Southern blot analyses of genomic DNA fragments suggest there are five different classes of phenylalanine ammonia-lyase (PAL, EC 4.3.1.5) genomic sequence in tomato (Lycopersicon esculentum). Isolation and subsequent sequence analysis of three examples from genomic libraries reveal highly homologous coding sequences but also a surprisingly high frequency of single point mutations which would truncate protein synthesis. The nucleotide sequence for one of the genes (PAL5) encodes a normal polypeptide of 721 amino acids, interrupted by a 710-base pair intron in the codon for amino acid 139. In contrast, premature stop codons, 363 triplets from the end in PAL1 and 304 triplets from the end in PAL3 would result in substantially (51-43%) shorter polypeptides that are consistent with the protein polymorphism, recently reported in alfalfa (Jorrin, J., and Dixon, R. A. (1990) Plant Physiol. 90, 447-445) but ascribed to protein degradation. S, mapping of the mRNA termini and polymerase chain reaction analysis of cDNA transcripts indicate that at least one of these truncated coding sequences is expressed, strongly suggesting that at least some of the shorter polypeptides constitute original gene products with a potentially important function.

Because of its important roles in plant development and pathogen defense phenylalanine ammonia-lyase (PAL, EC 4.3.1.5) has become one of the most studied enzymes with respect to plant secondary metabolism (for review see Dixon et al. (1983) and Jones (1984)). Lamb and his co-workers (Edwards et al., 1985) first isolated an encoding sequence as a cDNA clone from elicitor-treated bean cells. Based on this probe, gene isolations have been successful in four plant species, arabidopsis (Ohl et al., 1990), bean (Cramer et al., 1989), parsley (Lois et al., 1989), and rice (Minami et al., 1989). In each case, a family of at least four genes has been identified. Furthermore, at least four forms of active phenylalanine ammonia-lyase containing subunit variants have also been identified (Bolwell et al., 1986). A recent, more detailed analysis of this polymorphism in alfalfa (Jorrin and Dixon, 1990) indicated that the native enzyme was a tetramer of MW 311,000 with a subunit of MW 79,000. Other polypeptides of MW 71,000, 67,000, and 56,000 also were observed but were considered to be products of degradation. Because PAL is an important activity with respect to environmental stress responses including UV protection and pathogen defense (Dixon et al., 1983; Hahlbrock and Grisebach, 1979; Jones, 1984), the regulation of this gene family is also being actively examined. For example, studies in parsley have indicated both tissue-specific and stress-specific regulation (Loise et al., 1989).

In the course of studies on the role of PAL in Verticillium infections, we have been examining the PAL gene family in near-isolines of tomato which differ in their resistance to this pathogen. Our studies indicate that a fungal pathogen may gain access to this host not simply by passively avoiding defense mechanisms but also by actively inhibiting them through, for example, a suppression of PAL gene expression (Lee et al., 1992). To examine this possibility further we are isolating and characterizing the tomato PAL gene family. The present study indicates that heterogeneity in the PAL protein family not only results from differences in the gene sequences but also differences in protein size due to truncated gene expression.

**MATERIALS AND METHODS**

**Plant Materials—**Tomato plants (Lycopersicon esculentum cv. Bonny Best) were grown from seed in a growth chamber. For RNA extraction, 5-6-week-old tomato leaves were cut in half to induce PAL gene transcripts. After 12 h wounded leaves were harvested and immediately frozen in liquid nitrogen. For DNA extraction, 5-6-week-old plants were etiolated for 1-2 days to reduce the starch content prior to DNA extraction.

**Hybridization Analyses—**Genomic DNA was prepared using a method similar to that described by Vallejos et al. (1986). Ten to 20 g of finely chopped leaf tissue were homogenized with an Omni-mixer homogenizer at speed 5 (Serval Inc., Newton, CT) for 20-30 s with 100-200 ml of homogenization buffer (100 mM Mes, pH 5.5, 2.5 mM dithiothreitol, 2.5 mM EDTA, 10 mM NaCl, 10 mM KCl, and 0.25 M sucrose). The homogenate was filtered through four layers of cheesecloth and one layer of miracloth (Calbiochem). Triton X-100 was added to a 1% concentration, and the filtrate was cleared by centrifugation at 200 x g for 15 min. The pellet was resuspended in 5 ml of homogenization buffer, mixed with an equal volume of lysine buffer (0.1 M Tris-HCl, pH 7.5, 10 mM EDTA, 2 mM NaCl, 2% SDS), and incubated at 65°C for 20 min. This lysate was then subjected to organic solvent (phenol/chloroform, 1:1) extraction. An equal volume of 95% ethanol containing 2% potassium acetate was added to the aqueous phase, and the DNA was spooled out with a glass rod by gently swirling. Spooled DNA was dissolved in TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM NaEDTA) and repurified by a CsCl density gradient centrifugation. Purified DNA digested with EcoRI restriction endonuclease was fractionated on a 0.8% agarose gel and transferred onto a nylon membrane (Zetabind from CUNO Inc., Meriden, CT) by electrobolting (Bittner et al., 1980). A 580-bp highly conserved PAL gene EcoRI fragment from tomato (Lee et al., 1992) was labeled by nick translation (Rigby et al., 1977) and incubated with membranes in 6X SSPE, 5X Denhardt's, 0.5% SDS, and 100 μg/ml carrier RNA.
Truncated Phenylalanine Ammonia-lyase Expression in Tomato

