Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 11088–11092, November 1992
Plant Biology

Gene family for an elicitor-induced sesquiterpene cyclase in tobacco

(heterologous gene expression/isoprenoids/terpene cyclases)

PETER J. FACCHINI* AND JOSEPH CHAPPELL†

Plant Physiology/Biochemistry/Molecular Biology Program, Agronomy Department, University of Kentucky, Lexington, KY 40546-0091

Communicated by Eric E. Conn, August 3, 1992 (received for review June 8, 1992)

ABSTRACT **The initial step in the conversion of the isoprenoid intermediate farnesyl diphosphate to the sesquiterpenoid phytoalexin capsidiol in elicitor-treated tobacco tissues is catalyzed by an inducible sesquiterpene cyclase [5-*epi*-aristolochene synthase (EAS)]. Two independent cDNA clones (cEAS1 and cEAS2) encoding EAS were isolated from an elicitor-induced tobacco cDNA library by differential hybridization and subsequently were characterized by hybrid selection–*in vitro* translation. Insertion of cEAS1, a partial cDNA clone encoding 175 C-terminal amino acids, into an *Escherichia coli* expression vector resulted in accumulation of a fusion protein immunodetectable with EAS-specific polyclonal antibodies. The cDNA clones were used to isolate two full-length EAS genes that mapped 5 kilobases (kb) apart on one 15-kb genomic clone. The nucleotide sequences of the structural gene components were identical from 388 base pairs (bp) upstream of the transcription initiation site to 40 bp downstream of the translation termination codon, suggesting a relatively recent duplication event. The genes consist of 1479-bp open reading frames, each containing five introns and specifying 56,828-Da proteins. The N-terminal amino acid sequence deduced from the genomic clones was identical to the first 16 amino acids of the EAS protein identifiable by Edman degradation. RNA blot hybridization with cEAS1 demonstrated a mRNA induction time course consistent with the induction of the EAS mRNA translational activity with maximum levels 4–6 h after elicitation. EAS mRNA was not detected in control cells. DNA blot-hybridization analysis of genomic DNA revealed a copy number of ≈12–15 for EAS-like genes in the tetraploid tobacco genome. The conservation of a putative allelic prenyl diphosphate binding motif is also discussed.**

Sesquiterpenoids are a structurally diverse class of isoprenoids found in plants, some fungi, and bacteria, but not in vertebrates, which have important implications in plant–plant (1), plant–insect (2, 3), and plant–pathogen (4, 5) interactions. Despite an extensive appreciation for the structure–function relationships of plant sesquiterpenoids, the sesquiterpenoid biosynthetic pathway and its regulation are not well understood (4, 6). Although many enzymes of general isoprenoid biosynthesis have been measured in plants, few enzymes suspected of being rate-limiting have been identified (4, 6). Moreover, little is known about the mechanisms regulating the activity and absolute levels of plant isoprenoid biosynthetic enzymes with the exception of a diterpene cyclase (7). This is in contrast to the extensive molecular analysis of regulatory mechanisms controlling isoprenoid metabolism in mammalian systems (8).

Tobacco cell suspension cultures respond to treatment with elicitors such as fungal cell wall hydrolysates or cellulase by the *de novo* synthesis and secretion of antibiotic sesquiterpenoids, primarily the phytoalexin capsidiol (9, 10). Central to capsidiol biosynthesis is the induction of a sesquiterpene cyclase [5-*epi*-aristolochene synthase (EAS)] that catalyzes the cyclization of *trans,trans*-farnesyl diphosphate (*FPP*) to the bicyclic intermediate 5-*epi*-aristolochene (9). The induction of sesquiterpene cyclase enzyme activity and, hence, sesquiterpenoid biosynthesis has also been correlated with the suppression of squalene synthetase enzyme activity and consequent decline in sterol biosynthesis (9, 10). The induction of one enzyme and the suppression of the other are interpreted as an important regulatory mechanism controlling end-product formation, since the two enzymes are positioned at putative branch points in isoprenoid metabolism (9).

Tobacco EAS is a soluble, apparently cytoplasmic enzyme determined by SDS/PAGE to be a single polypeptide with a molecular mass of ≈60 kDa (11). The induction of EAS enzyme activity has been correlated with the absolute amount and the *de novo* synthesis of the enzyme protein (12). Changes in the *de novo* synthesis rate of the EAS protein were also correlated with changes in the EAS mRNA translational activity. Further investigation of the regulation of isoprenoid metabolism and in particular the sesquiterpenoid branch pathway in plants will require detailed molecular analysis of the EAS gene(s), especially those sequences involved in transcriptional regulation. Reported here is the isolation and characterization of cDNA and genomic clones‡ for EAS from tobacco.

MATERIALS AND METHODS

Cell Cultures, Elicitor Treatment, and Assays. Tobacco (*Nicotiana tabacum* L. cv KY14) cell suspension cultures in their rapid growth phase were treated with cellulase (*Trichoderma viride*, Type RS, Onozuka) as elicitor to a final concentration of 20 μg/ml (13). Control and elicitor-treated cells were collected by gentle vacuum filtration and used for measurements of EAS enzyme activity (12) and the level of EAS mRNA (14, 15).

