The *phtE* Locus in the Phaseolotoxin Gene Cluster Has ORFs with Homologies to Genes Encoding Amino Acid Transferases, the AraC Family of Transcriptional Factors, and Fatty Acid Desaturases

Yuan Xin Zhang and Suresh S. Patil

Department of Plant Pathology College of Tropical Agriculture and Human Resources, and the Biotechnology Program, Pacific Biomedical Research Center, University of Hawaii, 1993 East-West Road, Honolulu 96822, U.S.A.

Accepted 17 August 1997.

A cluster of genes involved in the production of phaseolotoxin, a phytotoxin produced by *Pseudomonas syringae* pv. *phaseolicola*, contains eight (*phtA* through *phtH*) complementation groups (Y. X. Zhang, K. B. Rowley, and S. S. Patil, J. Bacteriol., 175:6451-6458, 1993). In this study, sequencing of the region encompassing the *phtE* locus revealed six putative open reading frames (ORFs), each preceded by a putative ribosomal binding site, and all oriented in the same direction. Reverse transcription–polymerase chain reaction suggested that the *phtE* locus is transcribed as one large (6.4 kb) transcript, indicating that the ORFs constitute an operon. Primer extension analysis showed that the transcript begins at a T, located 31 bp upstream of the ATG codon of ORF1. Comparison of the sequences of the putative ORFs with the sequences of known genes revealed that ORF3, encoding a protein containing 395 amino acids, has 55% similarity to the acetylornithine aminotransferase gene from *Escherichia coli*, and the ornithine aminotransferase genes from other organisms. A lysine residue that is a binding site for pyridoxal phosphate and an arginine residue that is a binding site for the α -carboxylate group of the substrate are conserved in ORF3. These data suggest that ORF3 encodes a protein involved in the biosynthesis of ornithine, a constituent of phaseolotoxin. ORF5, encoding a peptide of 378 amino acid residues, possesses a helix-turn-helix motif at the C-terminal end that is characteristic of the AraC family of transcriptional factors, and there is a possible leucine zipper at the N-terminal end of this peptide. ORF6, encoding a protein of 327 amino acids, has about 40% similarity with the fatty acid desaturase gene, *desA*, of *Synechocystis* Pcc6803 and considerable similarity with fatty acid desaturase genes from other organisms. ORF6 and *desA* show very similar hydropathy profiles and both contain a copper binding signature. Computer searches did not

discover significant homologies in the data base for the other ORFs, but hydropathy analysis showed that all of them contain one to several hydrophobic domains, suggesting that the gene products of these ORFs may be membrane associated.

Phaseolotoxin, a well-studied bacterial phytotoxin, is produced by *Pseudomonas syringae* pv. *phaseolicola*, the causal agent of halo blight of beans (*Phaseolus vulgaris* L.). It contains a tripeptide, consisting of ornithine, alanine, and homo-arginine and an inorganic group (*N*'-sulfodiaminophosphinyl) attached to the *N*⁶ amino group of ornithine (Moore et al. 1984). The toxin inhibits ornithine carbamoyltransferase (OCTase), a key enzyme involved in the biosynthesis of arginine. The inactivation of the OCTase in the toxin-exposed tissues of plants results in arginine deficiency, chlorosis, and growth inhibition (Patil 1974).

Our laboratory has been interested in the elucidation of the molecular genetic mechanisms involved in the production (and regulation) of phaseolotoxin by *P. syringae* pv. *phaseolicola*. Recently, Zhang et al. (1993) reported the genetic organization of a cluster of genes harbored by a previously described (Kamdar et al. 1991) genomic cosmid clone, pHK120, that complements all Tox⁻ mutants induced by EMS, UV, and Tn5 mutagenesis. Tn5 mutagenesis of pHK120, marker exchange of pHK120::Tn5 plasmids in the chromosome of the wild-type strain, and pair complementation analysis revealed that there are a minimum of eight toxin loci (*phtA* through *H*) in pHK120 (Fig. 1). The largest locus among the eight transcriptional units is *phtE*, which spans a 6.4-kb region; the others range from 1 to 2.5 kb in length.

Small peptides such as phaseolotoxin have been reported to be synthesized by specialized multienzyme complexes, and not by the ribosomal mechanism (Marahiel 1992; Martin and Liras 1989). These multienzyme complexes, called peptide synthetases, are very large protein molecules, ranging from 60 to 600 kDa. We reasoned that if pHK120 harbors a peptide synthetase gene involved in the synthesis of the phaseolotoxin tripeptide, it would most likely be present in the *phtE* locus.

Here we report the complete nucleotide sequence of the DNA in the *phtE* locus. It contains six open reading frames

Corresponding author: Suresh S. Patil; Telephone: (808) 956-7280; Fax: (808) 956-4768; E-mail: suresh@pbrc.hawaii.edu

Present address of Yuan Xin Zhang: Lilly Research Labs. Corporate Center, Eli Lilly and Company, Indianapolis, IN 46285, U.S.A.

Nucleotide sequence data are to be found at GenBank as accession number U27310.

(ORFs) arranged in a structure suggestive of an operon. However, none of the ORFs shares similarity with known peptide synthetase genes. The deduced amino acid sequences of ORF3, ORF5, and ORF6 show a high degree of similarity to acetylornithine aminotransferase (and ornithine aminotransferase), the AraC family of transcriptional factors, and fatty acid desaturases, respectively. These data indicate that the product of ORF3 may be involved in the regulated production of ornithine for phaseolotoxin synthesis, the product of ORF5 may be a transcriptional regulator, and the product of ORF6 may be involved in the secretion of phaseolotoxin.

RESULTS

The ORFs predicted from the DNA sequence of the *phtE* locus.

Both strands of the DNA encompassing the *phtE* locus were sequenced and a 7,088-bp sequence was identified (Fig. 2). Six complete and two truncated ORFs, all with the same orientation (Fig. 1), are present in this sequence. Two computer-predicted, prokaryotic, ρ -independent, terminator-like sequences (from nucleotide [nt] 569 to nt 581, and from nt 6305 to nt 6320; Fig. 2) are found before ORF1 and after ORF6, respectively. No terminator-like sequence is located between ORF1 and ORF6. Six complete ORFs, from ORF1 through ORF6, are indicated in Figure 2 with putative start codons and ribosomal binding sites underlined. The arrangement of ORF4, ORF5, and ORF6 is interesting in that each putative ribosomal binding site is located within the preceding ORF, indicating the possibility of translational co-regulation. No promoter-like sequence was found preceding ORF1, but a promoter-like sequence, with TATAAT as the -10 region and TcGCA as the -35 region, was found between nt 6615 and nt 6589. This putative promoter may serve as the RNA poly-

merase binding site for the next gene, starting with the GTG codon at position 6675. The complete sequence of this ORF was not determined because this region was outside the *phtE* locus. Codon preference analysis (data not shown) with the codon usage table of published *Pseudomonas* genes revealed that ORF1 though ORF5 have a pattern of codon usage similar to that of most *Pseudomonas* genes (average codon preference >0.8). However, the pattern of codon usage for ORF6 is not characteristic of most *Pseudomonas* genes (average codon preference = 0.5), indicating that this gene may have been obtained from another species in the process of evolution.

The *phtE* transcript.

As shown in Figure 2, the *phtE* locus spans a 6.4-kb region and harbors six ORFs. To determine whether the six ORFs encode a single large transcript, we first attempted Northern (RNA) blot analysis with total RNA from wild-type cells grown at the permissive temperature (18°C) and the insert in pYZH102 (Fig. 1) as the probe. These experiments did not yield reliable results, possibly because most of the large transcript was degraded during isolation. We then used a reverse transcription-polymerase chain reaction (RT-PCR) kit (Titan One Tube RT-PCR Kit; GIBCO BRL, City, State), which is claimed by the manufacturer to yield RT-PCR products of between 12 and 20 kb from eukaryotic templates, also without success. Therefore, we decided to carry out an RT-PCR strategy involving primers designed to determine whether short RNA templates that continue from one ORF to the next are present in the total RNA isolated from *P. syringae* pv. *phaseolicola*.

Total RNA was isolated from the transconjugant of the wild-type strain 4612-1 containing the plasmid pHK120 on the assumption that multiple copies of pHK120 in the transconjugant would result in an increase in the copy number of the transcript(s) of interest. This RNA was used in RT-PCR

Fig. 1. Transcriptional units within pHK120 that are involved in the production of phaseolotoxin and open reading frames (ORFs) in the *phtE* locus. At the top are the inserts of previously described subclones used in this study, the restriction map of the pHK120 insert, and eight transcriptional loci (rectangular boxes; Zhang et al. 1993). H: *Hind*III, S: *Sma*I, B: *Bam*HI. Below are the ORFs in the *phtE* locus and their transcriptional orientation.

