JOURNAL OF BACTERIOLOGY, Dec. 1988, p. 5479–5488
0021-9193/88/125479-10$02.00/0

Vol. 170, No. 12

Molecular Analysis of a Pathogenicity Locus in *Pseudomonas syringae* pv. *syringae*†

PRADIP MUKHOPADHYAY, JANET WILLIAMS, AND DALLICE MILLS*

Department of Botany and Plant Pathology, Oregon State University, Corvallis, Oregon 97331

Received 1 June 1988/Accepted 22 August 1988

One of the chromosomal regions of *Pseudomonas syringae* pv. *syringae* encoding pathogenicity factors had been mapped into a 3.9-kilobase-pair fragment in previous studies. Promoter probe analysis indicated the existence of a promoter near one end of the fragment. DNA sequencing of this fragment revealed the existence of a consensus promoter sequence in the region of the promoter activity and two open reading frames (ORFs) downstream. These ORFs, ORF1 and ORF2, encoded putative polypeptides of 40 and 83 kilodaltons, respectively. All ORF1::Tn*5* as well as ORF2::Tn*5* mutant strains were nonpathogenic on susceptible host bean plants and were unable to elicit hypersensitive reactions on nonhost tobacco plants. The deduced amino acid sequence of the 83-kilodalton polypeptide contained features characteristic of known integral membrane proteins. Fusion of the *lacZ* gene to ORF2 led to the expression of a hybrid protein inducible in *Escherichia coli*. The functions of the putative proteins encoded by ORF1 and ORF2 are unknown at present.

Genetic studies of determinants of pathogenicity in phytopathogenic bacteria have been undertaken for a wide variety of organisms (23, 32) including *Pseudomonas* (2, 5, 8, 20, 39, 41). Isolation of mutants affected in their behavior on plants led to the identification of different types of genes involved in plant-pathogen interactions. The genes required for both the expression of disease symptoms on host plants and the development of the hypersensitive reaction on nonhost plants (16) have been designated *hrp* genes, whereas the name *dsg* has been attributed to genes responsible only for disease development (20). The hypersensitive reaction is considered a generalized expression of resistance by the plants to pathogens and is associated with only limited multiplication and spread of the pathogen surrounding the infected area (15). On the other hand, pathogenicity toward the susceptible host is considered to be the ability of the pathogen to establish itself in the plant, resulting in rapid multiplication and widespread invasion by the organism. Hence, characterization of these *hrp* genes should reveal important aspects of the plant-pathogen interactions.

The *Pseudomonas syringae* group of phytopathogenic bacteria contains various pathogens that cause diseases on the foliage of plants (12). Among these pathogens is an ecotype of *Pseudomonas syringae* pv. *syringae*, which is the causal agent of brown spot disease of *Phaseolus vulgaris* L., the common bean. Several regions of the bacterial genome that are involved in the pathogenicity of *P. syringae* pv. *syringae* R32 were identified in our laboratory by Tn*5* mutagenesis (2). One of these mutants, PS9021, failed to incite disease symptoms on bean and to cause a hypersensitive reaction on nonhost plants such as tobacco. The mutant did not grow in planta (4) but grew on minimal agar medium, and, unlike the parental strain, which has firm and smooth colonies, it exhibited mucoidal colony morphology. Hence, the mutant appeared to be affected in one or more *hrp* genes. In this paper we describe the complete DNA sequence of a *hrp* locus affected in the mutant, transcriptional analysis of the pathogenicity region in the cosmid pOSU3105 (30), and construction of a fusion protein for raising antibodies against a pathogenicity determinant of this *hrp* locus.

* Corresponding author.

† Technical paper no. 8625 from the Oregon Agricultural Experiment Station.

MATERIALS AND METHODS

Bacterial strains and plasmids. Bacterial strains and plasmids used in this study are listed in Table 1.

Media and antibiotics. *P. syringae* pv. *syringae* strains were grown at 28°C in MaNY medium as previously described (2); *Escherichia coli* strains were grown at 37°C in LB medium. The media were supplemented with the following antibiotics (obtained from Sigma Chemical Co.) as required: kanamycin, 50 μg/ml; penicillin, 250 μg/ml; tetracycline, 15 μg/ml; and chloramphenicol, 20 μg/ml. MacConkey agar medium was used for detecting galactokinase expression and was obtained from Difco Laboratories. L-agar plates were poured with 40 μg of the chromogenic substrate 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) per ml and 200 mM isopropyl-β-D-thiogalactopyranoside (IPTG) to allow visualization of β-galactosidase expression.

DNA manipulations. Restriction enzymes, T4 DNA ligase, and exonuclease III were obtained from Bethesda Research Laboratories, Inc., and used as specified by the manufacturer. T4 DNA polymerase (New England BioLabs, Inc.) was used to fill in protruding 3′ termini (19) of DNA. The Klenow fragment of DNA polymerase I (International Biotechnologies, Inc.) and [^{35}S]dATP (Du Pont, NEN Research Products) were used for DNA sequencing reactions. Mung bean nuclease and synthetic *Nco*I linkers were obtained from Pharmacia, Inc.

DNA sequencing. Portions of the *hrp* locus identified by the mutation in PS9021 were subcloned (see Results) into the bacteriophage vector M13mp18 (42) in both orientations to allow sequencing of each strand of DNA. Sequential overlapping deletions of DNA in each clone were generated by the exonuclease III-mung bean nuclease method of Henikoff (13). Each of these new clones was sequenced by the dideoxynucleotide chain termination method (36). Problems arising from compressions in the gels were overcome by substituting 7-deazaguanosine-5′-triphosphate for dGTP (26). The DNA sequence and deduced amino acid sequences

TABLE 1. Bacterial strains and plasmids

Strain or plasmid	Genotype or phenotype	Source or reference
E. coli		
HB101	*proA2 leuB6 thi-1 lacY1 hsdR hsdM recA13 supE44 rpsL20 galK2 arg-14*	21
N100	*galK2 recA13*	22
JM105	*thi rpsL endA sbcB15 hspR4* Δ(*lac-proAB*) (F′ *traD36 proAB lacI*^{q}*ZΔM15*)	42
P. syringae pv. *syringae*		
R32	Wild-type pathogenic isolate from bean	2
PS9020	Spontaneous Sm^r derivative of R32	2
PS9021	Nonpathogenic mutant of PS9020 obtained by Tn*5* mutagenesis	2, 30
PS3150 to PS3162	Derivatives of PS9020 with different Tn*5* insertions on the chromosome	24, 25
Plasmids		
pKO4	pBR322 derivative with promoterless *galK* gene	22
pKO6	Same as pKO4, but with polylinker site reversed	22
pIJ3100	RSF1010 derivative with promoterless *cat* gene	31
pOSU3105	Broad-host-range cosmid pVK102 with *P. syringae* pv. *syringae* DNA encompassing the *hrp* locus	30
pOSU3125	pBR322 derivative with left *Hin*dIII fragment cloned between promoter for *lac* operon and *galK* gene	25
pOSU3126	Same as pOSU3125, but with the *Hin*dIII fragment in reverse orientation	25
pOSU4101	pBR322 derivative with *tac* promoter, consensus RBS, and truncated *hrpM* locus with 8-bp *Nco*I linker	This study
pOSU4102	Same as pOSU4101, but with 10-bp *Nco*I linker	This study
pOSU4103	Same as pOSU4101, but with 12-bp *Nco*I linker	This study
pMLB1034	pBR322 derivative with polylinker site at codon 8 of *lacZ* gene	38
pOSU4104	pMLB1034 with the *P. syringae* pv. *syringae* DNA insert from pOSU4101	This study
pOSU4105	pMLB1034 with the *P. syringae* pv. *syringae* DNA insert from pOSU4102	This study
pOSU4106	pMLB1034 with the *P. syringae* pv. *syringae* DNA insert from pOSU4103	This study

were analyzed by using computer programs described by Mount and Conrad (7, 27).