Library Constructions and Screening. A cDNA library was constructed in pcDNAII (Invitrogen, San Diego) by using poly(A)$^+$ RNA from 4-h-elicitor-treated tobacco cells and was screened by differential hybridization (16) with radiolabeled first-strand cDNAs complementary to either control or elicitor-induced poly(A)$^+$ RNAs. Clones exhibiting preferential hybridization to elicitor-specific probe were further characterized by the hybrid selection–*in vitro* translation–immunoprecipitation technique (17). A putative EAS cDNA (cEAS1) identified as such was used to isolate additional cDNA and genomic clones. A λEMBL3 genomic library, constructed from *Mbo* I partially digested tobacco (*N.*

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: EAS, 5-*epi*-aristolochene synthase; IPTG, isopropyl 1-thio-β-D-galactoside; *FPP*, *trans,trans*-farnesyl diphosphate; TS, trichodiene synthase.

*Present address: Institut de recherche en biologie végétale, 4101 est rue Sherbrooke, Université de Montréal, Montréal, PQ, H1X 2B2, Canada.

†To whom reprint requests should be addressed.

‡The sequence reported in this paper has been deposited in the GenBank data base (accession number L04680).

tabacum L. cv NK326) hypocotyl DNA (Clontech) was screened according to a standard protocol (15). Genomic DNA fragments were subsequently inserted into pBluescript KS(+) for further characterization (15).

Isolation and Analysis of Nucleic Acids. Total RNA was prepared by the guanidine thiocyanate/CsCl method (14), and poly(A)$^+$ RNA was isolated by oligo(dT)-cellulose chromatography. Total RNA (20 μg) was fractionated on 1.0% agarose gels containing formaldehyde and transferred to nylon membranes (15). RNA blot hybridizations were performed with nick-translation-radiolabeled (15) cEAS1 as probe at 42°C in 50% (vol/vol) formamide/5× SSPE at pH 8.0 (1× = 0.15 M NaCl/10 mM NaH_2PO_4/1 μM EDTA, pH 8.0)/1× Denhardt's solution (0.02% polyvinylpyrrolidone/0.02% bovine serum albumin/0.02% Ficoll)/0.5% SDS/100 mg of denatured salmon sperm DNA per ml. Tobacco leaf genomic DNA was isolated (18), digested with various restriction endonucleases, electrophoresed on a 0.8% agarose gel, and transferred to nylon membranes. DNA blots were hybridized with random-primer-radiolabeled (15) cEAS1 at 60°C in 0.25 M sodium phosphate buffer, pH 8.0/7% SDS/1% bovine serum albumin/1 mM EDTA. Both DNA and RNA blots were washed at 45°C, twice with 2× SSC/0.1% SDS and twice with 0.2× SSC/0.1% SDS (15) (1× SSC = 0.15 M NaCl/0.015 M sodium citrate, pH 7). Relative mRNA transcript levels were estimated from RNA blot autoradiograms with a video densitometer (MilliGen/Biosearch, Ann Arbor, MI).

Fine Mapping of EAS Structural Genes. Intron junctions were mapped by nucleotide sequence comparison of cDNA and genomic clones. A 5′ intron not covered by the partial cDNA clones was identified by using a cDNA fragment amplified by reverse transcription–polymerase chain reaction (RT-PCR) with primers specific to exonic sequences on either side of the intron. The RT reaction was catalyzed with Moloney murine leukemia virus reverse transcriptase at 42°C by using elicitor-induced poly(A)$^+$ RNA primed with the antisense 3′ oligonucleotide (5′-TGAGTCCTTACATGTGAAGC-3′). First-strand cDNAs were amplified by 35 cycles of the PCR after addition of the sense 5′ primer (5′-AGAAATTGATGAGATTTTGG-3′) and adjustment of reaction conditions for *Taq* DNA polymerase. The purified 222-base pair (bp) PCR product was sequenced directly by using the same primers.

The transcription initiation site was identified by S1 nuclease protection analysis performed with elicitor-induced poly(A)$^+$ RNA and PCR-amplified double-stranded DNA probe (15). The PCR primers were specific to genomic sequences located 106 bp downstream (5′-CCAAAATCTCATCAATTTCT-3′) and 282 bp upstream (5′-TTGCAAACTATGAAGAAGAG-3′) of the ATG codon corresponding to the predicted N-terminal methionine. The primers were 5′-end-labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP prior to use in the PCR.

Nucleotide and Amino Acid Sequencing. Sequencing of single- and double-stranded DNAs was by the dideoxynucleotide chain-termination method (19) with Sequenase (United States Biochemical) or an automated fluorescence-based system (Applied Biosystems) or both. Both strands of insert DNAs were sequenced from their ends using T7 and T3 primers. In addition, 20-mer oligonucleotides synthesized according to obtained sequence information were used directly as primers for further sequencing.

N-terminal amino acids of the EAS protein were obtained by Edman degradation on a pulse-liquid-phase sequenator (Applied Biosystems). Purified EAS protein was cleaved with cyanogen bromide, and the products were separated by SDS/PAGE and then transferred to a poly(vinylidene difluoride) membrane. A Coomassie blue-staining band of ≈20 kDa was used for sequencing.