3 AGG AAC TTC GAA TTC GAG GCC GAT AGC CGA
33 TCC AGA CCA AAC CTA TTC AAG CAT AGC CAT
63 GAC CTG TAT GTG AGC GGC TCA AGA GTT GGT
93 GTC GAA TCA GCA ATA TCG GGT GGG ATT TAC
123 GAG TTG GAC TTC TGC TTC TTT CAG TCC GCG
153 GGT GCG CAG CAG AGT AAT AGT CGT GAA GAG
183 TGC ACC GCC AGA AGT GTC GTT ATT CGA TAT
213 GTC GGG GGG CGT AAA CCC TTT GAT ATG GAC
243 TGT GTC CAG TCG GAT ATG AGT GCC CTT TCA
273 AGT TGC CTG GTT AAT ACG CTG GAT CTC GAT
303 TCT TCC ACT CGT GAA CTT TCG GAG CTC CCT
333 AGC TGG ACG TGG AAT CAT TTT GAT ATC GCA
363 TAGagctgtttgattttatttttcagataaattgttgca
402 tgatgtttgttctgctgttaggtttaagtggctccgctc
441 gataagcggcggtgcaatccagatttaattgttattgctg
480 atattccccctcaggtgcgatgcacttatgtggtctgga

519 agtttttcggcgcccggtgtgcccgggaagtaaaatgataag
P-A -----> <---- Termination site
558 gacattgaaatggaattgtcttccgctgtgcaagatttc
597 gttaaaaagaatgcacaaagactttctcggggctcgaaat
636 ataggcttcaacttctctcaatgaagttgcttcccgctcag
675 tgcgatccctggaacacttctactcaccagttccggcggtc
↓ Transcriptional start
714 caTggtatttgccaccatggccagcagatattgat
747 **ATG** TTG TGC GTG GTT **GAC GAG CCT GTC AGC**
ORF1 ⇒ P-A'
777 ACC GAG CAA ATG GCA ACC CAG ATT CAT CAG
807 GAA GGT CTG CAT CTC GAA CTG TTG CTG TTC
837 AAT GCT GCC GAT GTT CAC TTG GCC CTT GGC
867 GAA CTC ACC AGC CTG GCG CAA AGC CCA GCG
897 CTG ATC TGC CTG CAA AAT TAT TCA AGA TGG
927 GAT AAG CGC AAC AAG GTT CCG CGC AAA TAT
957 GTC GAG CGC TTG GTC TAT GGC GTG ACC AAT
987 GAT CTC ACC ACT CCT TAT CTC TCG GCA CTT
1017 GGC GAC CTG TGT GTC GTG TCT CGC AAA
1047 GAC CTG GAT GTT TTC AGA GTG AAT GTC GCC
1077 TGC TCA CGG CTG GCG CAA CGG TCT GAC GAG
1107 CCT TTT GCG GCG ATC GGC TAT TTG ATG AAT
1137 GCA CTA TGC GCA GCC ATG AAC TCG CTG CTT
1167 TCG ATG TCG GGT TGG GTG TTG TCC AAC AAG
1197 AAG TGG ACG CTG CGT CGC TGG AAT AAA AGC
1227 CGC GAC CTG GTT TTC TTG GAT CTT TGC CCT
1257 ACC ACC TGG ACG GAG GTT TCA GCG CTC TGG
1287 GCA GAC CTT GCG AAA GGT CCA GCG GTG TTG
P-1
1317 GAC GAA GCC CTG TCA GCG ATC GAG ACG TTA
1347 TTG CAG CAA ATG ACA GGC AGG CTG GCT CCG
1377 GAA CTC CTT CGT GAT GAA CCA TTT ACC GGG
1407 TTG CAG GGA AAT GTT GCG CGG TAT CCG TTG
1437 CAC GCC AAG CCG CTC TGG TCT GCG CTG GCA
1467 ACG GTT ATG TAC GAT TGG CCG AGA ACG TGC
1497 CGG CCG GCG TAACgctggcgagctgtttcgccgac
1533 tgtgccagttgcccgcgctctctgcccctegaacacctgc
1572 aagcgatccgcgcagggacctggagtcacatcattgcaac
1611 gccagtcgggggcccagcgccagctaaggagcagc**ATG**
P-2 ORF2 ⇒
1650 TTT GAA GCA CGA CAA GAC TCA ACC TTG CGT
1680 TGG TTT CCT CGG CTG ACC GGA GGC GTG GGG

1710 GTG GAG GGC AAT AGC ATG GCC CGC GCC ATC
P-1'
1740 GTC AGC GCG GCA TGG CTG GTG ATG TCC GAA
1770 CTG TAT GCC TAC CTC GAA GAT CTG GAA GGC
1800 GCC ATG GAT GCT CCG GAT GCA TCG GTG TTG
1830 ATT AAA GTG AAG ATT GCT GAA CTG CTG GTG
1860 CAA ATC GAC TGC ACG TTG GGG CGC ACG GCG
1890 GTG CTC GAC GAA GAG CAC AGA TTG CCG TGG
1920 TTA CTT GAG TAC GGC TTG TGC GAG GTG ATC
1950 AAT TTA CCC GGG GCT GAC ATG GCG CGC TTG
1980 CTG GGT CTG TTC GCA GCG AAT GAC GCA ACT
2010 GAA ATC CGT CGA GTA AGC CAG CTG ATA CGT
2040 GAT TTG ATT GCG GCA TTC CCG GGC GAG TTG
2070 GTC GAT TCG CTG CAA GCG CAC AAC CAG GGA
2100 GGC GTT CTG CGT TTT CTT CGG TCT TCC GAC
2130 AAG GCG TGT ACC GCG TTG GGT TGC GAT GCC
2160 AGT TTT CTG GTG CCA CTG ATG AAA TCC CTA
P-3
2190 TGAaactgaactgactggacggtctgtcccggtgcgga
P-2'
2229 agcaatccgagcgtaaaagcgcttcggtgatcgctatg
2268 gtgctgtcttcggtcgctcaacagggcctg **ATG** TTG
ORF3 ⇒
2303 AAG ATG TCG GGG TTG TCC GCA TCC GAA GCG
2333 CGT GGC GAC GGT GCC TGG GTC GAA GAT ACT
2363 CAG GGT GGA CGC TGG CTG GAC TTC GGT TCG
2393 TTC GGT GTG CAC CTA CTT GGC CAT AGT CAT
2423 TCA GGC GTG GTA TCT GCC CTG GTG GAA CAG
P-4
2453 ATC CAG CGT TTT GGG CTT TCC GAC GAA AAT
2483 CTG TCG AAC **GAA** CCC ATC GTA CTG GCG GCC
P-3'
2513 GAG CGG CTA TTG GTC ATG GCG GGA CCT GAG
2543 AAG GAT AAG GTG ATT TTC GGC AAT ACC GGT
2573 TCC GAA GTC GTC GAA GCG GCA TTG AAA CTT
2603 GCC AGG ATC GTC ACT GGA CCG CGG AGG GTC
2633 ATT GCC TTC GAG CAG GCT TAC CAC GGG CGC
2663 ACC GCG GCT GCC TTG AGT GTC AGT CAT GGC
2693 TAC ATG CGG CAT GCC GGT CTT CTG ACT GAA
2723 GGT GAC GTG GTT TTT TGT CCA ATC GAC GAT
2753 CTC GAT GCT GTC GCC AAC GCG CTG GAG GCA
2783 GGC GAC ATC **GCG GCG ATC ATT ATT** GAG CCC
P-5
2813 **ATA** CAG GGC GAG GGC GGG ATA CGC CCT GTG
P-4'
2843 GAG AAT GAA TTT CTT CGA GGT CTT GGC GAC
2873 CTG GCC CAG GCT TTT GAC TGC CGC TTG ATA
2903 TTT GAC GAG CTC CAA ACG GGC CTG GGG CGT
2933 ATC GGT AGC CTG CGT GGA CCG GGT CCT TGC
2963 GAC ATC CTG CTG TTG GGC AAG GTT CTG GGC
2993 GGT GGT GTT TTC CCG ATC TCG GCA GCC TTG
3023 TTT ACG TCG AAG CGT TTT GAT GCG GCT GCC
BamHI
3053 **CGG GAT** CCG GTA GTC CAC GCT TCG TCT TTT
3083 GCC GGT AGT GCC CTG GCC GGT GCC GTC GTC
3113 AAC GCG GTG CTG GAT GTT GTG TCC GCT CCC
3143 GAT TTC ACC CTC AGG GTG CAG CGC TTA GGC
3173 GAT CGG GTC TTG AGC TAT CTC GGC CAG CGC
3203 CTT GCA GAT AAT CCC TTG ATT ACC GAC ATC

Fig. 2. Nucleotide sequence of the DNA encompassing the *phtE* locus. Nucleotide sequences encoding putative open reading frames (ORFs) are shown in uppercase. Start codons are underlined and shown in bold; putative ribosomal binding sites are underlined. Promoter-like sequences are shown in bold and marked by the symbol ^ underneath; terminator-like sequences are overlined with a dotted line. Transcriptional start site is marked by a vertical arrow above the sequence. Sequences of synthetic primers (P) used for reverse transcription-polymerase chain reaction are underlined; arrows indicate the orientation of primers. Restriction sites for *Bam*HI and *Hind*III are shown. Putative helix-turn-helix domain characteristic of the AraC family signature in ORF5 is italicized and underlined. Leucine zipper-like sequence occurs between nucleotide 3967 and nucleotide 4053 in ORF5 (see Results).