11825

at 42 °C for 18–24 h. Complementary fragments were identified by autoradiography.

Genomic Library Construction and Gene Isolation—Purified tomato DNA was digested partially or completely with EcoRI restriction endonuclease and used to construct genomic libraries in the Charon 3A vector (Williams and Blattner, 1979). A partial cDNA clone (pPAL5) of a bean PAL gene (Edwards et al., 1985), kindly provided by C. Lamb (Salk Institute, San Diego, CA), was labeled by in vitro nick translation and used to select complementary fragments using plaque hybridization (Maniatis et al., 1982) under the conditions described above.

Sequence Analyses—Fragments complementary to the bean probe were subcloned into a M13mp18 vector for sequencing by the dideoxy chain termination method (Sanger et al., 1977). All primers used for sequencing were synthesized using a Cyclone "Plus" automated oligonucleotide synthesizer (Milligan/Biosearch, Novato, CA). Regions that gave compression in the sequencing gels or contained frameshift deletions and stop codons were sequenced in both directions and analyzed by the chemical degradation methods of Maxam and Gilbert (1980).

S1 Nuclease Protection Analyses—S1 nuclease protection analyses were performed on transcript termini using gene-specific uniformly labeled or 3'-end-labeled probes. Uniformly labeled probes were made by primer extension as described by Sharrocks and Hornby (1991). For the PAL5 gene, a gene-specific primer (20-mer) at the 3'-end (328 bp after stop codon) was used to prime an extension into the coding sequences beyond the stop codon using (α-32P)dATP. After addition of unlabeled nucleotides and further incubation, the DNA was resolved by hybridization with HpaII. The labeled strand (434 bp) was purified by electrophoresis on a 6% gel. For PAL1 and PAL3, specific DNA fragments were made by PCR using two different primers to include the terminal part of and the 3'-end noncoding regions which follow. To label these probes, the DNA fragments were cut with Sau3AI restriction endonuclease, and the protruding termini were filled in using Klenow enzyme and [α-32P]dATP. The probes were subsequently purified on 5% polyacrylamide gels. For S1 nuclease digestion, 20–50 μg of total RNA, purified from wounded tomato leaves as described by Lee et al. (1992), was hybridized to the labeled probes in 20 mM Pipes, pH 6.8, containing 50% formamide, 0.5 mM NaCl and 5 mM EDTA at 42 °C for 12–18 h. Digestions were performed as described by Maniatis et al. (1982), and the products were analyzed on 6% denaturing polyacrylamide gels.

PCR Amplification of cDNA—RNA was reverse-transcribed into cDNA as described by Gerald (1987). In a total reaction volume of 20 μl, 1 μg of total RNA was mixed with 0.1 μg of random hexamucleotides (Pharmacia, Uppsala, Sweden), 0.5 mM of each deoxynucleotide, 50 mM Tris-HCl, pH 8.3, 75 mM KCl, 1–2 pmol of two different gene-specific primers, and 200 units of Moloney murine leukemia virus (M-MLV) reverse transcriptase (Bethesda Research Laboratories). The reaction was carried out at 37 °C for 60 min and 2 μl of the reaction mixture were used to amplify the resulting cDNA by PCR as described by Alwine et al. (1983), and by Wang et al. (1989). A typical 100-μl reaction mixture consisted of 10 μl of 10 X PCR buffer (100 mM Tris-HCl, pH 8.3, 500 mM KCl, 15 mM MgCl2, 0.1% gelatin), 0.5 μl of bovine serum albumin, 0.5 mM each of dATP, dGTP, dTTP, and dCTP, 15–20 pmol of the two gene-specific primers, 1–5 μCi of [α-32P]dATP, 2 μl of cDNA reaction mixture, and 2 units of Taq polymerase (Promega Corp., Madison, WI). The reaction consisted of a 5-min denaturation cycle at 94 °C, followed by 30 repetitive series of reaction cycles consisting of denaturation at 94 °C for 1 min, annealing at 55 °C for 1 min, and synthesis at 72 °C for 3 min. After the PCR, the reaction mix was precipitated with 2.5 volumes of ethanol containing 2% potassium acetate at −20 °C, and aliquots were analyzed by electrophoresis on a 6% polyacrylamide gel. To eliminate genomic DNA contamination of the RNA extract, HindIII restriction endonuclease was used to cut the contaminating DNA after cDNA synthesis.