Immunoblot analysis. Overnight cultures of the strains to be analyzed were diluted 100-fold in L broth with penicillin and grown to an A_{600} of 0.5 to 0.6. At this time, 2 mM IPTG was added to each culture, and the cultures were grown for an additional 1 h. A 1-ml sample of each culture was microcentrifuged for 1 min, and the pellet was suspended in 200 μl of the sample buffer (0.625 M Tris [pH 6.8], 2% sodium dodecyl sulfate, 10% glycerol, 2% β-mercaptoethanol) and boiled for 3 min. A 10-ml sample of each lysate was subjected to sodium dodecyl sulfate–7% polyacrylamide gel electrophoresis as described by Laemmli (18). After separation, the protein bands were transferred to nitrocellulose filters (pore size, 0.2 μm; Schleicher & Schuell, Inc.) in a high-field electroblotting apparatus (Bio-Rad Laboratories) as previously described (9). Prestained molecular size standards (Bethesda Research Laboratories) were used to mark migrations of proteins on the blot. The transblotted filter was probed with a 1:10,000 dilution of mouse anti-β-galactosidase antibody (Promega Biotec, Inc.). Immunochemical staining of the bands was performed with an alkaline phosphate-conjugated protoblot system (Promega Biotec) designed for use with mouse antiserum. At present, the hybrid protein is being purified in preparative amounts.

RESULTS

The pathogenicity locus affected in PS9021 is of interest for two principal reasons. First, the mutant is prototrophic and is able to utilize the same range of sugars and nitrogen sources as the wild type (data not shown), but is unable to grow in planta. Second, the mutant is altered in colony morphology, indicating that the gene(s) mutated may be involved in the synthesis of surface-associated or extracellular products, which, in turn, may have a role in the recognition essential for bacterial pathogenicity. Previous results (30) revealed a DNA sequence from a cosmid library of *P. syringae* pv. *syringae* R32 that complemented all of the altered phenotypes of PS9021. The approximate extent of this *hrp* locus was determined by site-directed mutagenesis with Tn*5* in *E. coli* (24), followed by marker exchange of these mutations into the chromosome of PS9020, a strain isogenic to R32. Among the resulting strains was a group of six Hrp$^-$ mutants, PS3151 to PS3156 (Table 1; Fig. 1), that mapped within a 3.9-kilobase-pair (kb) *Hin*dIII fragment. Since the Tn*5* insertions in PS9021 and PS3153 mapped to approximately the same site (24, 30), this *hrp* locus was thought to be affected in the original mutant PS9021. Two proteins of approximately 37 and 85 kilodaltons (kDa) (25) (Fig. 1D) have been expressed from this locus in one direction in *E. coli* maxicells.

Detection of promoter activity in the locus. Various derivatives of the 3.9-kb *Hin*dIII fragment (Fig. 1B) encompassing the *hrp* locus were generated by using *Sal*I and *Bgl*II sites present in the fragment. These smaller fragments were cloned in both orientations in the appropriate polylinker sites upstream of a promoterless galactokinase gene (*galK*) in the promoter probe plasmids, pKO4 and pKO6 (22). These clones were transformed into the indicator strain, *E. coli* N100, and tested on MacConkey plates with galactose as the sole carbon source. The clones with promoter activity formed red colonies. Promoter activity was detected in the *Hin*dIII-*Sal*I and *Hin*dIII-*Bgl*II (Fig. 1B) fragments from the right end of the locus, and the direction of transcription was inward (Fig. 1C). To verify that the promoter activity detected in *E. coli* is truly representative of that in *P. syringae* pv. *syringae*, the two *Hin*dIII-*Bgl*II fragments (Fig. 1B) generated from the 3.9-kb fragment and various deriva-

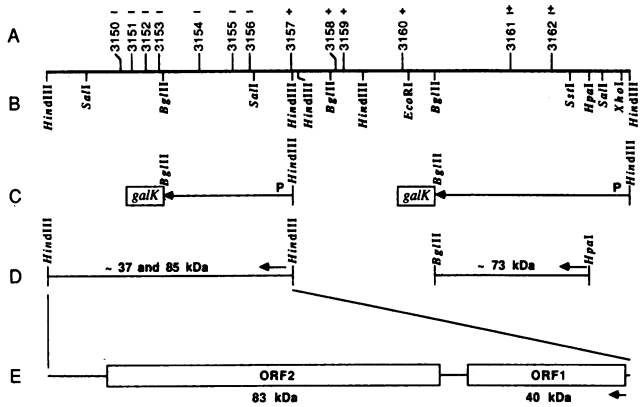

FIG. 1. Genetic structure of the 8.5-kb DNA fragment of *P. syringae* pv. *syringae* R32 present on the cosmid pOSU3105. The terminal *Hin*dIII fragments are 3.9 kb each and are separated by a 0.7-kb *Hin*dIII fragment. (A) Sites of Tn*5* insertions following marker exchange into the chromosome of *P. syringae* pv. *syringae* R32 and corresponding phenotypes on plants. Symbols: +, Hrp^+; −, Hrp^-; ±, retarded disease symptoms on bean, cultivar Red Mexican; 3150 to 3162, designations for the sites of Tn*5* insertions. (B) Restriction map of the 8.5-kb DNA fragment. (C) Transcriptional analysis of the 8.5-kb fragment in *E. coli* with a promoterless *galK* gene. Arrows indicate directions of transcription, and the box indicates the location of the *galK* gene in each case. (D) Expression of proteins from various subclones of the 8.5-kb fragment. Arrows indicate the directions of expression, and the estimated sizes of these proteins are shown above the respective clones. (E) ORFs present in the left half of the 8.5-kb fragment. The direction of transcription is shown by the arrow. The sizes of putative polypeptides encoded by these ORFs are indicated below the respective ORFs.

tives of these fragments were individually cloned into a polylinker site upstream of a promoterless chloramphenicol transacetylase gene in the broad-host-range plasmid pIJ3100 (31). The clones obtained were transformed (28) into *P. syringae* pv. *syringae* R32 to allow selection for the Sm^r marker present on pIJ3100, and individual colonies were then tested for the chloramphenicol resistance phenotype to detect any promoter activity present in the cloned fragment. The clones that carried a functional promoter in *E. coli* also conferred resistance to 15 to 20 μg of chloramphenicol per ml in *P. syringae* pv. *syringae*. These results suggest that promoter activity, functional in both *E. coli* and *P. syringae* pv. *syringae*, resides near the right end of the 3.9-kb fragment.

DNA sequence analysis. The plasmid pOSU3126 (Table 1), used in maxicell studies, contains the 3.9-kb *Hin*dIII fragment (Fig. 1) encompassing the *hrp* locus cloned into a polylinker site. An internal *Bgl*II site is located 2.1 and 1.8 kb away, respectively, from the *Hin*dIII sites at the right and left ends of the 3.9-kb fragment (Fig. 1B). A DNA fragment in pOSU3126 that extends from the *Bam*HI site in the polylinker to the internal *Bgl*II site 2.1 kb downstream was subcloned in both orientations at the *Bam*HI site of M13mp18. Similarly, a 1.8-kb *Bam*HI-*Bgl*II fragment that extends from the *Bam*HI site in the polylinker through the rest of the 3.9-kb fragment was subcloned from pOSU3125 (Table 1) into M13mp18 in both orientations. A 2.7-kb *Sal*I fragment (Fig. 1B) within the 3.9-kb fragment was also cloned in both orientations in M13mp18 to verify the DNA sequences around the *Bgl*II site that was used as an endpoint in previous clones. Unidirectional deletions of the 2.1-, 1.8-, and 2.7-kb fragments were generated from either end of each fragment by the exonuclease III deletion method (see Materials and Methods). Sequential overlapping clones with endpoints separated by ca. 150 base pairs (bp) were used to perform the DNA sequencing of each strand of the fragments. The DNA sequence of the strand containing the promoter activity (sense strand) and the deduced amino acid sequences are shown in Fig. 2.