Figure 2 continued on next page

experiments as a template to test whether a given pair of primers leads to the synthesis of the expected fragment. Use of the primer pair A and A' (Fig. 2) did not lead to the synthesis of a DNA fragment (data not shown), indicating that the 5' end of the transcript of *phlE* does not reach the binding site for

Figure 2 continued from previous page

[illegible]

Figure 2 continued on next page

3A,B lane 12), indicating that it was an artifact. These results indicate that the 3' end of the transcript lies downstream of the binding site for primer 12, which is located within the untranslated region (Fig. 2). The rest of the primer pairs used all led to production of fragments of expected sizes, and all of them hybridized to specific probes (Fig. 3A and B). Because the RT-PCR-generated fragments either overlap with another fragment or reside within a given ORF (Fig. 3C), the results presented above collectively suggest that the *phtE* locus encodes a large transcript that covers all six ORFs.

Primer extension.

To determine the transcriptional start site, total RNA from the wild-type strain G50-1 grown at 18 or 28°C and from transconjugants of 4612-1 containing pHK120 (18°C) was used as the template for primer extension. The synthetic oligo, primer A' (5'-CGGTGCTGACAGGCT-3'), was used as the primer. A major band was detected in all three lanes (Fig. 4, lanes 1, 2, and 3) that represent different RNA sources: from the wild-type strain grown at 18°C (Fig. 4, lane 1), wild-type strain grown at 28°C (Fig. 4, lane 2) and 4612-1::pHK120 grown at 18°C (Fig. 4, lane 3), indicating that the expression of the *phtE* locus is not temperature regulated. This band corresponds to nucleotide A in the sequencing ladder, indicating that the nucleotide T in the complementary strand, located 31 bp upstream of the ATG codon of ORF1 (Fig. 2), is the transcriptional start site. Examination of the upstream region of this start site did not reveal a sequence identical to the canonical sequences (TTGACA-17bp-TATAAT and GG-10bp-GC) recognized by σ^{70} and σ^{54} RNA polymerase holoenzyme (Deretic et al. 1989), respectively. However, the -10 and -35 regions of transcriptional start, CAGTT and TGGAAC, have some similarity with the *Escherichia coli* canonical promoter

sequence, raising the possibility that it may serve as the promoter for the *phtE* locus.

The deduced amino acid sequence of ORF3.

ORF3 encodes a putative protein containing 394 amino acids (aa), the largest gene product in the *phtE* locus. We named the putative ORF3 product Argd_Pphas. A BLAST search (Altschul et al. 1990) of the Swissprot data base revealed that the ORF3 amino acid sequence has significant similarity to acetylornithine aminotransferase (ACOAT) and ornithine aminotransferase (OAT) from *E. coli* (Heimberg et al. 1990), *Saccharomyces cerevisiae* (Degols 1987; Degols et al. 1987; Heimberg et al. 1990), and *Bacillus subtilis* (Calogero et al. 1994; Gardan et al. 1995; O'Reilly and Devine 1994); to OAT of *Vigna aconitifolia* (Delauney et al. 1993); and to OAT of mouse (G. Manenti, M. Garibaldi, M. A. Pierotti, G. Della Porta, and T. A. Dragani, unpublished data; EMBL/GenBank/DBJ data banks accession number: P29758), and rat (Oyama et al. 1990). When the amino acid sequences of the genes encoding these enzymes were aligned by the Pileup program of GCG, five highly conserved domains were detected (Fig. 5). A lysine residue (K) found in aspartate aminotransferase (AAT; Kirsch et al. 1984) and in histidinol phosphate aminotransferase (HPAT; Hsu et al. 1989) to which the cofactor pyridoxal phosphate (PLP) is believed to bind is conserved in all of these gene products (Fig. 5, marked by #). An arginine residue (R) believed to be involved in the binding of the α -carboxylate group of the substrate is also conserved in all these gene products (Fig. 5, marked by *). The N-terminal sequence of ORF3 is significantly different from that of all the mammalian OAT genes (data not shown), which is not surprising since the mammalian OATs are mitochondrial enzymes and the genes encoding them have leader sequences that are responsible for mitochondria targeting (Inana et al. 1986).

Figure 2 continued from previous page

```

        -----> <-----Term.site
6286  acactatccccggaagaaggcgtggtccagcctcccgaa
6325  ggcggtttttacctgggctctaccgtcgaatgtatcgccg
6364  ggctggatttttgcgagtgaaactctacccaatctatcga
6403  cagcgctatgtggcatgttctattcaaacacgcgcgccac
6442  tggggcgcacaggggcaacgatctgctggacgctggaaa
6481  ttggtggtcactcaaccgctgctgtatctatgcagggacga
        P-12'
6520  agattgccaaagtgtgtttctggaaaaattttggtgccg
6559  ttcgcgctatgaagggcggttatgcaggctcgcaaaagt
        P-14
6598  tcgttccttcaagaattataatggaccaaccatgatatt
        P-14'
6637  gacgggtaatagaataaccaatcgcttattgcaggggat
6676  atc GTG ATC TCT CCT TTC TCA AAG AAT AGC
6706  ATC AAT CCA AAT AGC TAT AAT TTT CGC CTG
6736  CAT AAT CAT ATG AAA GTC TAT GAG GAC GGC
6766  ATT ATC GAC GTG AAG TCG GAG CAG CCA ACC
6796  CGT GAT ATT TAT ATC GGC TCT GAA GGG TTT
6826  GTT CTC GAA CCG ATG CGG CTG TAC CTT GCC
6856  AGT TCA ATC GAG ACC ATG GGC AGT ACT CGT
6886  TTT GTA CCG ACT TAT GCG GCG CGT TCT TCC
6916  ATT GCT CGA CTC GGC ATG TTT ATC AAC CTG
        P-14'
6946  TCG GCA CCG CTT GGT GAT ATA GGA TTT ATA
6976  GGC CGT TGG ACC ATT CAA CTT TTG GCC GTC
7006  AAT AGA ATT CGG GTA TAC CTG GCA TGG ATA
7036  TTG GCA ATG ATG TTC TGG CAT GTT CAA GGC
7066  GAG ATC GAG CTG TAT ACG CAA GT 7088

```

The deduced amino acid sequence of ORF5.

When the amino acid sequence of ORF5 was used for a BLAST search of the Swissprot and PIR data bases, no gene product in the data bases showed significant similarity to ORF5 over long stretches. Several gene products, such as MS2, a male sterility protein from *Arabidopsis thaliana* (Aarts et al. 1993), and ANFa, a nitrogen fixation protein from *Azotobacter vinelandii* (Joerger et al. 1989), show similarity to ORF5 over regions no longer than 60 aa and the similarity is around 40%. However, when a motif search of ORF5 was done, a helix-turn-helix domain (Fig. 2, italicized and underlined) that is characteristic of the AraC family of transcriptional factors in bacteria was found. As in the proteins from most of this family, the helix-turn-helix motif is located in the C-terminal two-thirds of the ORF5 sequence. At the N-terminal end of this peptide, there is a stretch (nt 3967 and nt 4053; Fig. 2) of four leucine residues, each separated from the next by six amino acids, except for the second and the third leucine residues, which are separated by seven amino acids. This structure resembles the leucine zipper found in many eukaryotic (and a few prokaryotic) proteins (Sasse-Dwight and Gralla 1990; Weissbach and Brot 1991) that are either transcriptional factors or are involved in protein-protein interactions.

The deduced amino acid sequence of ORF6.

A computer search in the GeneEMBL data base revealed that the amino acid sequence of ORF6 is similar to that of the

Fig. 3. Results of the reverse transcription-polymerase chain reaction (RT-PCR) experiment. RNA was isolated from transconjugant 4612-1(pHK120). Two micrograms of total RNA was treated with 40 units of RNase-free DNase I and used as the template for RT-PCR amplification in a final volume of 100 μ l. **A**, Twenty micrograms of each RT-PCR product was separated on a 2% agarose gel. **A** and **B**, Primers used (Fig. 2) were 1 and 1' (lane 1), 2 and 2' (lane 2), 3 and 3' (lane 3), 4 and 4' (lane 4), 5 and 5' (lane 5), 6 and 6' (lane 6), 7 and 7' (lane 7), 8 and 8' (lane 8), 9 and 9' (lane 9), 10 and 10' (lane 10), 11 and 11' (lane 11), 12 and 12' (lane 12), 12 and 14' (lane 13), 14 and 14' (lane 14), primers provided with the RT-PCR kit (lane 16, positive control), and 3 and 3' (lane 16, no RT added, negative control). Primer binding sites are shown in Figure 2. Shown at the left (**A**) is the DNA marker ladder (with molecular weights in bp). **B**, Southern blot analysis of the fragments as shown in **A**. Lane 1 to 11 were hybridized with ^{32} P-labeled insert of pYZH102 (Fig. 1). Lanes 12 to 16 were hybridized with ^{32} P-labeled insert of pYZH202 (Fig. 1). RT-PCR fragment generated by primer pair 10 and 10' is not shown. However, in a separate experiment the expected fragment was observed. A fragment was present in lane 12 in the agarose gel but not in the blot, indicating that this fragment is an artifact. **C**, Regions in the *phlE* locus covered by expected fragments generated in the RT-PCR experiment. Numbers at top of each line correspond to the lanes in **A** and **B**.

fatty acid desaturase of *Synechocystis* Pcc6803 (Wada et al. 1990), Δ -12 desaturase of *Anabaena variabilis* (Sakamoto et al. 1994), β -carotene ketolase of an *Alcaligenes* sp. (Misawa et al. 1995), Δ -15 lineoyl desaturase of *Limnanthes douglasii* (Bhella and MacKenzie 1995), fatty acid desaturase of *Arabidopsis thaliana* (Okuley et al. 1994), oleate desaturase of *Brassica juncea* (Singh et al. 1995), and fatty acid desaturase of *Nicotiana tabacum* (Hamada et al. 1994). Alignment of these sequences (Fig. 6) performed with the Pileup program of GCG revealed two conserved regions, the first located at aa 58 to 112 of ORF6, and the second at aa 174 to 247 of ORF6, which contains a conserved copper-binding domain, CH-IN-H-HHL (Chepuri et al. 1990).