Expression of Partial EAS cDNA in *Escherichia coli*. Two oligonucleotides (5′-GGGGATCCCTAAAGGAAGTAGTAAGAAATT-3′ and 5′-GGGGAATTCCGGTCCTAGTTAGGAAGTCCA-3′) specific to cEAS1 and designed to include 5′ *Bam*HI and 3′ *Eco*RI restriction sites, respectively, were used to prime the PCR with cEAS1 as template. The amplified 700-bp product was inserted into the *Bam*HI/*Eco*RI sites of the expression vector pGEX-2T (Pharmacia). *E. coli* JM101 cells harboring the pGEX-EAS1 construct were grown in Luria–Bertani medium to OD_{600} = 0.6, and expression of the fusion protein was induced by addition of isopropyl 1-thio-β-D-galactoside (IPTG) to a final concentration of 1 mM. Cells were collected 1 h after addition of IPTG. Proteins were solubilized in SDS sample buffer, separated by SDS/PAGE, transferred to nitrocellulose, and immunodetected as described by Leary *et al.* (20).

RESULTS

Isolation and Characterization of Clones. Differential hybridization (15) and hybrid selection–*in vitro* translation (17) techniques were used to isolate a cDNA clone complementary to the EAS mRNA. The *in vitro* translational activity of EAS mRNA has been detected only in RNA preparations from elicitor-treated cells, suggesting the presence of EAS transcripts exclusively in induced mRNA populations (12). Hence, a cDNA library was prepared to poly(A)$^+$ RNA from 4-h-elicitor-induced tobacco cell cultures. Clones exhibiting preferential hybridization to the radiolabeled first-strand cDNA probe synthesized from the induced mRNA pool, versus the probe prepared to control mRNAs, were subsequently used for RNA blot hybridizations. Twelve of 46 clones hybridized to an elicitor-induced mRNA with an induction time course consistent with the EAS mRNA translational activity and did not hybridize to control mRNAs. Nine clones that did not cross-hybridize to one another were further characterized by hybrid selection–*in vitro* translation. One cDNA clone (cEAS1) hybrid-selected a mRNA from the population of mRNAs isolated from elicitor-treated cells which, when translated *in vitro* in the presence of [^{35}S]methionine, directed the synthesis of an ≈60-kDa protein immunoprecipitable with EAS-specific antibodies. The cDNA library was rescreened with cEAS1 as probe, and a second cDNA (cEAS2) was recovered (Fig. 1*C*). Nucleotide sequence analysis showed that the two partial cDNAs were independent transcripts.

A tobacco genomic library in λEMBL3 was screened by using cEAS1, and eight independent genomic clones were isolated. Restriction endonuclease analysis suggested that each genomic clone represented a different gene. One 15-kb genomic clone was selected for further characterization because the cDNA probe hybridized to two locations on the DNA fragment separated by ≈6 kb (Fig. 1*A*). Subsequent analysis indicated that the two regions contained 2517 bp of identical nucleotide sequence, which were identified as duplicated EAS genes (gEAS3 and gEAS4). The nucleotide and predicted amino acid sequences of gEAS3 are shown in Fig. 2. Each gene consists of an open reading frame of 1479 bp divided among six exonic segments and interrupted by five introns with a range in size of 75–155 bp (Figs. 1*B* and 2).

The two structural genes (gEAS3 and gEAS4) exhibited 100% nucleotide sequence identity in both coding and noncoding regions, from 388 bp upstream of the transcription-initiation site to 40 bp downstream of the translation-termination codon. The partial cDNAs (cEAS1 and cEAS2) each exhibited 94% and 95% identity to the structural genes and 95% and 93% identity relative to each other at the nucleotide and amino acid levels, respectively, to a position 40 bp downstream of the termination codon. The nucleotide sequences of all four genes beyond the conserved 40 bp 3′ to

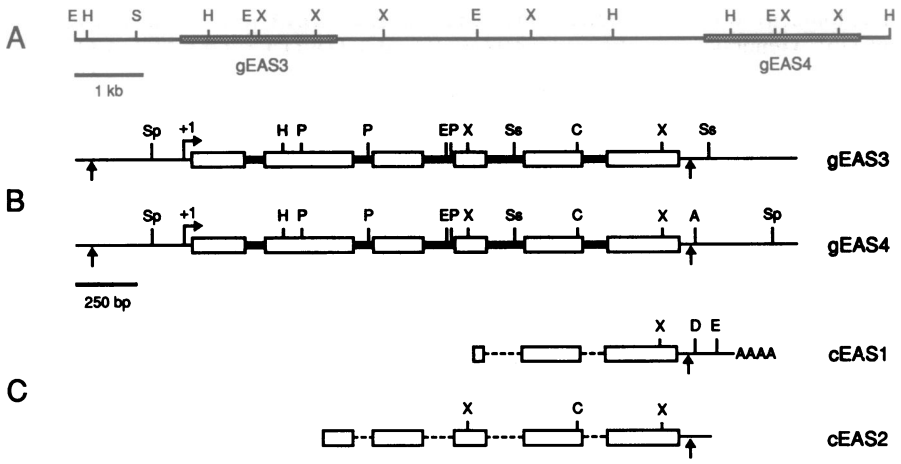

FIG. 1. Restriction and structural maps for EAS genomic and cDNA clones. (*A*) Restriction map of the cloned 12-kilobase (kb) genomic fragment showing the location of two independent structural genes (shaded boxes). (*B*) Structural maps of cyclase genes indicating sequence divergence at the 3′ termini. Coding regions are represented by open boxes with introns shown as thick lines. Horizontal arrows refer to transcription start sites; vertical arrows indicate the start and stop points for homology among the four genes. (*C*) Structural maps of partial cDNAs. Dashed lines indicate putative intron locations. A polyadenylate tract was found on only one clone as indicated. Restriction sites are as follows: A, *Acc* I; C, *Cla* I; D, *Dra* I; E, *Eco*RI; H, *Hin*dIII; P, *Pst* I; S, *Sal* I; Sp, *Spe* I; Ss, *Sst* I; X, *Xba* I.

