JOURNAL OF BACTERIOLOGY, Oct. 1988, p. 4846–4854
0021-9193/88/104846-09$02.00/0

Vol. 170, No. 10

Characterization and Expression of Two Avirulence Genes Cloned from *Pseudomonas syringae* pv. *glycinea*

S. TAMAKI,[1] D. DAHLBECK,[2] B. STASKAWICZ,[2] AND N. T. KEEN[1]*

Department of Plant Pathology, University of California, Riverside, California 92521,[1] and Department of Plant Pathology, University of California, Berkeley, California 94720[2]

Received 14 March 1988/Accepted 22 July 1988

Two avirulence genes, *avrB* and *avrC*, from race 0 of *Pseudomonas syringae* pv. *glycinea*, were sequenced and found to encode single protein products of 36 and 39 kilodaltons, respectively. The proteins had neither recognizable signal peptide sequences nor significant stretches of hydrophobic amino acids that might indicate membrane association. Both *avrB* and *avrC* had relatively low position 3 and overall G+C contents, which suggests that they may have been recently introduced into *P. syringae* pv. *glycinea*. The deduced amino acid sequences of the proteins encoded by *avrB* and *avrC* shared 42% identical amino acids. However, when introduced into race 4 of *P. syringae* pv. *glycinea*, each gene directed a unique pattern of hypersensitive reactions on several differential soybean cultivars. The *avrC* protein was overproduced in *Escherichia coli* cells and deposited as insoluble inclusion bodies in the cell cytoplasm. The *avrC* protein could be solubilized with urea-octyl glucoside treatment, but neither the solubilized protein nor the intact inclusion bodies elicited a hypersensitive reaction in soybean leaves.

Pseudomonas syringae pv. *glycinea* is the causal agent of bacterial blight of soybeans. This host-pathogen system has proven useful for the study of mechanisms that determine hypersensitive resistance and race specificity. Several races of the pathogen have been defined by their unique patterns of incompatible and compatible reactions on a series of differential soybean cultivars (1), and one plant resistance gene has been defined by classical genetic crosses (16). The dominant resistance gene, *Rpg1*, corresponds to a single avirulence gene, *avrB*, which occurs in two different races of *P. syringae* pv. *glycinea* and genetically defines a gene-for-gene interaction (21). The pathogen multiplies rapidly in inoculated leaves of a compatible soybean cultivar, resulting in water-soaked lesions. An incompatible plant reaction results in a rapid, necrotic, hypersensitive reaction (HR), followed after several hours by the accumulation of phytoalexins at the infection site and the restriction of pathogen multiplication (8, 13).

Pathogen avirulence genes corresponding to single dominant disease resistance genes in plant hosts have been demonstrated in many studies (5). The genetic complementarity of avirulence genes and resistance genes may further denote the existence of a corresponding biochemical complementarity wherein avirulence gene products or their catalytic products are recognized as elicitors by putative plant receptors encoded by plant disease resistance genes (7). To test this hypothesis and to study the basis of gene-for-gene complementarity, we began to clone avirulence genes from bacterial pathogens, characterize these genes, and investigate the properties of their protein products. We initially cloned (20) and sequenced (17) an avirulence gene from race 6 of *P. syringae* pv. *glycinea* called *avrA*. More recently, Staskawicz et al. (21) cloned two additional avirulence genes, designated *avrB* and *avrC*, from race 0 of *P. syringae* pv. *glycinea*. One of these genes, *avrB*, was identical or closely related to a single avirulence gene detected in race 1 of *P. syringae* pv. *glycinea*. In this paper, we present the DNA sequences of the *avrB* and *avrC* genes as well as describe overexpression of the *avrC* gene in *Escherichia coli* cells.

* Corresponding author.

MATERIALS AND METHODS

Bacterial strains and plasmids. A *P. syringae* pv. *glycinea* race 4 strain naturally resistant to ampicillin and selected for resistance to rifampin (Table 1) was maintained on King medium B (KMB) agar containing 100 μg of rifampin and 50 μg of ampicillin per ml and was grown at 28°C. *E. coli* cells were generally cultured on LB broth or agar media supplemented with appropriate antibiotics. Plasmid clones of pRK415 or pDSK519 (Table 1) were conjugated from *E. coli* into *P. syringae* pv. *glycinea* race 4 *rif amp* cells via triparental matings with the helper plasmid pRK2013 (4). Mating mixtures were allowed to incubate at 28°C for 12 h on KMB agar and were plated onto KMB agar containing 100 μg of rifampin, 50 μg of ampicillin, and 25 μg of tetracycline per ml. Colonies which grew on the final medium were restreaked, and single-colony isolates were used for plant inoculations. Plasmids containing *avr* genes and associated DNA were designated pAVR, which we have registered with the Plasmid Registry Center (12).