Independently, Hatziloukas et al. (1995) sequenced an *EcoRI* fragment from the *tox-argK* gene cluster from another strain of *P. syringae* pv. *phaseolicola* that encompasses the region between nt 3556 and nt 6308 (Fig. 2) and identified an ORF (DES-1755) of 584 aa residues. Bestfit analysis between our sequence and that of Hatziloukas et al. (1995) revealed five insertions in their sequence that were not present in our sequence. The following mismatches were found between the sequence of Hatziloukas et al. (1995), referred to here as the Hatz sequence, and our original sequence: an extra G at nt 804 in the Hatz sequence (nt 4357 in our sequence); an extra A at nt 1159 in the Hatz sequence (nt 4712 in our sequence); an extra T at nt 1184 in the Hatz sequence (at nt 4735 in our sequence); an extra A at nt 1190 in the Hatz sequence (at nt 4741 in our sequence); an extra A at nt 1385 in the Hatz sequence (at nt 4934 in our sequence). To resolve these differences we designed three different primer pairs covering the relevant regions of the sequence and re-sequenced these regions by cycle sequencing as described in Materials and Methods. The first primer pair used was 5'-CCGCACTGATCTATGTTTC-3' and 5'-TCTTGCTGTCACCCACTAC-3', which covered nt 4344 to 4462 of our sequence. The second primer pair was 5'-GACAAGCGTATCGGCTAC-3' and 5'-CTTGATGACCTGTGGCAG-3', which covered nt 4679 to 4981 of our sequence. The third primer pair was 5'-ATGCAGTAATGCCCGAAG-3' and 5'-GAACATCAATCTGCCAGCC-3', which covered nt 5807 to 6057 of our sequence. The new sequence data confirmed our previous sequence and therefore our designation of ORF6 (Fig. 2). The predicted size of the ORF6 product is 35 kDa, close to that of *desA* (Wada and Murata 1990). It is also close to the DES-1755 product (34.2 kDa) expressed in *E. coli* under T7 promoter control (Hatziloukas et al. 1995), but is much smaller than the predicted size of the DES-1755 product of about 61 kDa based on its size (584 aa). This difference between the predicted versus the expressed size of the DES-1755 product must be due to the five extra bases found in the DES-1755 sequence, which resulted in a frame shift leading to an increase in its size. Two membrane-spanning regions are perfectly aligned between the *desA* product and the ORF6 product, indicating that these regions are crucial for membrane association of the ORF6 product (data not shown). A similar observation has been made by Hatziloukas et al. (1995).

The deduced amino acid sequence of the other ORFs.

No clues as to the possible functions of the ORFs 1, 2, and 4 in the *phtE* locus are available because the computer searches did not discover any gene sequences homologous to

these ORFs. However, the hydropathy analysis showed that all of these ORFs contain one to several hydrophobic regions (data not shown), indicating that their products may be membrane associated.

DISCUSSION

The nucleotide sequence data of the *phtE* locus reported here shows that there are six ORFs in this locus and that they are arranged in a tight cluster. The organization of these ORFs is suggestive of an operonic structure. This is based on the following points: (i) a terminator-like sequence is located upstream of ORF1, which we assume acts as a terminator for the ORF upstream of ORF1; (ii) a terminator-like sequence is present at the C-terminal end of ORF6 but not following ORFs 1 through 5; (iii) primer extension shows that the transcriptional start is 31 bp upstream of the ATG codon of ORF1, and the -10 and -35 regions of this start have some similarity to the *E. coli* promoter sequence, suggesting that they may be part of the promoter controlling all six ORFs; (iv) a promoter-like sequence is located downstream of ORF6, which presumably

Fig. 4. Transcriptional start site of the *phtE* locus. Total RNA isolated from G50-1 grown at 18°C (lane 1), G50-1 grown at 28°C (lane 2), and 4612-1(pHK120) grown at 18°C (lane 3) was used as a template in the reverse transcription reaction. The primer used was A' (5'-CGGTGCTGACAGGCTC-3'; Fig. 2). Arrowhead indicates possible transcriptional start site of message. At left, DNA sequence ladder (complementary to the mRNA sequence) generated from pHK120 DNA with the same primer as that in the primer extension experiment. Nucleotide A (transcriptional start site) is marked with an *.

Fig. 5. Alignment of ArgD_Pphas amino acid sequence with that of acetylornithine aminotransferases (ACOATs) and ornithine aminotransferases (OATs) from various organisms. Sequences are aligned with the Pileup program of GCG and shaded with the Boxshad program available on the Internet. ARGD_ECOLI: ACOAT from *Escherichia coli*, ARGD_BACSU: ACOAT from *Bacillus subtilis*, ARGD_YEAST: ACOAT from *Saccharomyces cerevisiae*, OAT_ECOLI: OAT from *E. coli*, OAT_BACSU: OAT from *Bacillus subtilis*, OAT_YEAST: OAT from *Saccharomyces cerevisiae*, OAT_VIGAC: OAT from *Vigna aconitifolia*, OAT_MOUSE: OAT from *Mus musculus*, OAT_RAT: OAT from *Rattus norvegicus*. Five regions with high similarity among all the sequences are marked by a line on top. The lysine (K) residue believed to be involved in pyridoxyl phosphate binding is marked with an #. Also marked (*) is an arginine (R) residue that is probably the binding site for the α -carboxylate group of the substrate.

Figure 5 continued on next page

Figure 5 continued from previous page

		IV	
		#	
ARGD_ECOLI	191	LLVFDEYOCGAYMHYGVTPDILTSAKALGGG.FPISAMLTAEIASAFHPGS..HGSTYG	
ARGD_BACSU	177	LLIIDEIQTGAYEHFGLSPDIITVAKGLGNG.FPVGAVIGKKQGEAFTPGS..HGTTFG	
ARGD_YEAST	205	IVIHDEIQTCCAYLPSEAHPDIFTSKALGNG.FPIAATLVNEKYNNALRVGD..HGTTYG	
OAT_ECOLI	204	LMILDEYQTGACEHENVOPIILCLAKALGGGVMPIGATIAATEEVFSVLFDPNPFHTITTFG	
OAT_BACSU	187	LFIADIEIQTGD..WDGIVPDMYILGKALGGGVFPIISCAADREILGVFNPGS..HGSTFG	
OAT_YEAST	196	LLIVDEIQTGDHYKAEAKPDIVLLGKALSGGLVPVSCVLSHDIMSCFTPGS..HGSTFG	
OAT_VIGAC	221	LFIADIEIQTGDH..EDVKPDLILGKALSGGLYPVSAVLRDDHIMDCIQPGL..HTAMDV	
OAT_MOUSE	220	LFIADIEIQTGDH..ENVRPDMVLLGKALSGGLYPVSAVLCDDIIMLTIKPGE..HGSTYG	
OAT_RAT	220	LFIADIEIQTGDH..ENVRPDMVLLGKALSGGLYPVSAVLCDDIIMLTIKPGE..HGSTYG	
ARGD_PPHAS	167	RLIFDEIQTG...RVDVPCDILLGKVLGGGVFPIISAAFTSKRFDAAARDPVVHASSFA	
CONSENSUS	241	l i DEIQtG v pDil laKaL gGvfPisaml i pg Hgstyg	
		V	
		*	
ARGD_ECOLI	248	AAFDIINTPEVLEGIQAKRQRFVDHIQKIDQQYDV.FSDIRGMGLLITCAEDFLYAGAEAG	
ARGD_BACSU	234	ATLQIVFQPDFLQEAADKCAFLKEOLE.AELKSPF.VKQIRGKGLMLCTHEDIIAELQTLG	
ARGD_YEAST	262	YVLDTIADEAFLKQVSKSDIIOKRIREIQAKYPNQIKTIRGKGLMLCAEEVIKKARELG	
OAT_ECOLI	264	ATINVLEQNLPAQAEQKQDMILDGFRQLAREYDPLVQEARCKGMLMATEGYNFASEMFR	
OAT_BACSU	243	ASLEVEDEKLADRSLLEGEYFKSEESIDS...PVIKVIRGRGLFIGVEARPYCERLKE	
OAT_YEAST	254	AALVIRDEKLCQRAAQKSSFIAQLKALQAKSNGLIISVIRGMGLLTAIVAWDLCLLMKD	
OAT_VIGAC	277	SRYVVRVARERCENAQIOCTYLRKEINTLPK...DVVPVIRGKGLNATVAVWDVCLN...	
OAT_MOUSE	276	AALVIEEENLAENADKMCAILRKEIMKLP...DVVTSVIRGKGLNATVAVWKVCLRLRD	
OAT_RAT	276	AALVIEEENLAENADKMCAILRKEIMKLP...DVVTAVIRGKGLNATVAVWKVCLRLRD	
ARGD_PPHAS	224	AVLDVVSAPDFTLRVQRLCDRVLSYLGQRLADNP.LITDIRGOGLMICVEIIEAAKR...	
CONSENSUS	301	a ldvi g l l i lv iRGkGll gi	
ARGD_ECOLI	307	VMVLNAG...PDVMRFAPSLVVEDADIDEGMQRFAHAA.....	
ARGD_BACSU	292	LLVLPAG...PERDSAAAPLTVTKDEIAEAASKLKQAA.....	
ARGD_YEAST	322	LLITTAG...KSTVRFVPEALTIEDLHEEGMDAFKAAE.....	
OAT_ECOLI	324	QRVLVAGTLNNAKTIRIEPPLTLITIQCELVKAARKAA.....	
OAT_BACSU	300	EGLLCKET..HDTVIRFAPPLIISKEDLDWALEKIKHVER.....	
OAT_YEAST	314	QGLLAKPT..HGDHIRLAPPLVISEEDLOTGVEIACID.....	
OAT_VIGAC	331	...LCKPT..HGDHIRFATGHHRGTDPRMC.QYYQKYH.....	
OAT_MOUSE	333	NGLLAKPT..HGDHIRLAPPLVIKEDEIRESVEIINKTIL.....	
OAT_RAT	333	NGLLAKPT..HGDHIRLAPPLVIKEDEIRESVEIINKTIL.....	
ARGD_PPHAS	280	.DLLAFCLTDPKVLRLYPFAVISDADLETG.EHFCSWEIPLCQQ*	
CONSENSUS	361	vl vmr appl v ddi m h v	