the translation stop codon were completely divergent (Fig. 1). A putative "TATA" box was identified 25 bp upstream of the transcription start site (Fig. 2). Potential polyadenylylation signals were also observed in the region of gEAS3 3′ to the stop codon. Similar signals were detected in the divergent region of gEAS4.

```
CCGTCGCCGACTTCTCCCCTAGTCTCTGGGGTGATCAGTTCCTTTCATTCTCCATTGATAATCAGGTAATTTAACTAAT  -134
ACTAGTATTCTTTATTTATAGTTTGTTCTCCATTGATAATCAGGTAGTTTATTTATGTTGAACAACATTAATT       -55
                                                                    >>>
TTGCTAATTTCAGTTTAATGTACATTACATATAGGTTGCGGAAAAGTATATATATGCTCAAGAGATTGAAGCATTGAAG   25
GAACAAACGAGGAGT ATG CTG TTA GCA ACC GGA AGG AAA TTG GCC GAT ACA TTG AAT TTG ATT    85
                M  L  L  A  T  G  R  K  L  A  D  T  L  N  L  I
GAC ATT ATT GAA CGC CTT GGT ATA TCC TAC CAC TTT GAG AAA GAA ATT GAT GAG ATT TTG   145
 D  I  I  E  R  L  G  I  S  Y  H  F  E  K  E  I  D  E  I  L
GAT CAG ATT TAC AAC CAA AAC TCA AAC TGC AAT GAT TTG TGC ACC TCT GCA CTT CAA TTT   205
 D  Q  I  Y  N  Q  N  S  N  C  N  D  L  C  T  S  A  L  Q  F
CGA TTG CTC AGG CAA CAC GGT TTC AAC ATC TCT CCT  GGTAAGTTCATCATGAAGTTGTTAAAATTA   271
 R  L  L  R  Q  H  G  F  N  I  S  P
TTATCCATTTATTGGAAGAAGGCTAATTCATCTTGAGTTTTCTTTCTTGAAATACCA    GAA ATT TTC AGC AAA  343
                                                              E  I  F  S  K
TTC CAA GAT GAA AAT GGC AAA TTC AAG GAG TCT CTT GCT AGT GAT GTC TTA GGA TTA TTA   403
 F  Q  D  E  N  G  K  F  K  E  S  L  A  S  D  V  L  G  L  L
AAC TTG TAT GAA GCT TCA CAT GTA AGG ACT CAT GCT GAC GAT ATC TTA GAA GAC GCA CTT   463
 N  L  Y  E  A  S  H  V  R  T  H  A  D  D  I  L  E  D  A  L
GCT TTC TCC ACT ATC CAT CTT GAA TCT GCA GCT CCA CAT TTG AAA TCT CCA CTT AGG GAG   523
 A  F  S  T  I  H  L  E  S  A  A  P  H  L  K  S  P  L  R  E
CAA GTG ACA CAT GCC CTT GAG CAA TGT TTG CAC AAG GGT GTT CCT AGA GTC GAG ACC CGA   583
 Q  V  T  H  A  L  E  Q  C  L  H  K  G  V  P  R  V  E  T  R
TTC TTC ATC TCA TCA ATC TAT GAC AAG GAA CAA TCG AAG AAT AAT GTG TTA CTT CGA TTT   643
 F  F  I  S  S  I  Y  D  K  E  Q  S  K  N  N  V  L  L  R  F
GCC AAA TTG GAT TTC AAC TTG CTC CAG ATG TTG CAC AAA CAA GAA CTT GCT CAA GTA TCA   703
 A  K  L  D  F  N  L  L  Q  M  L  H  K  Q  E  L  A  Q  V  S
AGG TG CGATATATAAAAACGATGAACCCTTTTTGATTCATCATATCTCAAGTACTCATGTTAATTTCTTATGCTGCA   780
 R  W
GGTG  G TGG AAA GAT TTG GAT TTT GTA ACA ACA CTT CCA TAT GCT AGA GAT CGA GTA GTT   839
        W  K  D  L  D  F  V  T  T  L  P  Y  A  R  D  R  V  V
GAA TGC TAC TTT TGG GCA TTA GGA GTT TAT TTT GAG CCT CAA TAC TCT CAA GCT CGC GTC   899
 E  C  Y  F  W  A  L  G  V  Y  F  E  P  Q  Y  S  Q  A  R  V
ATG CTC GTT AAG ACC ATA TCA ATG ATT TCG ATT GTC GAT GAC ACC TTT GAT GCT TAC GGT   959
 M  L  V  K  T  I  S  M  I  S  I  V  D  D  T  F  D  A  Y  G