The sense strand contains two open reading frames (ORFs), ORF1 and ORF2, that code for putative polypeptides of 40 and 83 kDa, respectively (Fig. 1E). These data are in good agreement with the sizes of the polypeptides (ca. 37 and ca. 85 kDa) expressed in *E. coli* maxicells in earlier studies (25). An *E. coli* consensus promoter sequence (34) is present upstream of both ORFs (Fig. 2), where promoter activity was detected. However, the −10 region of this promoter overlaps with the first translational initiation codon, ATG, of ORF1, and no promoter activity was detected in the 0.7-kb *Hin*dIII fragment present immediately upstream. Initiation of translation at the following in-frame ATG of ORF1 present at nucleotide 405 (Fig. 2) will yield a putative 28-kDa polypeptide (see Discussion). The translational initiation codon ATG of ORF2 is located 205 nucleotides downstream of the translational stop codon of ORF1. This ATG is flanked by a 9-bp inverted repeat, and a consensus ribosome-binding site (RBS), GGAGGA (37), is located immediately upstream (Fig. 2). Two DNA sequences identical to 7 and 6 bp of the left repeat sequence are located further upstream between ORF1 and ORF2. A segment of the DNA sequence located 175 nucleotides downstream of the translational termination codon of ORF2 contains all the features of a transcriptional terminator. Specifically, there is an 11-bp inverted repeat that can form a putative stem by base pairing of the mRNA, leaving a 5-base loop with six T nucleotides immediately downstream of the stem-and-loop structure. Hence, the DNA sequence of the 3.9-kb fragment contains the features of a polycistronic operon.

Features of the putative polypeptides encoded by the ORFs. To determine whether the polypeptides encoded by the ORFs share any features of the proteins with known functions, we performed a computer search for amino acid sequence homology with proteins in the GenBank and European Molecular Biology Laboratory data bases. The amino acid sequence deduced from ORF2 did not show homology with any known protein. To further investigate