acts as a promoter for the ORF downstream of ORF6; and (v) ORF3 through 6 overlap such that the ribosomal binding site for one ORF is within the C-terminal end of the preceding ORF, indicating translational coupling. Similar translationally linked ORFs have been reported for the operon in the *bap* gene cluster involved in the biosynthesis of bialaphos (Raibaud et al. 1991). Even though we failed to detect a message of the expected size in Northern blots, probably because large prokaryotic transcripts are easily degraded, the results of the RT-PCR experiments shown in Figure 3 suggest (but do not prove) that the message encoded by the *phtE* locus covers all six ORFs and is continuous. A similar method was used by To et al. (1994) in a study about the same time this study was being carried out, in which they reported an operon organization of genes involved in the synthesis of carotenoids in *Erwinia herbicola*. These two independent studies showed that the RT-PCR-based method can be used to test whether several genes are transcribed within a single transcript, if the primers are carefully designed and appropriate controls are carried out.

Attempts to obtain a full length RT-PCR product covering all six ORFs using primers based on the sequences of ORF1 and ORF6, total RNA from the wild-type strain and a recently introduced RT-PCR kit (Titan One Tube RT-PCR system, Boehringer Mannheim, Indianapolis, IN) were unsuccessful, further indicating the instability of the RNA template. To conclude definitively that the six ORFs in the *phtE* locus are arranged in an operon, further studies involving polar and promoter mutations will have to be conducted.

Proteins encoded by genes in an operon are often functionally related, viz., the products of the genes in the *virB* operon of *Agrobacterium tumefaciens* (Ream 1989) or the products encoded by genes in the *lacZ* operon of *E. coli* (Beckwith 1987a). An advantage of the operon organization is that the production of the transcript that is expressed as a polycistronic message can be coordinately regulated (Beckwith 1987b). Although the functions of the products of the *phtE* locus are not understood, the fact that all of them except the product of ORF3 and ORF5 contain one to several hydrophobic regions

Fig. 6. Alignment of the amino acid sequence of ORF6 with various desaturase genes. Sequences are aligned with the Pileup program of GCG and shaded with the Boxshad program available on the Internet. FAD2_ANVA: Δ -12 desaturase from *Anabaena variabilis*, CRTW: β -carotene ketolase from *Alcaligenes* sp., DESA_SYNY3: fatty acid desaturase from *Synechocystis* sp., LDLDES: Δ -15 lineoyl desaturase from *Limnanthes douglasii*, FAD_ARATH: Ω -6 fatty acid desaturase from *Arabidopsis thaliana*, BJODES: oleate desaturase from *Brassica juncea*, NTFAD3: Ω -3 fatty acid desaturase from *Nicotiana tabacum*. Two regions with high similarity among all the sequences are marked with lines above the sequence; the copper binding signature in domain II is marked with *.

Figure 6 continued on next page

Figure 6 continued from previous page

		II			
				*	***
FAD2_ANVA	213	LFGTTHCDYPKWVEVLCHDINVHVP	PHHLSVAIPSYNLRLAYSSIQENW	CD	...FWS
CRTW	164	PLSLITC	...FHFGGYHHEHHLHPHV	PWN	...
DESA_SYNY3	216	LNGTVHCDYPRWVEVLCHDINVHVP	PHHLSVAIPSYNLRLAHGS	LKENW	GP...FNW
LDLDES	285	L.TTLDRDYGLINNTHHDI	.GTHVIHHLFPQIPHYHLVEATQA	KPIFKLPFHLIDV	LL
FAD2_ARATH	243	L.ATVDRDYGILNKVFHNITDTHVA	HHLFSTMPHYNAMEATKA	KPILED	...WYVAMY
BJODES	244	L.ATVDRDYGILNKVFHNITDTHVA	HHLFSTMPHYHAMEVTKA	KPILED	...WVKAMW
NTFAD3	231	L.TTVDRDYGLFNNTHHDI	.GTHVIHHLFPQIPHYHLREATKA	AKPVLC	KIPFHLVKDLT
ORF6	211	TERFLWA	...PCHINRHW	EHLCPCQCP	RTYR...CLLHDW
CONSENSUS	301	l	ti	dy	v Hv HHL iP y k g l
FAD2_ANVA	269	KLITDECO	YQTDVGYQPFKDY	YAGR	...
CRTW	190	...R	PRTRKTGGRA	*	...
DESA_SYNY3	272	QQISGQCH	YDPEHGYRTFGSL	KKV	...
LDLDES	343	KSLKRDHFV	...PDTGDIVYYQSDP	QISGSLKPE	*
FAD2_ARATH	298	REAKECIYVEPDREGDK	KGWYNNKL	...	
BJODES	299	REAKECIYVEPDRQGEK	KGWYNNKL	*	...
NTFAD3	289	RSMKQDHYV	...SDSGEIVFYQTD	PHIFRSAPKDE	
ORF6	259	CRSCTPTPLPHGWQ	IDVRVSQRMAGVLR	RSFHDADG	*
CONSENSUS	361	l			

suggests that they are membrane-associated proteins, and may be functionally related. Phaseolotoxin production in *P. syringae* pv. *phaseolicola* is temperature regulated and some of the toxin genes may not be expressed at 28°C due to the presence of a repressor that negatively controls these genes at this temperature (Rowley et al. 1993). However, the genes in the *phtE* locus are not among those regulated by temperature, because the same primer extension product was obtained with RNA templates prepared from cells grown at either 18 or 28°C (Fig. 4, lanes 1, 2, and 3). Nevertheless, it is possible that high temperature affects the expression of the operon in the *phtE* locus quantitatively, which cannot be determined from the results of a primer extension experiment. In this study we have not attempted to determine whether any of the proteins encoded by the six ORFs in the *phtE* locus are expressed in *P. syringae* pv. *phaseolicola*. However, in a previous study (Zhang et al. 1993) we showed that 11 separate Tn5 marker exchange mutants distributed throughout *phtE* were all Tox⁻, indicating the products of the ORFs in this locus are expressed in vivo.

The deduced amino acid sequence of ORF3 (Argd_Pphas) shows significant similarity to acetylornithine aminotransferase (ACOAT), an enzyme that catalyzes the formation of *N*-acetylornithine and α -ketoglutarate from *N*-acetylglutamate semialdehyde and glutamate (Alonso and Rubio 1989). The pyridoxal phosphate and α -carboxylate substrate binding sites are conserved in Argd_Pphas. These results strongly suggest that Argd_Pphas is a homologue of ACOAT, and that it may be involved in the production of ornithine, a constituent of phaseolotoxin. With homology not only to ACOAT but also to ornithine aminotransferase, Argd_Pphas may belong to the large aminotransferase family. Even though Argd_Pphas may be involved in the synthesis of ornithine, Zhang (1994) has previously reported that the marker exchange mutants containing Tn5 insertions upstream of Argd_Pphas are not auxotrophic for arginine. This suggests that Argd_Pphas catalyzes the formation of ornithine needed for phaseolotoxin production, but is not used to produce the ornithine pool used in ar-

ginine biosynthesis, and that there is another, housekeeping ACOAT gene present elsewhere on the chromosome of *P. syringae* pv. *phaseolicola*.