RESULTS

Genes encoding the tomato PAL enzyme were identified and isolated using a 580-bp tomato PAL gene fragment previously shown to be homologous with bean PAL cDNA (Lee et al., 1992). As shown by hybridization analysis (Fig. 1), EcoRI endonuclease-digested genomic DNA contains five fragments which are complementary to this probe, the darkest band (i.e. shortest fragment) being the probe sequence itself. These fragments were identified as PAL1 (longest) to PAL5 (shortest). The results indicate that, in tomato, there may be as many as five different tomato gene classes which encode the PAL protein.

When genomic libraries of EcoRI fragments in λ Charon 3A or 4A vector were screened, three of the complementary fragments were isolated. As indicated in Figs. 2–4, subcloning into a M13mp18 vector followed by dideoxy sequence analyses determined that all three fragments were highly homologous and potentially could encode PAL protein (PAL1, -3, and -5). As shown in Fig. 3, further isolations of overlapping fragments and subsequent sequence analyses provided the complete nucleotide sequence of PAL5 which encodes a protein sequence that is clearly homologous with that of PAL proteins from other plants. At the DNA level about 67% of the sequence is homologous with that of bean (Cramer et al., 1989), and at the protein level about 81% of the amino acids are identical.

As indicated in Fig. 2, the coding sequence also was determined in the other two isolates (PAL1 and -3); a comparison of these sequences (Fig. 4) again indicates a very high sequence homology with PAL5 and other known PAL gene sequences. However, in contrast to PAL5, four stop codons and three single base deletions or insertions make it unlikely that either the PAL1 or PAL3 sequence encodes a normal PAL polypeptide subunit. To exclude a sequencing error or cloning artifact, sequences in the mutated regions were confirmed by sequencing in both directions using the dideoxy

FIG. 1. Hybridization analysis of EcoRI-digested tomato genomic DNA. Purified genomic DNA (2 μg) was digested with EcoRI, fractionated on a 0.8% gel, electroblotted onto nylon membrane, and hybridized with a labeled 580-bp EcoRI fragment encoding the carboxyl end of a tomato PAL enzyme (PAL5). The size of each fragment (in parentheses) was determined using known size markers which were also applied to the gel.

FIG. 2. Partial restriction maps and sequence analysis of genomic subclones representing three different members of the tomato PAL gene family. Protein-coding regions are indicated by the shaded boxes and arrows indicating the direction and extent of sequence determination using complementary primers or chemical degradation. Fragments are those described in Fig. 1; restriction enzyme cleavage sites are indicated for EcoRI and HindIII by E and H, respectively.
methods of Sanger and co-workers (1977) and resequenced using the chemical degradation methods of Maxam and Gilbert (1980). PAL1 also was resequenced as an independent fragment that was made from genomic DNA by PCR amplification.

The same sequence remained clear in every instance.

While it might be reasonable to assume that the presence of stop codons and frameshift mutations results in inactive gene sequences or pseudogenes, the distribution of mutations...
together with recent reports of subunit polymorphism (Jorrin and Dixon, 1990) raises the possibility that these sequences continue to encode protein. As indicated in Table 1, in each instance, the number of nucleotide changes upstream of the first codon is significantly lower than downstream. Furthermore, differences in the effect of these changes are even more striking when the potential protein sequence is compared. For example, in PAL3, the 4.7% upstream difference in DNA sequence results in only a 0.9% difference in the protein sequence, but the 9.6% downstream difference results in a 10.2% difference in the polypeptide. While this could be coincidental it at least suggests the possibility that the upstream sequence is expressed and remains functionally important.

In view of this possibility the termination regions also were examined by DNA sequencing and terminal mapping using S1 nuclease digestion. As indicated in Fig. 5, substantial sequence homology remains when allowance is made for short insertions or deletions, but this is significantly lower than in the coding sequences (Fig. 4). Nevertheless, consensus polyadenylation sites (Proudfoot and Brownlee, 1976) are clearly evident in each DNA sequence and even may be repeated in at least one of the genes (PAL1). In the PAL3 sequence, however, the consensus sequence (AATTAA) was slightly different than that observed in the other two genes (AA-
but not in sweet potato (Tanaka et al., 1989). Nucleotides are aligned for maximum sequence homology with the termination sites as determined by SI which follow are indicated with corresponding adenylate residue clusters. Consensus polyadenylation signals have been noted in bean (Edwards et al., 1985) and rice (Minami et al., 1989) PAL termination regions but not in sweet potato (Tanaka et al., 1989).