ACA GTT AAA GAA CTT GAG GCA TAC ACA GAT GCC ATA CAA AGG T ATGAACTTCATCAATTCACTT  1023
 T  V  K  E  L  E  A  Y  T  D  A  I  Q  R
ATTCCTTGATAGTGAATGTCGTCGTGAAAAGATTAAGACGAATTTCTACTCATTATAGTTGTGCTCTTTCAAAATGCAT  1102

GAATTCACCTTAATTTTGTCATCCTGCAGAT  GG GAT ATC AAC GAA ATT GAT CGG CTT CCT GAT TAC  1168
                                 W  D  I  N  E  I  D  R  L  P  D  Y
ATG AAA ATC AGT TAT AAA GCT ATT CTA GAT CTC TAC AAG GAT TAT GAA AAG GAA TTG TCT  1228
 M  K  I  S  Y  K  A  I  L  D  L  Y  K  D  Y  E  K  E  L  S
AGT GCC GGA AGA TCT CAT ATT GTC TGC CAT GCA ATA GAA AGA  GTATGTTCGAGCAACTTAACTA  1292
 S  A  G  R  S  H  I  V  C  H  A  I  E  R
TCGAAATACATTTTTTCCTTAATCCATTTCTCACTTTGGTTTACCTTGTGTTCGTCTTTTAGTGATTAGAAACTTGATA  1371

CAGTTCAATCAATATTTTCTAACACTTGAACACATATATGTTTTGTATTCACAG  ATG AAA GAA GTA GTA AGA  1443
                                                        M  K  E  V  V  R
AAT TAT AAT GTC GAG TCA ACA TGG TTT ATT GAA GGA TAT ATG CCA CCT GTT TCT GAA TAC  1503
 N  Y  N  V  E  S  T  W  F  I  E  G  Y  M  P  P  V  S  E  Y
CTA AGC AAT GCA CTA GCA ACT ACC ACA TAT TAC TAC CTC GCG ACA ACA TCG TAT TTG GGC  1563
 L  S  N  A  L  A  T  T  T  Y  Y  Y  L  A  T  T  S  Y  L  G
ATG AAG TCT GCC ACG GAG CAA GAT TTT GAG TGG TTG TCA AAG AAT CCA AAA ATT CTT GAA  1623
 M  K  S  A  T  E  Q  D  F  E  W  L  S  K  N  P  K  I  L  E
GCT AGT GTA ATT ATA TGT CGA GTT ATC GAT GAC ACA GCC ACG TAC GAG GTATGATTTGCATCT  1686
 A  S  V  I  I  C  R  V  I  D  D  T  A  T  Y  E
CAAGAAATTATATCATTATATGGGATTTGGACAAACAAAGTGTTGCGACGACAATTAAGGCAATATAAAAGCTAACCTT  1765

TAATTTATCTGCTTTCTAG GTT GAG AAA AGC AGG GGA CAA ATT GCA ACT GGA ATT GAG TGC TGC  1829
                    V  E  K  S  R  G  Q  I  A  T  G  I  E  C  C
ATG AGA GAT TAT GGT ATA TCA ACA AAA GAG GCA ATG GCT AAA TTT CAA AAT ATG GCT GAG  1889
 M  R  D  Y  G  I  S  T  K  E  A  M  A  K  F  Q  N  M  A  E
ACA GCA TGG AAA GAT ATT AAT GAA GGA CTT CTT AGG CCC ACT CCC GTC TCT ACA GAA TTT  1949
 T  A  W  K  D  I  N  E  G  L  L  R  P  T  P  V  S  T  E  F
TTA ACT CCT ATT CTC AAT CTT GCT CGT ATT GTT GAG GTT ACA TAT ATA CAC AAT CTA GAT  2009
 L  T  P  I  L  N  L  A  R  I  V  E  V  T  Y  I  H  N  L  D
GGA TAC ACT CAT CCG GAG AAA GTC TTA AAA CCT CAC ATT ATT AAC CTA CTT GTG GAC TCC  2069
 G  Y  T  H  P  E  K  V  L  K  P  H  I  I  N  L  L  V  D  S
ATC AAA ATT TGA GCTGCCATTTGTTGCTCATCTCAAGGAAACTTCATTCTTCTTTGTTGTATAATTACGTTGTTA  2144
 I  K  I  Z
GGGAGCTGTTGTGCAGTGCACTTCCTAAACTAGGACCTTCTTAAGATCCTTGTAATCACAATATTTTGTTCTCGATATT  2223
CATACCTGAATATACTTGTTTAACACTTCAAATGATAGAAATTAAAGTTCATCCGATTTCACTTCTGTGCTATCAGCAT  2302
ATATACATTTTTTCATTCCGTCCATTTTCTCCTCTTTGTAATTTCATGTATATATCTAGGATGTAATAATAACTCTAAA  2381
CAACAGTTTTATTTGGGATTTTTATCAAGCCATACACTATATGAAACCTTTTTACCCTCCCTACTCAAGTTTCA       2460
```