ACOAT has been reported to display substrate ambiguity in both *E. coli* (Billheimer et al. 1976) and *Pseudomonas aeruginosa* (van der Wauven and Stalon 1985). This enzyme utilizes acetylornithine as well as ornithine as substrates, thus playing a dual role in arginine biosynthesis and arginine catabolism (Voellmy and Leisinger 1975), respectively. Its real function in arginine catabolism is to transaminate succinylornithine (van der Wauven et al. 1988), an intermediate in a newly discovered catabolic pathway responsible for the breakdown of arginine in a number of bacteria (van der Wauven and Stalon 1985). In light of this, we propose that Argd_Pphas catalyzes the formation of ornithine needed for phaseolotoxin production with the arginine pool. The evidence that supports this hypothesis is that addition of arginine (50 mg/liter) to the culture medium increases the production of phaseolotoxin by three- to fourfold (Aslamkhan 1992).

The presence of an AraC family signature, a helix-turn-helix domain, and a possible leucine zipper in ORF5 suggests that the product of ORF5 may be a transcriptional factor. Most of the transcriptional regulators in the AraC family are individually transcribed; however, there are several examples, such as the *celD* gene of *E. coli* (Parker and Hall 1990), *lumQ* gene of *Photobacterium leiognathi* (Lin et al. 1995), *tcpN* gene of *Vibrio cholerae* (Ogiorman and Manning 1992), *perA* gene of *E. coli* (Gomez-Duarte and Kaper 1995), *pocR* gene of *Salmonella typhimurium* (Chen et al. 1994), and *exsA* gene of *Pseudomonas aeruginosa* (Frank and Iglewski 1991), where the regulators are embedded in the operon, similar to ORF5 in the *phtE* gene cluster. The existence of a possible leucine zipper in ORF5 indicates complexity of function since leucine zippers are found mainly in eukaryotic proteins, although they have been found in the σ^{54} (Sasse-Dwight and Gralla 1990) and the MetR (Weissbach and Brot 1991) proteins produced by *E. coli*. This may indicate that the ORF5 product may be able to form a homodimer or a heterodimer with another pro-

tein, as is the case with most members of the AraC family (Bustos and Schleif 1993; Gallegos et al. 1993).

ORF6 has some homology to *desA*, a gene encoding fatty acid desaturase in the cyanobacterium, *Synechocystis* Pcc6803. Independently, Hatziloukas et al. (1995) have also reported sequence data from part of the *tox-argK* gene cluster isolated from another strain of *P. syringae* pv. *phaseolicola* that has homology to *desA* (see Results). Although the amino acid identity between ORF6 and *desA* genes is not high, the striking resemblance in their hydropathy profiles strongly suggests that the ORF6 product has a fatty acid desaturase-like function. Fatty acid desaturase (or steroyl-CoA desaturase) generates unsaturated fatty acids for the synthesis of phospholipids located in the cytoplasmic membrane (Wada et al. 1990). Phospholipids consist of saturated and/or unsaturated fatty acids and the content of the unsaturated fatty acids affects the fluidity of the membrane. The higher the unsaturated fatty acid content in phospholipids, the more fluid the membrane, possibly due to the double bonds in the unsaturated fatty acids that prevent close packing of the acyl chains in the phospholipids (Cronan et al. 1987). Higher fluidity of the cell membrane facilitates transport of solutes across the membrane (Cronan and Rock 1987). Interestingly, the desaturase activity in many organisms, including *E. coli* (Cronan et al. 1987), the cyanobacterium *Synechocystis* Pcc6803 (Wada et al. 1990; Wada and Murata 1990), *Bacillus* (Dunkley et al. 1991), soybean (Cheesbrough 1990), *Tetrahymena pyriformis* (Skriver and Thompson 1979), and *Pseudomonas* E-3 (Wada et al. 1987) is temperature-dependent, i.e., the activity is low at higher temperature (37 to 45°C) but higher at lower temperature (18 to 25°C). It is interesting that the optimum temperature for toxin production in *P. syringae* pv. *phaseolicola* is around 18°C, which is also within the temperature range at which most desaturases, and possibly the ORF6 product, show the highest activity. Recently, our laboratory has isolated a mutant that contains a *TnphoA* insertion in ORF6 (K. B. Rowley and S. S. Patil, unpublished). This supports the suggestion that the ORF6 product is a membrane-associated protein.

That ORF6 is transcribed *in vivo* is shown by the data in Figure 3. When total RNA from *P. syringae* pv. *phaseolicola* grown at the permissive temperature was used as a template along with primer pairs 10 and 10' and 11 and 11' (both of which are within ORF6; Fig. 2) in an RT-PCR experiment, products of expected sizes were detected (Fig. 3, lanes 10 and 11). That the expression of ORF6 is essential for Tox⁺ phenotype of the wild-type strain can be assumed from the Tox⁻ phenotype of the previously reported (Zhang et al. 1993) marker exchange mutant HOA14, which has a Tn5 insertion located close to the 3' prime end of ORF6.

No specific function has been proposed for the products of the remaining ORFs because no significant homologous gene sequences were found in the data bases. Although ORF2 and ORF4 showed some homology with certain genes, the similarities were too limited to be significant. ORF1, ORF2, and ORF4 all have more than one membrane-spanning region, indicating that the gene products of these ORFs are membrane associated, and may be involved in the secretion of the toxin.

MATERIALS AND METHODS

Bacterial strains and plasmids.

P. syringae pv. *phaseolicola* 4612-1 is a Rif^r derivative of

the wild-type Tox⁺ strain PDDCC 4612 (Kamdar et al. 1991). pHK120 contains a 25-kb genomic DNA insert from the wild-type *P. syringae* pv. *phaseolicola* strain G50-1 that is Tox⁺ and Rif^r, in the vector pLAFR3 (Kamdar et al. 1991). pYZH102 contains a 7.0-kb *Hind*III fragment in pLAFR3, pYZH202 contains a 4.9-kb *Hind*III fragment in pLAFR3, and pYZHB37 contains a 3.7-kb *Hind*III-*Bam*HI fragment in the vector pT7/T3 α -18 (Fig. 1; Zhang et al. 1993).

DNA sequencing.

DNA was sequenced by the dideoxy-nucleotide chain termination method (Sanger et al. 1977), with the modified T7 DNA polymerase (Sequenase 2.0, Promega, Madison, WI). Subclones were generated in pT7/T3 α -18 or pT7/T3 α -19 and the inserts were sequenced from both ends, utilizing the primer binding sites in the vectors. A nested set of deletion clones was produced by digesting pYZHB37 with *Bam*HI and *Kpn*I, followed by digestion with *Exo*III (Henikoff 1984). The deletion clones were sequenced with the T7 sequencing primer. Synthetic primers (15-mers) were prepared with a 380B DNA synthesizer (Perkin-Elmer-Applied Biosystems, Branchburg, NJ) and used for sequencing most of the DNA in the *phtE* locus. Both strands of the DNA were sequenced with dGTP but some regions were sequenced with dITP to overcome compression caused by the high G+C content of some of the DNA inserts. After our work was completed, Hatziloukas et al. (1995) published sequence data on part of the *phtE* locus that differed from our sequence. To resolve these differences, we designed new primers for re-sequencing the relevant regions (see Results) and performed cycle sequencing with an automatic sequencer (Model 373A; Perkin-Elmer-Applied Biosystems, Branchburg, NJ).

Isolation of total RNA.

Total RNA was extracted from *P. syringae* pv. *phaseolicola* or its derivatives according to the method of Innes et al. (1993). Overnight bacterial cultures in KB medium were subcultured to an optical density of 0.3 (600 nm) and shaken at the appropriate temperature for 4 h. Ten-milliliter bacterial cells were harvested by centrifugation and quick frozen in dry ice. The cell pellets were resuspended in 1 ml of hot lysis buffer (50 mM Tris-HCl [pH 9.0], 50 mM EDTA, 300 mM sodium acetate, 0.625 % [wt/vol] sodium dodecyl sulfate) and 1 ml of hot phenol (65°C), mixed and boiled for 5 min. After centrifuging at 700 rpm (bench top clinical centrifuge, International Equipment, Needham, MA) for 5 min, the upper aqueous phase was removed and extracted twice with 1 volume of phenol, twice with 1 volume of phenol:chloroform, and once with 1 volume of chloroform. The RNA was precipitated with 2 volumes of cold EtOH and dried under vacuum. The pellets were resuspended in 96 μ l of DNase digestion buffer (20 mM Tris-HCl [pH 8.0], 10 mM MgCl₂) and digested with 4 μ l (40 units) of RNase-free DNase I (Boehringer Mannheim, Indianapolis, IN) at 37°C for 15 min. The DNase-treated RNA was extracted with phenol and chloroform, precipitated with EtOH, and dried as described above.

RT-PCR amplification of mRNA.