FIG. 2. (*See following page.*) Nucleotide sequence of the 3,854-bp *Hin*dIII fragment encompassing the *hrp* locus and the deduced amino acid sequences of putative polypeptides encoded on the sense strand. The nucleotides are numbered beginning at the *Hin*dIII site upstream of ORF1. The putative translational initiation site for each ORF is indicated by a solid arrow above the sequence, whereas both the start and stop sites are shown in capital letters in the amino acid sequences. A consensus sequence for a putative RBS upstream of ORF2 is indicated by smaller arrows above the sequence. Conserved regions of a consensus promoter sequence upstream of ORF1 are underlined. A 9-bp inverted repeat sequence flanking the putative initiation codon of ORF2 is indicated by arrows pointing inward. Part of this repeat (GCCGAG), which is reiterated in direct order upstream of ORF2, is identified by a shorter arrow below the sequence. Dotted lines below the sequence downstream of ORF2 indicate the stem of a putative stem-and-loop structure that may act as a transcriptional terminator. Vertical arrows above the sequence indicate the extent of ORF2 used in protein fusion experiments.

```
                                orf1 →
AAG CTT GGC GTT GCT CCT CTG ACC AGT ATG TTC CTG TTC GGC GCC AAC CAG CCT TCG CGT GTG CCT AAC TAC CGT CGT GAA CTG CAC GAT 90
                                    MET Phe Leu Phe Gly Ala Asn Gln Pro Ser Arg Val Pro Asn Tyr Arg Arg Glu Leu His Asp

TCC AGC GGT CTG TCG ATT CAG GCG GCC AAC GGT GAG TGG CTG TGG CGT CCG CTG AAC AAC CCT AAA CAT CTG TCC ATC AGC AGC TTC TCG 180
Ser Ser Gly Leu Ser Ile Gln Ala Ala Asn Gly Glu Trp Leu Trp Arg Pro Leu Asn Asn Pro Lys His Leu Ser Ile Ser Ser Phe Ser

GTC GAG AAC CCG CGT GGT TTC GGT CTG CTG CAA CGT GGC CGC GAC TTC AGC CAG TAC GAA GAC CTG GAT GAC CGC TAC GAC AAG CGT CCA 270
Val Glu Asn Pro Arg Gly Phe Gly Leu Leu Gln Arg Gly Arg Asp Phe Ser Gln Tyr Glu Asp Leu Asp Asp Arg Tyr Asp Lys Arg Pro

AGT GCC TGG ATC GAG CCG AAG GGC GAT TGG GGT AAA GGG ACT GTC GAG CTG GTC GAA ATT CCG ACT GCC GAC GAG ACC AAC GAC AAC ATC 360
Ser Ala Trp Ile Glu Pro Lys Gly Asp Trp Gly Lys Gly Thr Val Glu Leu Val Glu Ile Pro Thr Ala Asp Glu Thr Asn Asp Asn Ile

GTA GCT TAC TGG AAG CCT GAA ACG CTG GCC GAG CCT GGT CAG GAA ATG GCG TTC GAC TAC CGT CTG CAC TGG ACC ATG CAG GAA AAC TCG 450
Val Ala Tyr Trp Lys Pro Glu Thr Leu Ala Glu Pro Gly Gln Glu MET Ala Phe Asp Tyr Arg Leu His Trp Thr MET Gln Glu Asn Ser

ATT CAC TCG CCG GAT CTG GGC TGG GTC AAG CAG ACT CAA CGC TCC ATC GGT GAC GTG CGT CAG TCC AAC CTG ATC CGT CAG CCG GAC GGC 540
Ile His Ser Pro Asp Leu Gly Trp Val Lys Gln Thr Gln Arg Ser Ile Gly Asp Val Arg Gln Ser Asn Leu Ile Arg Gln Pro Asp Gly

AGC CTT GCC TTC CTG GTC GAC TTC GTG GGC CCG GTG CTG GCC GCA CTG CCG GAA GAC AAG ACC ATT CGC AGC CAG GTG ACC ACT GAC GAC 630
Ser Leu Ala Phe Leu Val Asp Phe Val Gly Pro Val Leu Ala Ala Leu Pro Glu Asp Lys Thr Ile Arg Ser Gln Val Thr Thr Asp Asp

AAC GTC GAG CTG GTG GAA AAC AAC CTG CGC TAC AAC CCG GTC ACC AAA GGT TAC CGC CTG ACC CTG CGT GTC AAG GTC AAG GAT TCC AGC 720
Asn Val Glu Leu Val Glu Asn Asn Leu Arg Tyr Asn Pro Val Thr Lys Gly Tyr Arg Leu Thr Leu Arg Val Lys Val Lys Asp Ser Ser

AAG CCG ACC GAA ATG CGC GCC TAC CTG TTG CGT GAA ATC CCT GCC GAA CCG GGC AAG GAA CCT GCG CTG CTC GTG GCT GAC AAA GCC GAA 810
Lys Pro Thr Glu MET Arg Ala Tyr Leu Leu Arg Glu Ile Pro Ala Glu Pro Gly Lys Glu Pro Ala Leu Leu Val Ala Asp Lys Ala Glu

GAG AAG AAG GCT GCC GCG AAG GAA GCT GCC AAG CCG GCA GTC TCC AAG GAG TCC GCC AAC GAC CAG GTA GAA ATC GCC AAG GCC GAC GCA 900
Glu Lys Lys Ala Ala Ala Lys Glu Ala Ala Lys Pro Ala Val Ser Lys Glu Ser Ala Asn Asp Gln Val Glu Ile Ala Lys Ala Asp Ala

CCC AAG CCG GAA GCT GCC AAG CCT GAG ACT GCC AAG TCC GAA GCT GGC AAG GCT GAC GCA GCC AAA GGC AAA GGC GAA GTC GCC AAG GCC 990
Pro Lys Pro Glu Ala Ala Lys Pro Glu Thr Ala Lys Ser Glu Ala Gly Lys Ala Asp Ala Ala Lys Gly Lys Gly Glu Val Ala Lys Ala

GAT GCA GGC AAA GCC GAC GCA TCC AAG GCT GAA GCA GCC AAG GAT AAG GAC GGT AAG GAA ATT CAG CAG CCT GAA ACC GAG GCA GCA CCC 1080
Asp Ala Gly Lys Ala Asp Ala Ser Lys Ala Glu Ala Ala Lys Asp Lys Asp Gly Lys Glu Ile Gln Gln Pro Glu Thr Glu Ala Ala Pro

ACC CAT CCG GAA CCG GCC AAG ACG TTG CAA GTC ATG ACC GAG ACC TGG AGC TAT CAG TTG CCG AGC GAT GAG TAA TTC TCT ACC GGT GCC 1170
Thr His Pro Glu Pro Ala Lys Thr Leu Gln Val MET Thr Glu Thr Trp Ser Tyr Gln Leu Pro Ser Asp Glu END

AATGTCTCTGAACGAGTACCTGGCGCATTTACCGATGAGCGACGAGCAGCGGGCAGAACTTGCCGGCTGCACGACCTT  CGCCGAG TTGCATTGAGCGACTGTCCGCGCAGCCGGTC
ACTGACCC TGCCGAG GCCGCTCAGGCTTCGGTGGGTCGCCGTCTG                                                                    1330
                              ←rbs→    orf2 →
ACG TGA CCA CGA GAT CAG CTG GAG GAC GCC GAG ATG CTC GGC GTC GAT GCC AGC GGT CGC CTG TGC CTG AAG GCT ACA CCA CCG ATT CGC 1420
                                    MET Leu Gly Val Asp Ala Ser Gly Arg Leu Cys Leu Lys Ala Thr Pro Pro Ile Arg

CGG ACC AAG GTC GTG CCA GAG CCA TGG CGC ACC AAC ATC CTG GTG CGC GGC TGG CGT CGC CTG ACC GGC AAG GGC AAC CCG CCC AAG CCC 1510
Arg Thr Lys Val Val Pro Glu Pro Trp Arg Thr Asn Ile Leu Val Arg Gly Trp Arg Arg Leu Thr Gly Lys Gly Asn Pro Pro Lys Pro

GAG CAC GAT GAT CTG CCG CGG GAT CTG CCG AAG GCG CGC TGG CGT ACC GTC GGT TCG ATC CGT CGC TAC ATC CTG CTG ATC CTC ATG CTG 1600
Glu His Asp Asp Leu Pro Arg Asp Leu Pro Lys Ala Arg Trp Arg Thr Val Gly Ser Ile Arg Arg Tyr Ile Leu Leu Ile Leu MET Leu

GGT CAG ACG ATC GTG GCT GGC TGG TAC ATG AAA GGC ATT CTG CCG TAT CAG GGC TGG TCG CTG GTT TCG CTC GAC GAA ATC ACC CGT CAG 1690
Gly Gln Thr Ile Val Ala Gly Trp Tyr MET Lys Gly Ile Leu Pro Tyr Gln Gly Trp Ser Leu Val Ser Leu Asp Glu Ile Thr Arg Gln

ACC TTT GTG CAG ACC GCC TTG CAG GTC ATG CCT TAT GCC TTG CAG ACC AGT ATT CTG TTG CTG TTC GGG ATT CTG TTC TGC TGG GTA TCG 1780
Thr Phe Val Gln Thr Ala Leu Gln Val MET Pro Tyr Ala Leu Gln Thr Ser Ile Leu Leu Leu Phe Gly Ile Leu Phe Cys Trp Val Ser

GCC GGT TTC TGG ACC GCG CTG ATG GGC TTC CTG GAA TTG CTC ACC GGT CGC GAC AAA TAC CGC ATC TCG GGT GCC AGT GCC GGC AAC GAG 1870
Ala Gly Phe Trp Thr Ala Leu MET Gly Phe Leu Glu Leu Leu Thr Gly Arg Asp Lys Tyr Arg Ile Ser Gly Ala Ser Ala Gly Asn Glu
```

```
CCG ATC GAA AAG GGC GCA CGT ACT GCG CTG GTC ATG CCG ATC TGC AAC GAA GAC GTG CCT CGG GTT TTC GCC GGT CTG CGC GCT ACG TTC 1960
Pro Ile Glu Lys Gly Ala Arg Thr Ala Leu Val MET Pro Ile Cys Asn Glu Asp Val Pro Arg Val Phe Ala Gly Leu Arg Ala Thr Phe

GAA TCG GTA GCG GCC ACG GGT GAC CTG GAT CGT TTC GAT TTC TTC GTG CTC AGT GAC ACC AAC GAA ACC GAC ATC GCC GTT GCC GAG CAA 2050
Glu Ser Val Ala Ala Thr Gly Asp Leu Asp Arg Phe Asp Phe Phe Val Leu Ser Asp Thr Asn Glu Thr Asp Ile Ala Val Ala Glu Gln

CAG GCG TGG CTG GAC GTG TGC CGC GAG ACC AAA GGC TTC GGC AAG ATC TTC TAC CGT CGC CGT CGC CGT CGC GTA AAA CGC AAA AGC GGC 2140
Gln Ala Trp Leu Asp Val Cys Arg Glu Thr Lys Gly Phe Gly Lys Ile Phe Tyr Arg Arg Arg Arg Arg Arg Val Lys Arg Lys Ser Gly

AAC CTC GAC GAC TTC TGC CGG CGC TGG GGC GGT GAC TAC CGC TAC ATG GTC GTG CTG GAC GCC GAC AGC GTC ATG AGC GGT GAG TGT CTG 2230
Asn Leu Asp Asp Phe Cys Arg Arg Trp Gly Gly Asp Tyr Arg Tyr MET Val Val Leu Asp Ala Asp Ser Val MET Ser Gly Glu Cys Leu

ACC AGT CTG GTT CGC CTG ATG GAA GCC ACG CCG GAC GCC GGT ATC ATC CAG ACC GCG CCA CGT GCG TCG GGC ATG GAC ACG CTG TAT GCA 2320
Thr Ser Leu Val Arg Leu MET Glu Ala Thr Pro Asp Ala Gly Ile Ile Gln Thr Ala Pro Arg Ala Ser Gly MET Asp Thr Leu Tyr Ala

CGC ATG CAG CAG TTC GCC ACC CGG GTC TAT GGT CCG CTG TTC ACT GCC GGT CTG CAC TTC TGG CAG CTG GGT GAA TCC CAC TAT TGG GGG 2410
Arg MET Gln Gln Phe Ala Thr Arg Val Tyr Gly Pro Leu Phe Thr Ala Gly Leu His Phe Trp Gln Leu Gly Glu Ser His Tyr Trp Gly

CAC AAC GCG ATC ATC CGC ATG AAG CCC TTC ATC GAG CAC TGC GCC CTG GCG CCG CTG CCC GGC AAA GGC GCA TTC GCC GGT GCG ATC CTC 2500
His Asn Ala Ile Ile Arg MET Lys Pro Phe Ile Glu His Cys Ala Leu Ala Pro Leu Pro Gly Lys Gly Ala Phe Ala Gly Ala Ile Leu

TCC CAC GAC TTC GTC GAA GCT GCG CTG ATG CGC CGT GCC GGC TGG GGC GTG TGG ATT GCC TAC GAC CTG CCA GGC AGT TAC GAA GAG TTG 2590
Ser His Asp Phe Val Glu Ala Ala Leu MET Arg Arg Ala Gly Trp Gly Val Trp Ile Ala Tyr Asp Leu Pro Gly Ser Tyr Glu Glu Leu

CCG CCT AAC CTG CTG GAC GAA CTC AAG CGT GAC CGT CGC TGG TGC CAC GGC AAC CTG ATG AAC TTC AGG CTG TTC CTG GTC AAG GGC ATG 2680
Pro Pro Asn Leu Leu Asp Glu Leu Lys Arg Asp Arg Arg Trp Cys His Gly Asn Leu MET Asn Phe Arg Leu Phe Leu Val Lys Gly MET

CAC CCG GTT CAC CGT GCG GTG TTC CTG ACC GGT GTG ATG TCT TAC CTG TCG GCA CCG TTG TGG TTC TTC TTC CTC GTG CTG TCC ACG GCT 2770
His Pro Val His Arg Ala Val Phe Leu Thr Gly Val MET Ser Tyr Leu Ser Ala Pro Leu Trp Phe Phe Phe Leu Val Leu Ser Thr Ala

TTG CTG GCG GTG AAC ACG CTG ATG GAG CCG ACC TAC TTC CTT GAA CCG CGT CAG CTG TAC CCG CTG TGG CCA CAA TGG CAC CCG GAA AAA 2860
Leu Leu Ala Val Asn Thr Leu MET Glu Pro Thr Tyr Phe Leu Glu Pro Arg Gln Leu Tyr Pro Leu Trp Pro Gln Trp His Pro Glu Lys

GCC GTT GCG TTG TTC TCG ACC ACC ATC GTC CTG CTG TTC CTG CCT AAA CTG CTC AGC GTC ATT CTG ATC TGG GCC AAG GGC GCG AAA GGC 2950
Ala Val Ala Leu Phe Ser Thr Thr Ile Val Leu Leu Phe Leu Pro Lys Leu Leu Ser Val Ile Leu Ile Trp Ala Lys Gly Ala Lys Gly

TTC GGT GGC AAG TTC AAG GTC ACC GTT TCG ATG CTG CTG GAA ATG CTC TTC TCG GTG CTG CTG GCT CCG GTG CGC ATG CTG TTC CAC ACA 3040
Phe Gly Gly Lys Phe Lys Val Thr Val Ser MET Leu Leu Glu MET Leu Phe Ser Val Leu Leu Ala Pro Val Arg MET Leu Phe His Thr

CGC TTC GTA CTG GCC GCT TTC CTG GGC TGG GCC GCG ACC TGG AAC TCG CCG CAG CGC GAC GAT GAT TCC ACG CCG TGG ATC GAA GCG GTG 3130
Arg Phe Val Leu Ala Ala Phe Leu Gly Trp Ala Ala Thr Trp Asn Ser Pro Gln Arg Asp Asp Asp Ser Thr Pro Trp Ile Glu Ala Val

AAG CGT CAT GGT CCG CAA ACC CTG CTG GGC GCG TGC TGG GCC TTG CTG GTG TTC TGG TTG AAC CCG AGC TTC CTG TGG TGG CTT GCG CCG 3220
Lys Arg His Gly Pro Gln Thr Leu Leu Gly Ala Cys Trp Ala Leu Leu Val Phe Trp Leu Asn Pro Ser Phe Leu Trp Trp Leu Ala Pro

ATC GTG GTG TCG TTG ATG CTG TCG ATT CCG GTG TCG GTG ATT TCC AGC CGT ACC AAT CTG GGC GTC AAG GCG CGT GAC GAG AAG TTC TTC 3310
Ile Val Val Ser Leu MET Leu Ser Ile Pro Val Ser Val Ile Ser Ser Arg Thr Asn Leu Gly Val Lys Ala Arg Asp Glu Lys Phe Phe

CTG ATT CCT GAA GAG TTC GAG CCG CCG CAA GAG CTG ATC TCG ACG GAT CGG TAC ACC TAC GAG AAC CGC TGG CAT GCG CTG AAG CAG GGC 3400
Leu Ile Pro Glu Glu Phe Glu Pro Pro Gln Glu Leu Ile Ser Thr Asp Arg Tyr Thr Tyr Glu Asn Arg Trp His Ala Leu Lys Gln Gly

TTC ATC CGC GCT GTG GTC GAC CCG CGC CAG AAC GCC CTG GCC TGC GCC CTG GCG ACG TCG CGT CAC GTC AGG CTC AGC CGA TTG AAG TGG 3490
Phe Ile Arg Ala Val Val Asp Pro Arg Gln Asn Ala Leu Ala Cys Ala Leu Ala Thr Ser Arg His Val Arg Leu Ser Arg Leu Lys Trp

TGC GTA TGG AGC GTG TCG ATC AGG CAC TCA AGG TCG GTC CGG CAA AAC TCG GCA ATC AGG AAC GCC TGA TGC TGC TGA GCG ACC CGG TCG 3580
Cys Val Trp Ser Val Ser Ile Arg His Ser Arg Ser Val Arg Gln Asn Ser Ala Ile Arg Asn Ala END

CCCTTGGCCGCTTGCACGAGCGCGTCTGGAGCGAAGGTCACGAAGAGTGGCTGGCCGCGTGGAGAGCTTCCATCGAAGCCGATCCACATGCGCCTCTGCTGCCTTTGCAGCCTGAAGGTA 3700
AAGCATCGGAGCCGGTTCCGGTCTAAAACCGCCCGGCTACGAAAAGA GCCCCTGACGC T T A G C GCGTCGGGGGC TTTTTT ATTGGGCTTTGTGGCAAAGACCCACGCAAG 3810
CGAGGGGGAGAGGCGAACTTGTTCGCGGCATGTTCGGAAAGCTT 3854
```