FIG. 2. Nucleotide and predicted amino acid sequence in single-letter code of the tobacco EAS gene (gEAS3). N-terminal amino acids obtained by Edman degradation of the purified protein are underlined. A putative "TATA" box is indicated with a wavy line; the transcription initiation site is marked overhead by >>>; potential polyadenylylation signals 3′ to the stop codon are underlined.

The predicted translation product initiated at the first in-frame ATG following the transcription start site has a molecular mass of 56.8 kDa, consistent with the value of 60 kDa determined for the EAS protein by SDS/PAGE (11). Additional evidence that the gene encodes the previously purified EAS protein (11) was obtained by comparing the empirically determined N-terminal amino acid sequence of EAS with the predicted sequence as shown in Fig. 2. Sixteen of the first 18 amino acids were identifiable by Edman degradation and corresponded with the predicted amino acid at each position.

Expression of EAS Polypeptide in *E. coli*. PCR amplification of the partial cDNA insert of cEAS1 using primers with specifically adapted 5′ *Bam*HI and 3′ *Eco*RI restriction sites allowed for ligation of the fragment into the open reading frame of the human glutathione *S*-transferase gene contained on the *E. coli* expression vector pGEX-2T. The 700-bp PCR product contains a termination codon 525 bp from the introduced 5′ *Bam*HI restriction site and encodes 175 amino acids at the C terminus of the EAS protein. Cell homogenates of *E. coli* JM101 transformed with the recombinant plasmid pGEX-EAS1 and induced with IPTG contained a 45-kDa fusion protein immunodetectable with mouse polyclonal antibodies prepared to purified EAS (Fig. 3). Only trace amounts of the polypeptide were detectable in uninduced cultures. Untransformed cultures and those transformed with the nonrecombinant pGEX-2T plasmid did not produce any proteins detectable with the EAS-specific antibodies.

Induction of EAS mRNA in Response to Elicitor. The induction of EAS mRNA was monitored directly by hybrid-

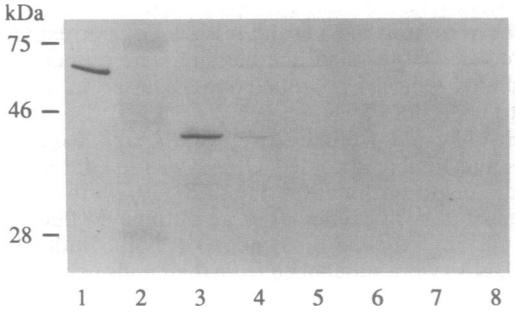

FIG. 3. Immunoblot detection of an EAS polypeptide fusion protein in cell homogenates of IPTG-induced *E. coli* transformants harboring the pGEX-EAS1 construct. Lanes: 1, purified EAS protein from elicitor-treated tobacco cell cultures; 2, molecular weight standards; 3 and 4, IPTG-induced (lane 3) and uninduced (lane 4) JM101 cells transformed with pGEX-EAS1; 5 and 6, IPTG-induced (lane 5) and uninduced (lane 6) JM101 cells transformed with nonrecombinant pGEX-2T; 7 and 8, IPTG-induced (lane 7) and uninduced (lane 8) nontransformed JM101 cells.

ization of the radiolabeled cDNA insert from cEAS1 to total RNA isolated from control and cellulase-treated cultures at various times after addition of elicitor (Fig. 4*A*). No EAS transcripts were detected in control RNA. Elicitor treatment resulted in the appearance of detectable levels of EAS mRNA in less than 4 h, with a maximum occurring 6 h after elicitation, followed by a subsequent gradual decline (Fig. 4*B*). EAS enzyme activity was also induced in elicitor-treated cells, but the rate of increase in enzyme activity was less than that of the EAS mRNA. Although the level of EAS mRNA was transient with a maximum at 6 h and a return to control levels within 24 h, EAS enzyme activity was detected for more than 48 h after elicitation.

Gene Family for EAS in Tobacco. Tobacco genomic DNA digested with either *Eco*RI or *Hin*dIII and probed with radiolabeled cEAS1 revealed 12 and 11 bands, respectively, suggesting the existence of numerous EAS-like genes (Fig. 5). In contrast, only one predominant band appeared in *Xba* I digests (Fig. 5). The size of the fragment (800 bp) is consistent with the separation distance of the two *Xba* I sites on the EAS genomic clones (Fig. 1). The predominant *Xba* I band was estimated by video densitometry to be 15 times more intense than the average band intensity in *Hin*dIII- or *Eco*RI-digested DNA, confirming a copy number of ≈12–15. The single intense *Xba* I band also suggests a high degree of structural similarity among all EAS genes in tobacco. Secondary bands apparent in the *Xba* I-digested DNA may represent hybridization of the cEAS1 probe to genomic fragments containing mostly 3′ untranslated gene-specific sequences. The divergent sequences on the four EAS cDNA and genomic clones begin 120 bp downstream of the 3′ *Xba* I site, a feature apparently common to all EAS genes (Fig. 1).

DISCUSSION

We report the isolation and characterization of genes encoding the sesquiterpene cyclase EAS from tobacco. The gene

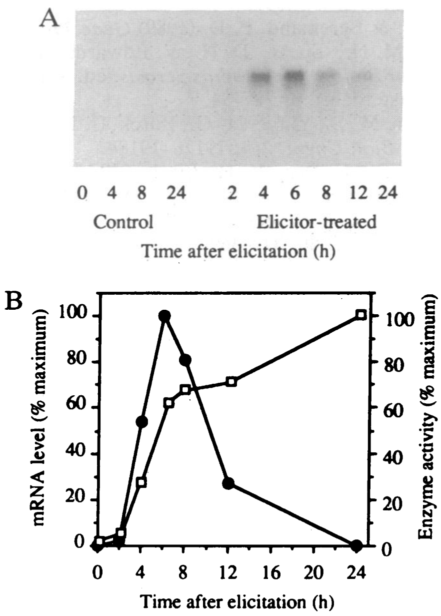

FIG. 4. RNA gel blot analysis of EAS gene expression in elicitor-treated tobacco cell suspension cultures. Total RNAs from cells at various times after treatment were separated on a denaturing agarose gel, transferred to a nylon membrane, and probed with cEAS1. (*A*) Autoradiogram of the RNA blot. (*B*) Time course of elicitor-induced EAS mRNA accumulation (●) and enzyme activity (□). EAS mRNA levels were quantified by video densitometry. EAS enzyme activity was measured in a fraction of the cells used for RNA extraction. Maximum specific EAS enzyme activity was 16.3 nmol of product per h per mg of protein.