Plant culture and inoculation methods. Soybean plants were grown from seed as previously described (13) in plant growth chambers at 22°C. Bacteria were grown overnight on KMB agar plates or in KMB broth supplemented with appropriate antibiotics at 28°C. Bacterial cells were suspended in water and adjusted to absorbance values of 1.0 or 0.1 (corresponding to ca. 2×10^9 and 2×10^8 CFU/ml, respectively). Primary soybean leaves which were nearly fully expanded were inoculated with a Hagborg apparatus (13). Solubilized protein preparations were infiltrated with the same apparatus. Plants were maintained in growth chambers at 22°C in all cases, and reactions were visually monitored daily for 5 days.

Recombinant DNA techniques. DNA-modifying enzymes were generally purchased from New England BioLabs, Inc.,

TABLE 1. Bacterial strains, plasmids, and phage used in this study

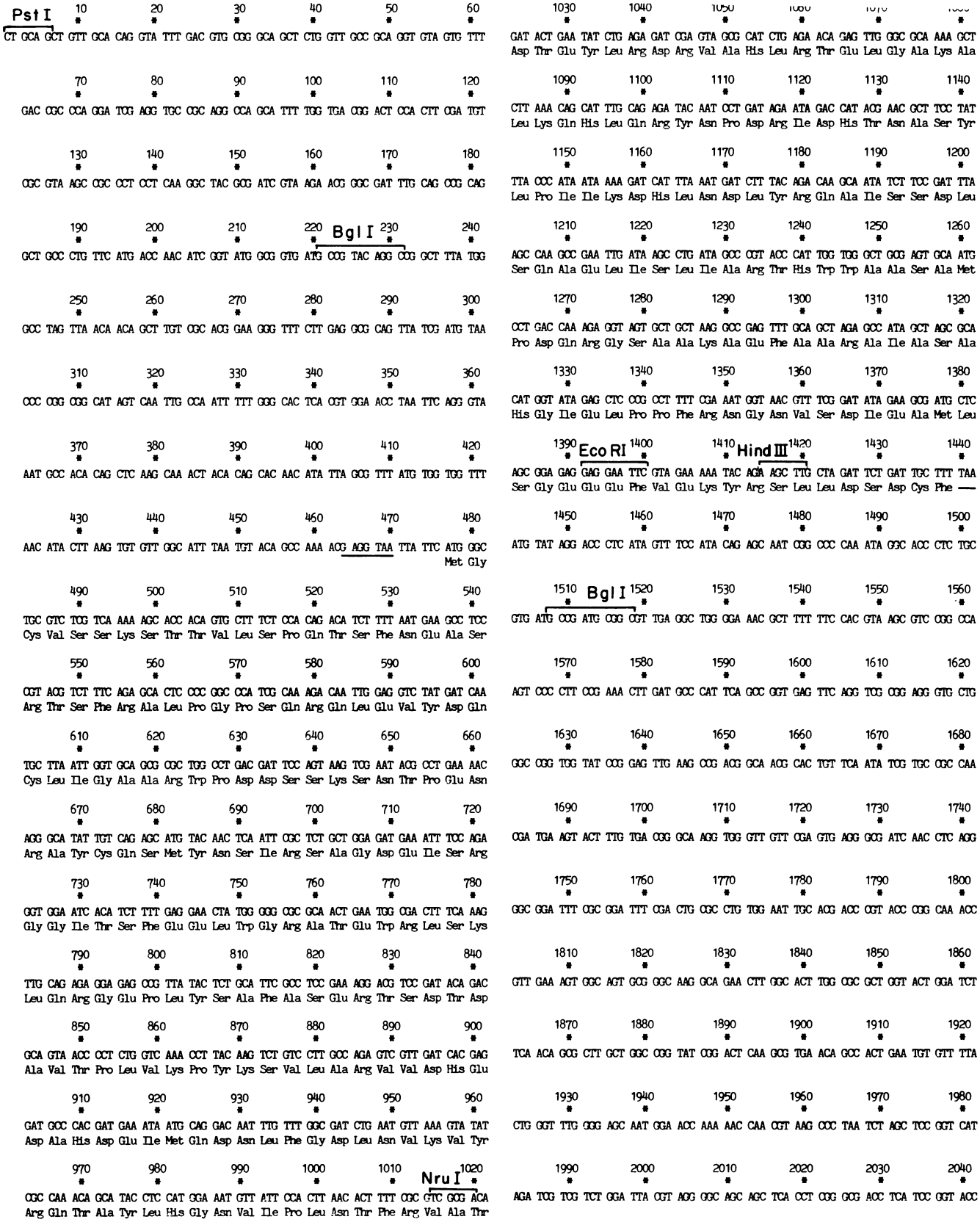
Bacterial strains, plasmids, or phage	Description	Reference or source
Bacteria		
P. syringae pv. *glycinea* race 4 *rif amp*		13
E. coli		
DH5α	*endA1 hsdR17* (r_K^- m_K^+) F$^-$ *supE44 thi-1* λ$^-$ *recA1 gyrA relA1* φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169	BRL[a]
JM-109	*recA1* Δ(*lac pro*) *endA1 gyrA96 thi-1 hsdR17 supE44 relA1* F′ *traD36 proAB*$^+$ *lacI*^{q}ZΔM15	23
MV1193	Δ(*lac-proAB*) *thi rpsL endA sbcB15 hspR4* Δ(*srl-recA*) *306*::Tn*10* (Tet^r) [F′:*traD36 proAB lacI*^{q}ZΔM15]	Messing, unpublished data
HB101	F$^-$ *hsdS20* (*hsdR hsdM*) *recA13 ara-14 proA2 lacY1 galK2 rpsL20* (Str^r) *xyl-5 mtl-1 supE44* λ−	14
BMH 71-18	Δ(*lac pro*) F′ *lacI*^{q}ZΔM15-*pro*$^+$	19
Plasmids		
pRK415	Shuttle vector based on pRK404	Keen et al.[b]
pDSK519	Shuttle vector based on pRSF1010	Keen et al.
pUC18, pUC19	Cloning vectors	23
pUC118, pUC119	pUC18 and pUC19 with the M13 replication origin	22
pUC129	pUC119 with additional polylinker cloning sites	Keen et al.
pUR292	*lacZ* fusion protein vector	19
pPSG0002	2.2-kb *Pst*I fragment cloned from race 0 of *P. syringae* pv. *glycinea* containing the *avrB* gene	21
pPSG0101	2.7-kb *Pst*I fragment cloned from race 0 of *P. syringae* pv. *glycinea* containing the *avrC* gene	21
pAVRB1	1,287-bp *Bgl*I fragment from pPSG0002 cloned in opposite orientation to promoter in pUC118	This study