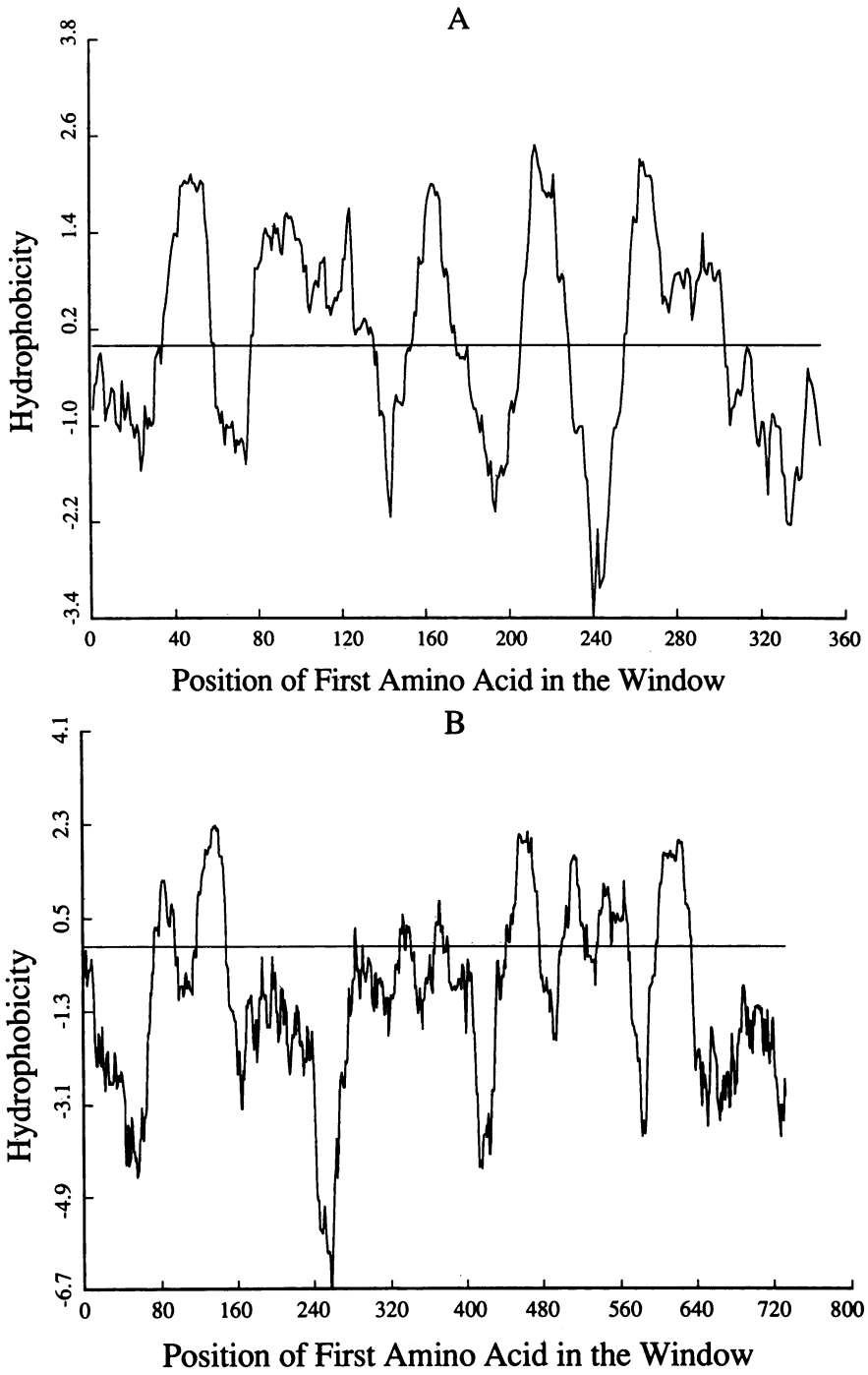

FIG. 3. Comparison of hydrophobicity profiles of the putative 83-kDa polypeptide and another integral membrane protein, rhodopsin. The plots were generated by using a window of 19 amino acids to identify nonpolar transbilayer helices in the amino acid sequences (9). Such putative membrane-spanning domains are indicated by peaks rising above the base line of the plot. Positive values on the ordinate represent hydrophobic regions, whereas negative values indicate hydrophilic regions of the protein. (A) Bovine rhodopsin. (B) 83-kDa polypeptide.