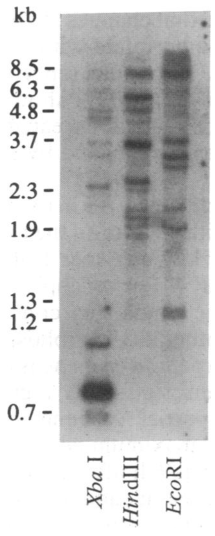

FIG. 5. DNA gel blot analysis of the EAS-like genes in tobacco. Genomic DNA (5 μg) was digested with *Eco*RI, *Hin*dIII, or *Xba* I; separated by electrophoresis; transferred to a nylon membrane; and probed with cEAS1 at high stringency.

for a fungal sesquiterpene cyclase, trichodiene synthase (TS), from *Fusarium sporotrichioides* has also been characterized (21). Although both enzymes use FPP as the sole substrate, comparison of the predicted amino acid sequences using the FASTDB search algorithm did not reveal significant overall homology between these proteins or any other entry in the Protein Identification Resource (May 1992) or Swiss-Prot (May 1992) protein sequence data bases. However, several other proteins that use F*PP* or related prenyl substrates contain aspartate-rich consensus sequences proposed as binding sites for the charged head group of these substrates (22). Sequences in both EAS and TS containing the putative consensus motif Asp-Asp-Xaa-Xaa-Asp were aligned with the proposed allylic prenyl diphosphate binding domains of F*PP* synthase (23) and squalene synthetase (24) from yeast (Fig. 6). EAS and TS exhibit 33% identity relative to each other and 33% and 23% identity, respectively, relative to F*PP* synthase among 18 amino acid residues that include the Asp-Asp-Xaa-Xaa-Asp motif. The apparent dissimilarity of the squalene synthetase domain has been suggested to reflect the unique head-to-head (1′–1) condensation catalyzed by squalene synthetase, in contrast to the more common head-to-tail (1′–4) condensations catalyzed by other prenyl binding proteins (24). The moderate sequence conservation of the EAS and TS aspartate-rich domains relative to the proposed consensus sequence of F*PP* synthase may indicate a similar substrate binding function. Additional experimentation, including site-directed mutagenesis, is necessary to further evaluate the function of this domain.

Transcriptional control of EAS enzyme activity was suggested previously by thiouridine-labeling experiments, which measured the *in vitro* translational activity for EAS mRNA in *de novo* synthesized RNA populations from elicitor-treated and control cells (12). EAS mRNA translational activity was detected only in RNAs from elicited cells with a maximum 4–6 h after addition of elicitor. The induction patterns for the steady-state level of EAS transcripts (Fig. 4) and the EAS mRNA translational activity (12) are identical. The minimal time lag for the induction of EAS enzyme activity subsequent to the increase in mRNA level (Fig. 4*B*) is also in agreement

```
EAS    239 I S M I S I V D D T F D A Y G T V K 256
TS     228 M S F Y K E F D D E R D Q I S L V K 246
Y-FPS  233 G E Y F Q I Q D D Y L D C F G T P E 251
Y-SS   215 G L F L Q K T N I I R D Y N E D L V 233
```

FIG. 6. Alignment of a proposed allylic prenyl diphosphate binding motif (22) in EAS, TS, yeast F*PP* synthase (Y-FPS), and yeast squalene synthetase (Y-SS). Amino acids asparagine (N) and aspartate (D) are considered a conservative substitution (24).

with earlier results, which showed that changes in the *de novo* synthesis rate of the EAS protein correlate with changes in mRNA translational activity (12). The results reported herein together with those reported previously (12) show that the induction of EAS enzyme activity in elicitor-treated tobacco cell suspension cultures is primarily regulated by transcriptional control of the EAS gene.

The rapid transcriptional induction of EAS in tobacco after elicitor treatment is consistent with reports on the expression pattern of genes encoding other key enzymes from diverse phytoalexin biosynthetic pathways. For example, in castor bean (*Ricinus communis* L.), the single step conversion of geranylgeranyl diphosphate, another allylic diphosphate isoprenoid intermediate, to the diterpenoid casbene is catalyzed by the elicitor-inducible and transcriptionally regulated enzyme casbene synthetase. A partial casbene synthetase cDNA was used to demonstrate maximum accumulation of transcripts in castor bean seedlings 6 h after treatment with a pectic elicitor (7). In addition, the accumulation of antibiotic furanocoumarins by members of the Leguminosae is dependent on the transcriptional induction of two important enzymes (25), phenylalanine ammonia lyase and chalcone synthase. Recently, the biosynthesis of cytotoxic benzophenanthridine alkaloids in the Papaveraceae and Fumariaceae was also correlated with elicitor-mediated transcriptional regulation of the berberine bridge enzyme (26).

