Acids Res.* **1987**, *15*, 9627−9640.
(26) Li, L.; Popko, J. L.; Zhang, X. H.; Osakabe, K.; Tsai, C. J.; Joshi, C. P.; Chiang, V. L. A novel multifunctional *O*-methyltransferase implicated in a dual methylation pathway associated with lignin biosynthesis in loblolly pine. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 5461−5466.
(27) Ibrahim, R. K. Plant *O*-methyltransferase signatures. *Trends Plant Sci.* **1997**, *2*, 249−250.
(28) Zubieta, C.; He, X.-Z.; Dixon, R. A.; Noel, J. P. Structures of two natural product methyltransferases reveal the basis for substrate specificity in plant *O*-methyltransferases. *Nat. Struct. Biol.* **2001**, *8*, 271−279.
(29) Vidgren, J.; Svensson, L. A.; Liljas, A. Crystal structure of catechol *O*-methyltransferase. *Nature* **1994**, *368*, 354−358.
(30) Maxwell, C. A.; Harrison, M. J.; Dixon, R. A. Molecular characterization and expression of alfalfa isoliquiritigenin 2′-*O*-methyltransferase, an enzyme specifically involved in the biosynthesis of an inducer of *Rhizobium meliloti* nodulation genes. *Plant J.* **1993**, *4*, 971−981.
(31) Takeshita, N.; Fujiwara, H.; Mimura, H.; Fitchen, J. H.; Yamada, Y.; Sato, F. Molecular cloning and characterization of *S*-adenosyl-L-methionine:scoulerine-9-*O*-methyltransferase from cultured cells of *Coptis japonica*. *Plant Cell Physiol.* **1995**, *36*, 29−36.
(32) Vernon, D. M.; Bohnert, H. J. A novel methyl transferase induced by osmotic stress in the faculative halophyte *Mesembryanthemum crystallinum*. *EMBO J.* **1992**, *6*, 2077−2085.

Received for review December 26, 2001. Revised manuscript received March 8, 2002. Accepted March 8, 2002. This work was supported in part by the Texas A&M Agricultural Experimental Station, Cotton Incorporated, and the U.S. Department of Agriculture, Agricultural Research Service.

JF011701Y