any feature of the 83-kDa protein shared with known proteins, we plotted the distribution of its hydrophilic and hydrophobic amino acids by using several different computer programs (17). The hydrophobicity plot, which depicts membrane-spanning helices of a protein (10, 11), revealed an interesting feature of this protein. The hydrophobicity plots of the 83-kDa polypeptide and a transmembrane protein, rhodopsin (29), are presented in Fig. 3. The existence of nonpolar transbilayer helices in the amino acid sequence of rhodopsin is shown by hydrophobic peaks (Fig. 3A). The presence of similar membrane-spanning helices in the plot of the 83-kDa polypeptide suggested that it might be a transmembrane protein (Fig. 3B). This plot also indicated that the most hydrophilic region, consisting of six arginine residues

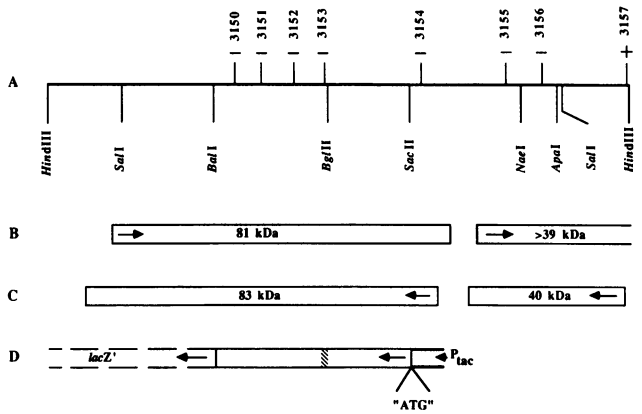

FIG. 4. Fine-structure map of the Tn*5* insertions in the chromosome of *P. syringae* pv. *syringae* R32 with respect to the positions of the ORFs on both sense and antisense strands of the *hrp* locus. The arrows indicate the directions of translation. (A) Fine-structure map of the 3.9-kb *Hin*dIII fragment, with the restriction sites indicated below. (B) Positions of ORFs on the antisense strand with respect to the sites of Tn*5* insertions and consequent phenotypes. The sizes of the putative polypeptides encoded by these ORFs are indicated within the boxes. (C) Same as panel B, but ORFs are on the sense strand. (D) Scheme for construction of a *hrpM'-lacZ'* fusion gene. The sites of the synthetic linkers ("ATG") and *lacZ'* gene (---) are indicated. Symbol: , the strongest antigenic determinant in *hrpM* predicted by computer analysis (see text). The position and direction of the *tac* promoter are shown by P_{tac} and an arrowhead, respectively.

located 2,104 nucleotides downstream of the left *Hin*dIII site, was apparently the strongest antigenic determinant of the polypeptide, as projected by Hopp and Woods analysis (14) of various antigenic determinants in known proteins.

An intriguing feature of this *hrp* locus is the presence of two ORFs on the antisense strand that encode putative polypeptides of 81 and >39 kDa (Fig. 4B). The second ORF on this strand extends beyond the *Hin*dIII site at the right end of the *hrp* locus, and therefore its putative size was not determined. These ORFs consist mostly of codons complementary to those present in the respective ORFs in the sense strand. The genetic data suggested that the putative products of these ORFs, if expressed, were not involved in pathogenesis. Tn*5* insertions designated 3155 and 3156 (Fig. 4A), which resulted in loss of pathogenicity, should not affect expression of the putative 81-kDa protein, and insertion 3157 in the following ORF (>39 kDa) did not affect pathogenicity. Moreover, expression of any polypeptide from these ORFs has not been established in maxicell studies (25); no detectable promoter activity is present upstream; and examination of codon usage in *E. coli* (1) reveals that products encoded by the antisense strand, unlike those encoded by the sense strand, will be poorly expressed (data not shown). Hence, the genes encoding the 83-kDa polypeptide, designated *hrpM*, and possibly the 40-kDa polypeptide appeared to be the major determinants of this pathogenicity locus.

Construction of a chimeric gene from the *hrpM* locus. Since the original mutant PS9021 had Tn*5* inserted in *hrpM*, we attempted to overexpress the gene in *E. coli* by cloning the entire locus downstream of the temperature-inducible λ p_L promoter and the IPTG-inducible *tac* promoter in expression vectors pCP3 (33) and pKK223-3 (Pharmacia, Inc.), respectively. No overexpression was observed in Coomassie blue-stained gels under induced conditions. To circumvent any problem associated with the expression of the wild-type 83-kDa protein in *E. coli*, we constructed a chimeric gene by

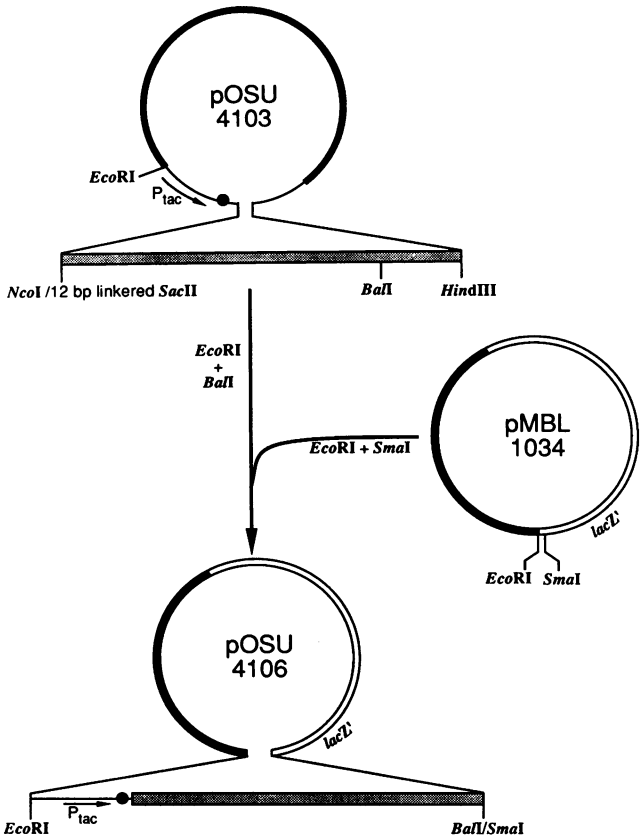

FIG. 5. Plasmid constructions for synthesis of a *hrpM'-lacZ'* fusion protein. Constructs generated for the 12-bp *Nco*I linker are shown. An identical strategy was used to generate fusions with 8- and 10-bp *Nco*I linkers. Symbols: ——, pKK233-2 sequences; ▬, pBR322 sequences; ═══, *lacZ'* and *lacY'* sequences of *E. coli*; ▦, *P. syringae* pv. *syringae* DNA (Fig. 4D). Positions of the *tac* promoter and a consensus RBS (●) upstream of the *hrpM'* sequence are shown. Restriction sites present in the sequences are indicated below the sequences, and those used in the fusion are shown on appropriate sides of the vertical arrow. Fusions of two restriction sites are shown as slashes.

fusing the *Sac*II-*Bal*I fragment (Fig. 4D) of the *hrpM* gene to codon 8 of the *lacZ* gene in the appropriate reading frame.

Initially, a pBR322 derivative plasmid, pOSU3125, containing the 3.9-kb *Hin*dIII fragment (Fig. 4A) was digested with *Sac*II, and the ends were filled in with T4 DNA polymerase. Synthetic linkers carrying the translational initiation codon ATG as part of a *Nco*I site (Pharmacia, Inc.) were then ligated to the blunt-ended *Sac*II site to introduce translational start codons. These linkers, of 8, 10, and 12 bp, created ATG start sites in the three possible reading frames at the *Sac*II site of the *hrpM* gene. The *Nco*I-*Hin*dIII digests of these clones were then ligated to plasmid pKK233-2 (Pharmacia), which was digested similarly. The resulting clones were screened for inserts that contain the *Bal*I site of the 3.9-kb fragment, and clones which contained the 8-, 10-, and 12-bp linkers were designated pOSU4101, pOSU4102, and pOSU4103, respectively. The *Nco*I site in plasmid pKK233-2 is situated at an optimum distance downstream of a consensus RBS, as well as the inducible *tac* promoter of *E. coli* (Fig. 5). Hence, the three clones contained translational start sites that should express truncated proteins upon induction from the three respective reading frames of the *hrpM* gene. However, if fusions were performed properly

and the DNA sequencing data were correct, the ATG in the 12-bp linker was the only initiation codon predicted to be properly aligned with *hrpM'* and *lacZ'*. An *Eco*RI site upstream of the *tac* promoter in pKK233-2 and a *Bal*I site near the 3' end of *hrpM* (Fig. 5) were used to release an *Eco*RI-*Bal*I fragment from each of the three clones that contained all the features of the 5' end of a gene, namely, the promoter, the RBS, and the truncated ORF beginning with an ATG codon. Each of these fragments was separately ligated to the *Eco*RI-*Sma*I digest of pMLB1034 (Table 1; Fig. 5), and the resulting plasmids were designated pOSU4104, pOSU4105, and pOSU4106 (Fig. 5). The ligation of *Bal*I and *Sma*I ends created a fusion of the 3' end of the truncated *hrpM* gene to codon 8 of the *lacZ* gene present in the plasmid pMLB1034.