One to 3 μ g of total RNA was used for the synthesis of cDNA by reverse transcriptase (Perkin Elmer, Nutley, NJ) with 2.5 pmol of random hexamers as the primer. To eliminate DNA

contamination (Dilworth and McCarrey 1992), the RNA samples were treated with 40 units of RNase-free DNase I (Boehringer Mannheim, Indianapolis, IN) in a solution containing 5 mM MgCl₂, 50 mM KCl, 10 mM Tris-HCl, 1 mM of dATP, dGTP, dTTP, and dCTP at 37°C for 1 h. The reaction mixtures were immediately heated at 99°C for 5 min to inactivate the remaining DNase I and cooled to 4°C. One microgram each of reverse transcriptase and RNase inhibitor was added to each of the reaction tubes, which were incubated at 42°C for 30 min. Primer pairs (15-mers), one identical and the other complementary to the mRNA sequence, were used in the amplification of the randomly generated cDNA with 25 units of *Taq* DNA polymerase (Perkin Elmer, Nutley, NJ) in a solution containing 2 mM MgCl₂, 50 mM KCl, 10 mM Tris-HCl, 0.2 mM dATP, dGTP, dCTP, and dTTP in a final volume of 100 µl. PCR was performed for 40 cycles with a setting of 1 min at 95°C, 1 min at 55°C, and 1 min at 72°C with the 9600 PCR Thermal Cycler (Perkin Elmer, Nutley, NJ). Twenty microliters of each of the PCR products was analyzed for the expected fragment on a 2% agarose gel. Southern blot analysis was performed with specific probes to confirm the PCR results.

Primer extension.

One microliter (about 100 ng) of primer was labeled with ³²P-dCTP and 10 units of T4 polynucleotide kinase in a solution containing 50 mM Tris-HCl [pH 7.5], 10 mM MgCl₂, 5 mM DTT, 0.1 mM spermidine at 37°C for 30 min. The labeled primer and 10 µg of total RNA were mixed together and incubated at 70°C for 10 min, and then immediately chilled on ice. cDNA synthesis (primer extension) was carried out in a solution containing 50 mM Tris-HCl [pH 8.3], 75 mM KCl, 3 mM MgCl₂, 200 mM dithiothreitol (DTT), and 200 units of SuperScript RT (GIBCO BRL, Gaithersburg, MD) at 42°C for 1 h. The reaction was terminated by adding 4 µl of 0.5 M EDTA. A 10-µl aliquot of each primer extension reaction was run on a 6% polyacrylamide gel along with a sequencing ladder, which was generated with the same primer in a sequencing reaction according to the standard chain-termination method of Sanger et al. (1977).

Computer-assisted analyses.

All computer analysis of DNA sequences was done with the GCG (Genetic Computer Group, Madison WI) program. Codon preference was analyzed with a codon frequency table for *Pseudomonas* genes. Searches for homologous sequences (Heijne 1987; Lipman and Pearson 1985) were conducted with Gene-EMBL data bases at both the DNA and the amino acid levels.

ACKNOWLEDGMENTS

Y.Z. was supported by the Water Resources Research Center of the University of Hawaii during this work. We thank Ed Bylina for useful discussions, Neil Reimer, Manager of the University of Hawaii Core Facility for help with the use of GCG programs and for providing synthetic primers, and Karla Rowley for carefully reading the manuscript and for discussions. We appreciate the help of Xue Mei Ou and Bonnie Akaka in the preparation of the manuscript.

LITERATURE CITED

Aarts, M. G., Dirkse, W. G., Stiekema, W. J., and Pereira, A. 1993. Transposon tagging of a male sterility gene in Arabidopsis. *Nature*

- 363:715-717.
- Alonso, B., and Rubio, V. 1989. Participation of ornithine aminotransferase in the synthesis and catabolism of ornithine in mice. *Biochem. J.* 259:131-138.
- Altshul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.
- Aslamkhan, A. 1992. Enzymology of the biosynthesis of the peptide moiety of phaseolotoxin. M.S. thesis. University of Hawaii, Honolulu, HI.
- Beckwith, J. 1987a. The operon: A historical account. Pages 1439-1443 in: *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology. G. C. Neidhardt, J. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger, eds. American Society for Microbiology, Washington, D.C.
- Beckwith, J. 1987b. The lactose operon. Pages 1444-1452. in: *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology. G. C. Neidhardt, J. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger, eds. American Society for Microbiology, Washington, D.C.
- Bhella, R. S., and MacKenzie, S. L. 1995. Nucleotide sequence of cDNA from *Limnanthes douglasii* L. encoding a Δ-15 linoleic acid desaturase. *Plant. Physiol.* 108:861
- Billheimer, J. T., Carnevale, H. N., Leisinger, T., Eckhardt, T., and Jones, E. E. 1976. Ornithine-transaminase activity in *Escherichia coli*: Its identity with acetylornithine-transaminase. *J. Bacteriol.* 127:1315-1323.
- Bustos, S. A., and Schleif, R. F. 1993. Functional domains of the AraC protein. *Proc. Natl. Acad. Sci. USA* 90:5638-5642.
- Calogero, S., Gardan, R., Glaser, P., Schweizer, J., Rapoport, G., and Debarbouille, M. 1994. RocR, a novel regulatory protein controlling arginine utilization in *Bacillus subtilis*, belongs to the NtrC/NifA family of transcriptional activators. *J. Bacteriol.* 176:1234-1241.
- Cheesbrough, T. M. 1990. Decreased growth temperature increases soybean stearoyl-acyl carrier protein desaturase activity. *Plant Physiol.* 93:555-559.
- Chen, P., Andersson, D. I., and Toth, J. R. 1994. The control region of the *pdu/cob* regulon in *Salmonella typhimurium*. *J. Bacteriol.* 176:5474-5482.
- Chepur, V., Lemieux, L., Au, C. T., and Gennis, R. B. 1990. The sequence of the *cyo* operon indicates substantial structural similarities between the cytochrome *o* ubiquinol oxidase of *Escherichia coli* and the *aa₃*-type family of Cytochrome *c* oxidases. *J. Biol. Chem.* 265:11185-11192.
- Cronan, J. E., Gennis, R. B., and Maloy, S. R. 1987. Cytoplasmic membrane. Pages 31-55 in: *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology. G. C. Neidhardt, J. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger, eds. American Society for Microbiology, Washington, D.C.
- Cronan, J. E., and Rock, C. O. 1987. Biosynthesis of membrane lipids. Pages 474-497 in: *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology. G. C. Neidhardt, J. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger, eds. American Society for Microbiology, Washington, D.C.
- Degols, G. 1987. Functional analysis of the regulatory region adjacent to the *cargB* gene of *Saccharomyces cerevisiae*. Nucleotide sequence, gene fusion experiments and cis-dominant regulatory mutation synthesis. *Eur. J. Biochem.* 169:193-200.
- Degols, G., Jauniaux, J. C., and Waime, J. M. 1987. Molecular characterization of transposable-element-associated mutations that lead to constitutive L-ornithine aminotransferase expression in *Saccharomyces cerevisiae*. *Eur. J. Biochem.* 165:289-296.
- Delauney, A. J., Hu, C. A. A., Kishor, P. B. K., and Verma, D. P. S. 1993. Cloning of ornithine delta-aminotransferase cDNA from *Vigna aconitifolia* by trans-complementation in *Escherichia coli* and regulation of proline biosynthesis. *J. Biol. Chem.* 268:18637-18678.
- Deretic, V., Konyecsni, W. M., Mohr, C. D., Martin, D. W., and Hibler, N. S. 1989. Common denominators of promoter control in *Pseudomonas* and other bacteria. *Bio/technology* 7:1249-1254.
- Dilworth, D. D., and McCarrey, J. R. 1992. Single-step elimination of contaminating DNA prior to reverse transcriptase PCR. *PCR Methods Appl.* 1:279-282.
- Dunkley, E. A., Jr., Clejan, S., and Krulwich, T. A. 1991. Mutants of *Bacillus* species isolated on the basis of protonophore resistance are deficient in fatty acid desaturase activity. *J. Bacteriol.* 173:7750-7755.