The detection of multiple bands in *Eco*RI- or *Hin*dIII-digested genomic DNA (Fig. 5), the isolation of numerous independent genomic and cDNA clones (Fig. 1), and the existence of two genes (gEAS3 and gEAS4) that map only 5 kb apart (Fig. 1) suggest that the EAS genes in tobacco are a partially clustered gene family comprised of ≈12–15 members. The duplication of genes for important rate-determining enzymes of phytoalexin biosynthesis has also been observed for phenylalanine ammonia lyase (27, 28) and chalcone synthase (29). Although the gene for chalcone synthase is present in only a single copy in parsley (*Petroselinum crispum*) and other nonlegume species (30), a gene family of six to eight partially clustered members has been identified in french bean (*Phaseolis vulgaris*) (28) and another legume, soybean (*Glycine max*) (25). Several of the chalcone synthase genes are transiently activated by elicitor treatment or illumination with UV light; however, a small set are expressed after exposure of etiolated hypocotyls to visible light. The divergent amino acid sequences of the predicted translation products of the EAS cDNAs and structural genes reported herein indicate the existence of multiple EAS isozymes in tobacco. This genetic polymorphism may position the enzymatic isoforms in separate temporal-, spatial-, or stress-specific regulatory networks to provide the optimum response to environmental cues.

We thank Drs. Arthur Hunt, John Shaw, and George Cheniae for critical review of the manuscript. We are grateful to Mike Russ at the University of Kentucky Macromolecular Structure Analysis Facility for the automated DNA sequencing. This work was supported by the Kentucky Agricultural Experiment Station, a National Science Foundation grant (DCB-9106629), and a Natural Sciences and Engineering Research Council of Canada Postdoctoral Fellowship to P.J.F. This is journal paper 92-3-151 from the Kentucky Agricultural Experiment Station.

1. Kalsi, P. S., Goyal, R., Talwar, K. K. & Chhabra, B. R. (1989) *Phytochemistry* **28**, 2093–2096.
2. Barnby, M. A. & Kloke, J. A. (1990) *J. Insect Physiol.* **36**, 125–131.
3. Gibson, R. W. & Pickett, J. A. (1983) *Nature (London)* **302**, 608–609.
4. Cane, D. E. (1981) in *Biosynthesis of Isoprenoid Compounds*, eds. Porter, J. W. & Spurgeon, S. L. (Wiley, New York), pp. 283–374.
5. Stoessl, A., Stothers, J. B. & Ward, E. W. B. (1976) *Phytochemistry* **15**, 855–872.
6. Gray, J. C. (1987) *Adv. Bot. Res.* **14**, 25–91.
7. Lois, A. F. & West, C. A. (1990) *Arch. Biochem. Biophys.* **276**, 270–277.
8. Goldstein, J. L. & Brown, M. S. (1990) *Nature (London)* **343**, 425–430.
9. Vögeli, U. & Chappell, J. (1988) *Plant Physiol.* **88**, 1291–1296.
10. Threlfall, D. R. & Whitehead, I. M. (1988) *Phytochemistry* **27**, 2567–2580.
11. Vögeli, U., Freeman, J. W. & Chappell, J. (1990) *Plant Physiol.* **93**, 182–187.
12. Vögeli, U. & Chappell, J. (1990) *Plant Physiol.* **94**, 1860–1864.
13. Chappell, J. & Nable, R. (1987) *Plant Physiol.* **85**, 469–473.
14. Glisin, V., Crkvenjakow, R. & Byus, C. (1974) *Biochemistry* **13**, 2633–2637.
15. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), 2nd Ed.
16. Hanahan, D. & Meselson, M. (1980) *Gene* **10**, 63–67.
17. Alwine, J. C., Kemp, D. J., Parker, B. A., Reiser, J., Renart, J., Stark, G. R. & Wahl, G. M. (1979) *Methods Enzymol.* **68**, 220–242.
18. Murray, M. & Thompson, W. F. (1980) *Nucleic Acids Res.* **8**, 4321–4325.
19. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **83**, 8073–8076.
20. Leary, J. J., Brigati, D. & Ward, D. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 4045–4049.
21. Hohn, T. & Beremand, P. D. (1989) *Gene* **79**, 131–138.
22. Ashby, M. N., Spear, D. H. & Edwards, P. A. (1990) in *Molecular Biology of Antherosclerosis*, ed. Attie, A. D. (Elsevier, Amsterdam), pp. 27–34.
23. Anderson, M. S., Yarger, J. G., Burck, C. L. & Poulter, C. D. (1989) *J. Biol. Chem.* **264**, 19176–19184.
24. Jennings, S. M., Tsay, Y. H., Fisch, T. M. & Robinson, G. W. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 6038–6042.
25. Hahlbrock, K. & Scheel, D. (1989) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **40**, 347–369.
26. Dittrich, H. & Kutchan, T. M. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 9969–9973.
27. Edwards, K., Cramer, C. L., Bolwell, G. P., Dixon, R. A., Schurch, W. & Lamb, C. J. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 6731–6735.
28. Cramer, C. L., Edwards, K., Dron, M., Liang, X., Dildire, S. L., Bolwell, G. P., Dixon, R. A., Lamb, C. J. & Schurch, W. (1989) *Plant Mol. Biol.* **12**, 367–383.
29. Ryder, T. B., Hedrick, S. A., Bell, J. N., Liang, X., Clouse, S. D. & Lamb, C. J. (1987) *Mol. Gen. Genet.* **210**, 219–233.
30. Herrmann, A., Schulz, W. & Hahlbrock, K. (1988) *Mol. Gen. Genet.* **212**, 93–98.