Expression of the fusion protein encoded by the *hrpM'-lacZ'* hybrid gene in *E. coli*. Plasmids pOSU4104, pOSU4105, and pOSU4106 were transformed into *E. coli* JM105. Since JM105 contains an F' episome containing a repressor for the *tac* promoter, induction of the *tac* promoter in this strain can be achieved by addition of IPTG. Strains JM105(pOSU 4104), JM105(pOSU4105) and JM105(pOSU4106) were grown to mid-log phase and induced with 2 mM IPTG for 1 h. Equal amounts of cell lysates (see Materials and Methods) from all these strains were subjected to Western immunoblot analysis with antibody raised against β-galactosidase. A putative polypeptide of 163 kDa was predicted to be encoded by the hybrid gene, and a protein of comparable size was detected only in the lysates of JM105(pOSU4106) (Fig. 6). Since the *lacZ'* gene was fused to the appropriate reading frame of *hrpM'* in pOSU4106, expression of the fusion protein in JM105(pOSU4106) lysates provided further verification of the DNA sequencing data. Expression of a truncated β-galactosidase protein (with 32 amino acids deleted) from the *lacZΔM15* gene (19) present on the F' episome in JM105 provided a convenient marker protein of 112 kDa, which was detected in each of the lysates in Western blot analysis. The amount of hybrid protein obtained in this induction assay was estimated to be 0.25% of the total protein present in the cell lysates (data not shown). At present, the fusion protein is being purified in preparative amounts for use in raising antibodies against the 83-kDa polypeptide.

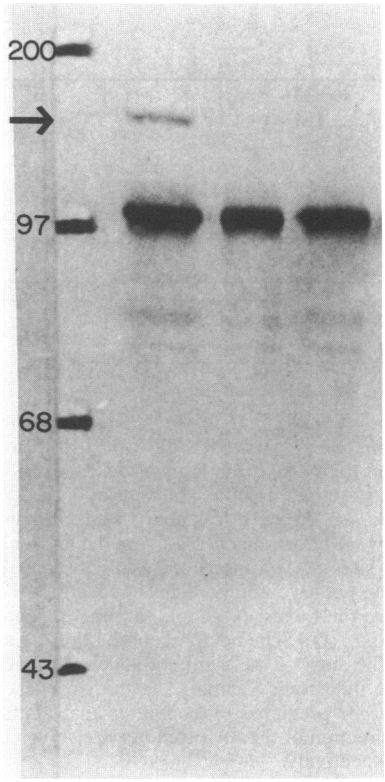

FIG. 6. Expression of the *hrpM'-lacZ'* fusion product in *E. coli*. Immunoblot analysis of postinduction cell lysates of *E. coli* JM105 containing the fusion gene in different reading frames was performed with antibodies to β-galactosidase (see text). Left lane, lysates of JM105(pOSU4106); middle lane, lysates of JM105(pOSU4105); right lane, lysates of JM105(pOSU4104). The position of the hybrid protein is indicated by the arrow. Numbers on the left represent molecular masses (in kilodaltons) of protein standards, whose positions in the blot are also shown.

DISCUSSION

A pathogenicity locus of *P. syringae* pv. *syringae* PS9020 was shown to be contained within a 3.9-kb *Hin*dIII fragment. When the DNA sequence of this region was determined and analyzed, two ORFs appeared to be important for the Hrp^+ phenotype of PS9020 on bean plants. The Tn*5* insertions in ORF2 resulted in a loss of the pathogenicity, indicating that the 83-kDa protein is a pathogenicity factor. Although the Tn*5* insertion designated 3157 appeared to be within the putative promoter sequence on the basis of a fine-structure restriction map of the Tn*5* insertions, this insertion did not affect the Hrp^+ phenotype of PS9020. The reason for such a phenotype is not clear at present, but it is possible that a promoter in Tn*5* is being used to express the downstream ORFs under these conditions (3). Although Tn*5* insertions were not obtained downstream of *hrpM* to reveal the extent of this pathogenicity locus in that direction, a DNA sequence 88 bp downstream of *hrpM* was characteristic of a transcription termination site, suggesting that *hrpM* was the last gene in any possible operon structure. The common features of bacterial termination sequences (35) include an inverted repeat, a G+C-rich sequence of 3 to 11 contiguous bases, and a run of 4 to 8 U residues following the G+C-rich region of the mRNA. A DNA sequence containing all these features was detected 175 bp downstream of the *hrpM* gene (Fig. 2). Moreover, in maxicell studies, expression of the *galK* gene was severely reduced when an insert containing this sequence was cloned upstream in the proper orientation (25). Thus, any other gene downstream of *hrpM* would probably belong to a separate genetic unit. Another interesting feature of the 83-kDa polypeptide is a 9-bp inverted repeat that flanks the putative ATG start codon. The left repeat separated the putative RBS from ATG (Fig. 2), suggesting that regulation of *hrpM* expression may occur at the translational level (43).

The role of the ORF1-encoded product in pathogenicity is as yet undetermined. The Tn*5* insertions in ORF1 led to mutants with the same phenotypes (nonpathogenic and mucoid) as those resulting from insertions within the *hrpM* gene downstream. These results suggested that the expression of the *hrpM* product was inhibited by Tn*5* insertions in the upstream ORF. Such inhibition might be the result of the polar effect of Tn*5* insertions in distal genes (3), since no separate promoter was detected upstream of *hrpM*. Alternatively, these Tn*5* insertions could have inactivated the ORF1 product, which might act as a positive regulator of *hrpM* gene expression. The first possibility does not necessarily involve the ORF1 product in pathogenicity, whereas the

second one implicates ORF1 as a regulatory gene of the *hrp* locus. A third possibility is that ORF1 is an independent *hrp* locus. Interestingly, computer analysis of amino acid homology with known proteins in the data bases revealed that the putative 40-kDa protein encoded by ORF1 exhibited 30 to 35% homology over a range of 80 to 130 amino acids with various DNA-binding proteins (data not shown), such as the regulatory histones H1 and H5 (6). Since proteins involved in the regulation of gene expression are DNA-binding proteins in many cases, it is possible that the ORF1-encoded product is involved in regulation of *hrpM* or other *hrp* loci.

Some ambiguity exists in determining the size of the wild-type protein encoded by ORF1. A protein estimated to be 37 kDa was detected in maxicell studies when an external promoter from the *lac* operon of *E. coli* was present upstream (25). In such a case, translation could begin at the first ATG codon of ORF1. However, if the consensus promoter sequence detected upstream of ORF1 is the promoter transcribing both ORF1 and *hrpM* in the chromosome of *P. syringae* pv. *syringae* R32, then the first ATG of ORF1, at nucleotide 28 (Fig. 2), is precluded as a translational start site, since it is part of the −10 region of this promoter. The next ATG codon, located 372 bases downstream, could initiate synthesis of a 27-kDa protein. A protein of that size could have gone undetected in *E. coli* maxicells, since it would be obscured in the sodium dodecyl sulfate-polyacrylamide gel by a comigrating *bla* gene product (40). Purification of the ORF1-encoded product and physical mapping of the transcriptional unit will reveal the size of the ORF1 product and the location of the native promoter for this *hrp* locus.

ACKNOWLEDGMENTS

We thank William Fickett for excellent technical assistance and the members of this laboratory for critical discussion of this work.