- Frank, D. W., and Iglewski, B. H. 1991. Cloning and sequence analysis of a trans-regulatory locus required for exoenzyme synthesis in *Pseudomonas aeruginosa*. *J. Bacteriol.* 173:6460-6468.
- Gallegos, M. T., Michan, C., and Ramos, J. L. 1993. The XylS/AraC family of regulators. *Nucleic Acids Res.* 21:807-810.
- Gardan, R., Rapoport, G., and Debarbouille, M. 1995. Expression of the *rocDEF* operon involved in arginine catabolism in *Bacillus subtilis*. *J. Mol. Biol.* 249:843-856.
- Gomez-Duarte, O. G., and Kaper, J. B. 1995. A plasmid-encoded regulatory region activates chromosomal *eaeA* expression in enteropathogenic *Escherichia coli*. *Infect. Immun.* 63:1767-1776.
- Hamada, T., Kodama, H., Nishimura, M., and Iba, K. 1994. Cloning of a cDNA encoding tobacco omega-3-fatty acid desaturase. *Gene* 147:293-294.
- Hatziloulas, E., Panopoulos, N. J., Delis, S., Prosen, D., and Schaad, N. W. 1995. An open reading frame in the approximately 28-kb *tox-argK* gene cluster encodes a polypeptide with homology to fatty acid desaturase. *Gene* 166:83-87.
- Heijne, G. 1987. *Sequence Analysis in Molecular Biology*. Academic Press, New York.
- Heimberg, H., Boyen, A., Crabeel, M., and Glansdorff, N. 1990. *Escherichia coli* and *Saccharomyces cerevisiae* acetylornithine aminotransferase: Evolutionary relationship with ornithine aminotransferases. *Gene* 90:69-78.
- Henikoff, S. 1984. Unidirectional digestion with *ExoIII* creates targeted breakpoints for DNA sequencing. *Gene* 28:351-359.
- Hsu, L. C., Okamoto, M., and Snell, E. E. 1989. L-Histidinol phosphate aminotransferase from *Salmonella typhimurium*. Kinetic behavior and sequence at the pyridoxal-P binding site. *Biochimie* 71:477-489.
- Inana, G., Totsuka, S., Redmond, M., Dougherty, T., Nagle, J., Shiono, T., Ohura, T., Kominami, E., and Katunuma, N. 1986. Molecular cloning of human ornithine aminotransferase mRNA. *Proc. Natl. Acad. Sci. USA* 83:1203-1207.
- Innes, R. W., Bent, A. F., Kunkel, B. N., Bisgrove, S. R., and Staskawicz, B. J. 1993. Molecular analysis of avirulence gene *avrRpt2* and identification of a putative regulatory sequence common to all known *Pseudomonas syringae* avirulence genes. *J. Bacteriol.* 175:4859-4869.
- Joerger, R. D., Jacobson, M. R., and Bishop, P. E. 1989. Two *nifA*-like genes required for expression of alternative nitrogenases by *Azotobacter vinelandii*. *J. Bacteriol.* 171:3258-3267.
- Kamdar, H. V., Rowley, K. B., Clements, D., and Patil, S. S. 1991. *Pseudomonas syringae* pv. *phaseolicola* genomic clones harboring heterologous DNA sequences suppress the same phaseolotoxin-deficient mutants. *J. Bacteriol.* 173:1073-1079.
- Kirsch, J. F., Eichele, G., Ford, G. C., Vincent, M. G., and Jansonius, J. N. 1984. Mechanism of action of aspartate aminotransferase proposed on the basis of its spatial structure. *J. Mol. Biol.* 174:497-525.
- Lin, J. W., Yu, K. Y., Chao, Y. F., and Weng, S. F. 1995. The *lumQ* gene is linked to the *lumP* gene and the *lux* operon in *Photobacterium leiognathi*. *Biochem. Biophys. Res. Comm.* 217:684-695.
- Lipman, D. J., and Pearson, W. R. 1985. Rapid and sensitive protein similarity searches. *Science* 227:1435-1441.
- Marahiel, M. A. 1992. Multidomain enzymes involved in peptide synthesis (minireview). *FEBS Lett.* 307:40-43.
- Martin, J. F., and Liras, P. 1989. Organization and expression of genes involved in the biosynthesis of antibiotics and other secondary metabolites. *Annu. Rev. Microbiol.* 43:173-206.
- Misawa, N., Satomi, Y., Kondo, K., Yokoyama, A., Kajiwar, S., Saito, T., Ohtani, T., and Miki, W. 1995. Structure and functional analysis of a marine bacterial carotenoid biosynthesis gene cluster of astaxanthin biosynthetic pathway proposed at the gene level. *J. Bacteriol.* 177:6575-6584.
- Moore, R. E., Niemczura, W. P., Kwok, O. C. H., and Patil, S. S. 1984. Inhibitors of ornithine carbamoyltransferase from *Pseudomonas syringae* pv. *phaseolicola*: Revised structure of phaseolotoxin. *Tetrahedron Lett.* 25:3931-3934.
- Ogierman, M. A., and Manning, P. A. 1992. Homology of TcpN, a putative regulatory protein of *Vibrio cholerae*, to the AraC family of transcriptional activators. *Gene* 116:93-97.
- Okuley, J., Lightner, J., Feldman, K., Yadav, N., Larke, E., and Browse, J. 1994. Arabidopsis *fad2* gene encodes the enzyme that is essential for polyunsaturated lipid synthesis. *Plant Cell* 6:147-158.
- O'Reilly, M., and Devine, K. M. 1994. Sequence and analysis of the citrulline biosynthetic operon *argC-F* from *Bacillus subtilis*. *Microbiology* 140:1023-1025.
- Oyama, R., Suzuki, M., Matsuzawa, T., and Titani, K. 1990. Complete amino acid sequence of rat kidney ornithine aminotransferase: Identity with liver ornithine aminotransferase. *J. Biochem.* 108:133-138.
- Parker, L. L., and Hall, B. G. 1990. Characterization and nucleotide sequence of the cryptic *cel* operon of *Escherichia coli* K12. *Genetics* 124:455-471.
- Patil, S. S. 1974. Toxins produced by phytopathogenic bacteria. *Annu. Rev. Phytopathol.* 12:259-279.
- Raibaud, A., Zalacain, M., Holt, T. G., Tizard, R., and Thompson, C. J. 1991. Nucleotide sequence analysis reveals linked *N*-acetyl hydrolase, thioesterase, transport, and regulatory genes encoded by the bialaphos biosynthetic gene cluster of *Streptomyces hygroscopicus*. *J. Bacteriol.* 173:4454-4463.
- Ream, W. 1989. *Agrobacterium tumefaciens* and interkingdom genetic exchange. *Annu. Rev. Phytopathol.* 27:583-618.
- Rowley, K. B., Clements, D. E., Mandel, M., Humphreys, T., and Patil, S. S. 1993. Multiple copies of a DNA sequence from *Pseudomonas syringae* pv. *phaseolicola* abolish thermoregulation of phaseolotoxin production. *Mol. Microbiol.* 8:625-635.
- Sakamoto, T., Wada, H., Nishida, I., Ohmore, M., and Murata, N. 1994. Identification of conserved domains in the delta-12 desaturases for cyanobacteria. *Plant Mol. Biol.* 24:643-650.
- Sanger, F., Nicklen, S., and Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.
- Sasse-Dwight, S., and Gralla, J. D. 1990. Role of eukaryotic-type functional domains found in the prokaryotic enhancer receptor factor σ^{54} . *Cell* 62:945-954.
- Singh, S., van der Heidi, T., McKinney, S., and Green, A. 1995. Nucleotide sequence of a cDNA from *Brassica juncea* encoding a microsomal Omega-6 desaturase (Accession No. x91139) (PGR 95-107). *Plant Physiol.* 109:1498.
- Skriver, L., and Thompson, G. A., Jr. 1979. Temperature-induced changes in fatty acid unsaturation of *Tetrahymena* membranes do not require induced fatty acid desaturase synthesis. *Biochem. Biophys. Acta* 572:376-381.
- To, K. Y., Lai, E. M., Lee, L. Y., Lin, T. P., Hung, C. H., Chen, C. L., Chang, Y. S., and Liu, S. T. 1994. Analysis of the gene cluster encoding carotenoid biosynthesis in *Erwinia herbicola* Eho13. *Microbiology* 140:331-339.
- van der Wauven, C., Jann, A., Haas, D., Leisinger, T., and Stalon, V. 1988. N^{δ} -succinylornithine in ornithine catabolism of *Pseudomonas aeruginosa*. *Arch. Microbiol.* 150:400-404.
- van der Wauven, C., and Stalon, V. 1985. Occurrence of succinyl derivatives in the catabolism of arginine in *Pseudomonas cepacia*. *J. Bacteriol.* 164:882-886.
- Voellmy, R., and Leisinger, T. 1975. Dual role for N^{δ} -acetylornithine 5-aminotransferase from *Pseudomonas aeruginosa* in arginine biosynthesis and arginine catabolism. *J. Bacteriol.* 122:799-809.
- Wada, M., Fufunaga, N., and Sasaki, S. 1987. Effect of growth temperature on phospholipid and fatty acid compositions in a psychrotrophic bacterium, *Pseudomonas* sp. Strain E-3. *Plant Cell Physiol.* 28:1209-1217.
- Wada, H., Gombos, Z., and Murata, N. 1990. Enhancement of chilling tolerance of a cyanobacterium by genetic manipulation of fatty acid desaturation. *Nature* 347:200-203.
- Wada, H., and Murata, N. 1990. Temperature-induced changes in the fatty acid composition of the cyanobacterium, *Synechocystis* PCC6803. *Plant Physiol.* 92:1062-1069.
- Weissbach, H., and Brot, N. 1991. Regulation of methionine synthesis in *Escherichia coli*. *Mol. Microbiol.* 5:1593-1597.
- Zhang, Y. X. 1994. Genetic organization of a cluster of genes involved in the production of phaseolotoxin, a phytotoxin produced by *Pseudomonas syringae* pv. *phaseolicola* and sequence analysis of the *phtE* locus. Ph.D. thesis. University of Hawaii, Honolulu, HI.
- Zhang, Y. X., Rowley, K. B., and Patil, S. S. 1993. Genetic organization of a cluster of genes involved in the production of phaseolotoxin, a phytotoxin produced by *Pseudomonas syringae* pv. *phaseolicola*. *J. Bacteriol.* 175:6451-6458.