pAVRB2	Same as pAVRB1 except in downstream orientation to vector *lac* promoter	This study
pAVRCp	2.7-kb *Pst*I insert from pPSG0101 cloned in *Pst*I site of pUC119 downstream of vector promoter	This study
pAVRC1	Sequencing deletion from pPSG0101 including 91 bp 5′ to start codon of *avrC* gene and continuing to downstream *Pst*I site of pPSG0101	This study
pAVRC4	1.2-kb fragment from 91 bp upstream of ATG start codon to *Not*I site downstream of the ORF; sequencing deletion was used for 5′ end of fragment; fragment was blunted with S1 nuclease and cloned downstream of vector promoter in *Sma*I site of pUC19	This study
pAVRC2	Same construct as pAVRC4 except cloned in pUC119	This study
pAVRC3	Same as pAVRC2 except opposite orientation	This study
pAVRC11	*Hin*dIII fragment from pAVRC4 cloned into pUR292 in antiorientation, leading to truncated fusion protein	This study
pAVRC12	*Hin*dIII fragment from pAVRC4 cloned into pUR292 in orientation leading to *lacZ-avrC* fusion protein	This study
pAVRC16	Fragment from leftward *Pst*I site of pPSG0101 to unique *Not*I site, containing *avrC*; cloned downstream of promoter in pUC129	This study
pAVRC19	ORF 2 of pPSG0101 from *Bgl*II site at base 1591 to *Bgl*II site at base 2494, cloned into *Bam*HI site of pUC119 (opposite orientation to vector promoter)	This study
pAVRC20	Same construct as pAVRC19 except *Bgl*II insert was cloned downstream of vector promoter	This study
Phage		
M13mp18, M13mp19	Phage for generation of single-strand DNA	23

[a] BRL, Bethesda Research Laboratories.
[b] N.T. Keen, S. Tamaki, D. Kobayashi, and D. Trollinger, Gene, in press.

Beverly, Mass., or Bio-Rad Laboratories, Richmond, Calif. Subcloning was generally performed on low-melting-point agarose gels by the method of Crouse et al. (2). Other general methods were as described by Maniatis et al. (14) or Keen and Tamaki (9).

DNA sequencing. *Pst*I fragments coding *avrB* or *avrC* were used for BAL 31 deletions, and these deletions were then subcloned in phage M13 and sequenced by the dideoxy-chain termination method (3). Both strands were sequenced to confirm data. DNA sequence data were analyzed by the computer programs of Pustell and Kafatos (18) as well as by the BIONET system supplied through Intelligenetics Corp., Mountain View, Calif.