Because the non-coding regions in each gene contained substantial sequence differences the presence of gene-specific mRNA could be probed. This was even true of the coding sequence which was used to demonstrate active gene expression. As shown in Fig. 6, polyadenylation sites could be demonstrated in at least two of the genes. In the normal PAL5 sequence, polyadenylation begins in an AGA sequence upstream of an adenylate cluster (AAAACAA) which is located 28 bp after the AATATA element (Fig. 5). The PAL1 gene is interesting not only because the polypeptide appears to be prematurely terminated, but the gene apparently contains two potential polyadenylation sites, both of which appear to be utilized. As shown in Fig. 6, the SI mapping reveals two mRNA termini both of which end 13–31 bp after an AATATA sequence element in a TCCT sequence which is also upstream of an adenylate cluster (AAAAGAA or AAAAGGA) identified by open boxes in Fig. 5. Since both bands disappear in the presence of heterologous RNA they are clearly the result of mRNA protection and not simple artifacts of S1 nuclelease digestion. All attempts to demonstrate termini with PAL3 were not successful (e.g. Fig. 6).

Because artifactual cleavages in adenylate-rich sequences have been reported with S1 nuclelease, the presence of complementary mRNA for PAL1 was further demonstrated by cDNA amplification using PCR technology. The cDNA was made from extracted tomato RNA using random priming and reverse transcriptase, and any contaminating genomic DNA was only present when reverse transcriptase was used to first make a cDNA copy, and genomic copies were not detected when four different varieties of tomato, S,GCR26, R,GCR218, Bonny Best, and Sweet 100, were examined with the same PCR assay, an equivalent fragment was produced which was confirmed by DNA sequencing (results not shown). The results clearly indicate that, consistent with the S1 mapping, a transcript of the anticipated size (360 bp) for the PAL1 gene was present and widely distributed in tomato. Again, attempts to demonstrate a PAL3 transcript were unsuccessful although the PCR reaction could readily be used to detect the PAL3 gene in genomic DNA (results not shown).
amplification of derived cDNA. These results clearly show that PAL1-truncated mRNA is not only transcribed but is actively utilized by ribosomes for protein synthesis.

**DISCUSSION**

Although a PAL gene family has been previously demonstrated in several other plant species, the present study raises new and intriguing questions about this important defense gene system. Recently reported PAL protein polymorphism (Jorrin and Dixon, 1990), which has been considered to be the product of natural protein degradation, instead appears to be partly or entirely due to truncated gene expression. Four lines of evidence support this possibility: the sizes of the putative polypeptide chains are similar to protein isoforms which were recently reported (Jorrin and Dixon, 1990); the protein sequence is more highly conserved prior to the premature stop codons; S1 nuclease mapping studies indicate the presence of actively transcribed full-length mRNAs; and, lastly, gene-specific amplifications of cDNA transcripts confirm the presence of transcribed RNA and clearly demonstrate that these transcripts are actively utilized in protein synthesis. While this must ultimately be confirmed by protein sequencing the results clearly raise an alternate interpretation of the subunit analyses. Although the PAL3 sequence is putatively less inactivated by stop codons or frameshift mutations than the PAL1 gene (Fig. 4), neither S1 nuclease protection studies nor PCR amplification could be used to demonstrate a full-length transcript. In the absence of such evidence this gene must be considered inactive but, in view of the PAL1 gene, its induction under as yet unidentified conditions should not be eliminated entirely.

The presence of actively transcribed truncated PAL1 sequences also raises questions pertaining to the regulation of this gene family. A variety of studies have shown that the PAL gene family responds differentially to fungal infections, environmental conditions, or developmental cues (Dixon et al., 1983; Hahlbrock and Grisebach, 1979; Jones, 1984; Liang et al., 1989a, 1989b). Enzymatic studies on PAL isoforms have also suggested differences in affinity for the substrate, L-phenylalanine (Jorrin and Dixon, 1990), which could be related to the regulation of phenylpropanoid biosynthesis. Again, the present study offers an alternate interpretation for the differences. Premature truncation in the polypeptide chain could result in enzymes which lack binding sites for substrates or even regulatory sites which are active in feedback inhibition. The truncated polypeptide subunits clearly should not be disregarded as simply products of protein degradation.

**REFERENCES**


Gerad, G. F. (1987) Focus 9, 5-6


Truncated Phenylalanine Ammonia-lyase Expression in Tomato