This research was supported by U.S. Department of Agriculture Science and Education Administration grant 85-CRCR-1-1771 from the Competitive Research Grants Office and by the Oregon Agricultural Experiment Station.

LITERATURE CITED

1. **Allf-Steinberger, C.** 1984. Evidence of coding pattern on the noncoding strand of the *E. coli* genome. Nucleic Acids Res. **12:** 2235–2241.
2. **Anderson, D. M., and D. Mills.** 1985. The use of transposon mutagenesis in the isolation of nutritional and virulence mutants in two pathovars of *Pseudomonas syringae*. Phytopathology **75:** 104–108.
3. **Berg, D., A. Weiss, and L. Grassland.** 1980. Polarity of Tn*5* insertion mutations in *Escherichia coli*. J. Bacteriol. **142:**439–446.
4. **Bertoni, G., and D. Mills.** 1987. A simple method to monitor growth of bacterial populations in leaf tissue. Phytopathology **77:**832–835.
5. **Boucher, C. A., F. Van Gijsegem, P. A. Barberis, M. Arlat, and C. Zischek.** 1987. *Pseudomonas solanacearum* genes controlling both pathogenicity on tomato and hypersensitivity on tobacco are clustered. J. Bacteriol. **169:**5626–5632.
6. **Coles, L. S., and J. R. E. Wells.** 1985. An H1 histone gene-specific 5′ element and evolution of H1 and H5 genes. Nucleic Acids Res. **13:**585–594.
7. **Conrad, B., and D. W. Mount.** 1982. Microcomputer programs for DNA sequence analysis. Nucleic Acids Res. **10:**31–38.
8. **Cuppels, D. A.** 1986. Generation and characterization of Tn*5* insertion mutations in *Pseudomonas syringae* pv. *tomato*. Appl. Environ. Microbiol. **51:**323–327.
9. **Deininger, C. A., G. M. Mueller, and P. K. Wolber.** 1988. Immunological characterization of ice nucleation proteins from *Pseudomonas syringae*, *Pseudomonas fluorescens*, and *Erwinia herbicola*. J. Bacteriol. **170:**669–675.
10. **Engelman, D. M., and T. A. Steitz.** 1984. On the folding and insertion of globular membrane proteins, p. 87–114. *In* D. B. Wetlaufer (ed.), The protein-folding problem. Westview Press, Boulder, Colo.
11. **Engelman, D. M., T. A. Steitz, and A. Goldman.** 1986. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. Annu. Rev. Biophys. Biophys. Chem. **15:** 321–353.
12. **Fahy, P. C., and A. B. Lloyd.** 1983. Pseudomonas: the fluorescent pseudomonads, p. 141–188. *In* P. C. Fahy and G. J. Persley (ed.), Plant bacterial diseases: a diagnostic guide. Academic Press, Ltd., Sydney, Australia.
13. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28:**351–359.
14. **Hopp, T. P., and K. R. Woods.** 1981. Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. USA **78:**3824–3828.
15. **Kiraly, Z.** 1980. Defenses triggered by invaders: hypersensitivity, p. 201–224. *In* J. Horsfall and E. Cowling (ed.), Plant disease, vol. 5. Academic Press, Inc. (London), Ltd., London.
16. **Klement, Z.** 1982. Hypersensitivity, p. 149–177. *In* M. S. Mount and G. H. Lacy (ed.), Phytopathogenic procaryotes, vol. 2. Academic Press, Inc., New York.
17. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:**105–132.
18. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
19. **Langley, K. E., M. R. Villarejo, A. V. Fowler, P. J. Zamenhof, and I. Zabin.** 1975. Molecular basis of β-galactosidase α-complementation. Proc. Natl. Acad. Sci. USA **72:**1254–1257.
20. **Lindgren, P. B., R. C. Peet, and N. J. Panopoulos.** 1986. Gene cluster of *Pseudomonas syringae* pv. "*phaseolicola*" controls pathogenicity of bean plants and hypersensitivity on nonhost plants. J. Bacteriol. **168:**512–522.
21. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
22. **McKenney, K., H. Shimatake, D. Court, U. Schmeissner, C. Brady, and M. Rosenberg.** 1981. A system to study promoter and terminator signals recognized by *Escherichia coli* RNA polymerase, p. 383–414. *In* J. G. Chirikjian and T. S. Papas (ed.), Gene amplification and analysis. Elsevier/North-Holland Press, New York.
23. **Mills, D.** 1985. Transposon mutagenesis and its potential for studying virulence genes in plant pathogens. Annu. Rev. Phytopathol. **23:**381–419.
24. **Mills, D., and F. Niepold.** 1987. Molecular analysis of pathogenesis of *Pseudomonas syringae* pv. *syringae*, p. 185–200. *In* S. Nishimura, C. P. Vance, and N. Doke (ed.), Molecular determinants of plant diseases. Japan Scientific Societies Press, Tokyo.
25. **Mills, D., F. Niepold, and M. Zuber.** 1985. Cloned sequence controlling colony morphology and pathogenesis of *Pseudomonas syringae*, p. 97–102. *In* I. Sussex, A. Ellingboe, M. Crouch, and R. Malmberg (ed.), Plant cell/cell interactions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
26. **Mizusawa, S., S. Nishimura, and F. Seela.** 1986. Improvement of dideoxy chain termination method of DNA sequencing by use of deoxy-7-deazaguanosine triphosphate in place of dGTP. Nucleic Acids Res. **14:**1319–1324.
27. **Mount, D. W., and B. Conrad.** 1984. Microcomputer programs for graphic analysis of nucleic acid and protein sequences. Nucleic Acids Res. **12:**811–817.
28. **Murooka, Y., H. Iwamoto, A. Hamamoto, and T. Yamauchi.** 1987. Efficient transformation of phytopathogenic strains of *Xanthomonas* species. J. Bacteriol. **169:**4406–4409.
29. **Nathans, J., and D. S. Hogness.** 1983. Isolation, sequence analysis, and intron-exon arrangement of the gene encoding bovine rhodopsin. Cell **34:**807–814.

30. **Niepold, F., D. Anderson, and D. Mills.** 1985. Cloning determinants of pathogenesis from *Pseudomonas syringae* pathovar *syringae*. Proc. Natl. Acad. Sci. USA **82**:406–410.
31. **Osbourn, A. E., C. E. Barber, and M. J. Daniels.** 1987. Identification of plant-induced genes of the bacterial pathogen *Xanthomonas campestris* pathovar *campestris* using a promoter-probe plasmid. EMBO J. **6**:23–28.
32. **Panopoulos, N., and R. Peet.** 1985. The molecular genetics of plant pathogenic bacteria and their plasmids. Annu. Rev. Phytopathol. **23**:381–419.
33. **Remaut, E., H. Tsao, and W. Fiers.** 1983. Improved plasmid vectors with a thermoinducible expression and temperature-regulated runaway replication. Gene **22**:103–113.
34. **Reznikoff, W. S., and W. R. McClure.** 1986. *E. coli* promoters, p. 1–33. *In* W. Reznikoff and L. Gold (ed.), Maximizing gene expression. Butterworth Publishers, Stoneham, Mass.
35. **Rosenberg, M., and D. Court.** 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. **13**:319–353.
36. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
37. **Shine, J., and L. Dalgarno.** 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71**:1342–1346.
38. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
39. **Sonlyai, G., M. Hevesi, Z. Banfalvi, Z. Klement, and A. Kondorosi.** 1986. Isolation and characterization of non-pathogenic and reduced virulence mutants of *Pseudomonas syringae* pv. *phaseolicola* induced by Tn*5* transposon insertions. Physiol. Mol. Plant Pathol. **29**:369–380.
40. **Sutcliffe, J. G.** 1978. Nucleotide sequence of the ampicillin resistance gene of *Escherichia coli* plasmid pBR322. Proc. Natl. Acad. Sci. USA **75**:3737–3741.
41. **Xu, P., S. Leong, and L. Sequeira.** 1988. Molecular cloning of genes that specify virulence in *Pseudomonas solanacearum*. J. Bacteriol. **170**:617–622.
42. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.
43. **Yanofsky, C.** 1981. Attenuation in the control of expression of bacterial operons. Nature (London) **289**:751–758.