Protein electrophoresis. *E. coli* cells carrying desired plasmids were grown overnight at 37°C in 15 ml of LB broth containing 50 μg of ampicillin with or without 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Bacterial cells were pelleted by centrifugation for 5 min at 5,000 × *g* and suspended in 2 ml of water per 15 ml of original culture; 50 μl of this suspension was mixed with 50 μl of 2.5× Laemmli sample solution (10). The mixture was boiled for 5 min, and samples were electrophoresed on 10% acrylamide gels which


PstI
BglI
NruI
EcoRI
HindIII
BglI

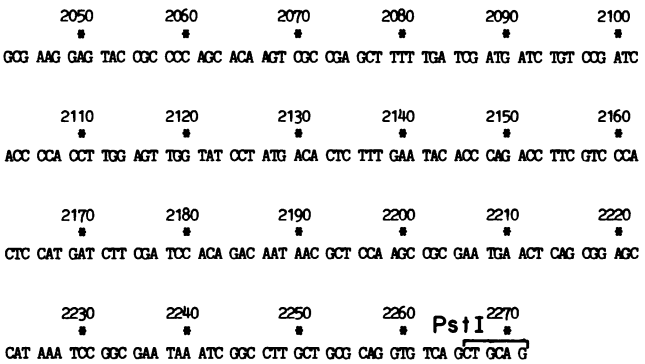


FIG. 1. Sequence of the *avrB* gene and its translated protein sequence. Selected restriction sites noted in the text are shown, and the Shine-Dalgarno box is underlined.

were then stained with Coomassie blue R250 and destained as described by Lane (11).

RESULTS

Sequencing of the *avrB* gene. Plasmid pPSG0002, containing a 2.2-kilobase-pair (kb) *Pst*I insert of *P. syringae* pv. *glycinea* race 0 DNA, was constructed by Staskawicz et al. (21). The DNA sequence of this fragment (Fig. 1) disclosed the presence of a single long open reading frame (ORF) of 963 nucleotides beginning at base 477 and ending at base 1439. No other reading frames of significant length were detected in either strand. The observed ORF was composed of 46% G+C residues, with a position 3 GC content of 40%. A Tn*5* insertion near the *Nru*I site (Fig. 1) was observed to abolish the *avrB* phenotype (data not shown). A *Pst*I-*Hin*dIII fragment lacking 24 base pairs (bp) of the 3′ terminus retained *avrB* activity; however, the corresponding *Pst*I-*Eco*RI fragment lacking an additional 21 bp from the 3′ end did not yield an HR in the appropriate soybean cultivars. These data confirmed that the observed ORF is responsible for the *avrB* phenotype.

The *avrB* ORF was preceded by an AG-rich Shine-Dalgarno box, and a sequence resembling the *E. coli* consensus −10 promoter element was observed at base 445, but no sequence resembling the *E. coli* −35 promoter element was present (Fig. 1). Identification of the *avrB* promoter elements, however, will require transcription studies as defined by S1 nuclease protection and primer extension experiments. The ORF initiating at position 477 was further subcloned as the *Bgl*I fragment occurring between bases 227 and 1514. This fragment was blunt ended with S1 nuclease, and both orientations were recovered by subcloning into the *Sma*I site of pUC118 (Table 1). The insert of both constructs was then transferred to the broad-host-range shuttle vectors pDSK519 and pRK415 with the flanking polylinker *Bam*HI and *Kpn*I sites such that the *avrB* ORF was oriented downstream from the vector *lac* promoters. The resultant plasmids were conjugated into race 4 of *P. syringae* pv. *glycinea*, and these cells were inoculated into soybean leaves. Bacteria containing the *avrB* constructs yielded HRs in the expected soybean cultivars, regardless of insert orientation in the vectors. Furthermore, the HRs appeared relatively rapidly (16 to 24 h postinoculation) and involved intense necrosis, similar to results seen in our previous experience with *avrB*, but unlike HRs occurring in response to the *avrC* gene (21).

The protein product of the single long ORF identified in Fig. 1 was predicted to contain 321 amino acids and have a molecular weight of 36,019. Attempts to confirm the production of this protein by sodium dodecyl sulfate-gel electrophoresis of lysates from *E. coli* HB101 or of DH5α cells carrying pUC118 constructs of *avrB* oriented downstream of the vector *lac* promoter have failed, since discernible protein bands were not observed. However, conditions have recently been devised for improved expression of *avrB* in *E. coli* (T. Huynh, D. Dahlbeck, and B. Staskawicz, manuscript in preparation).

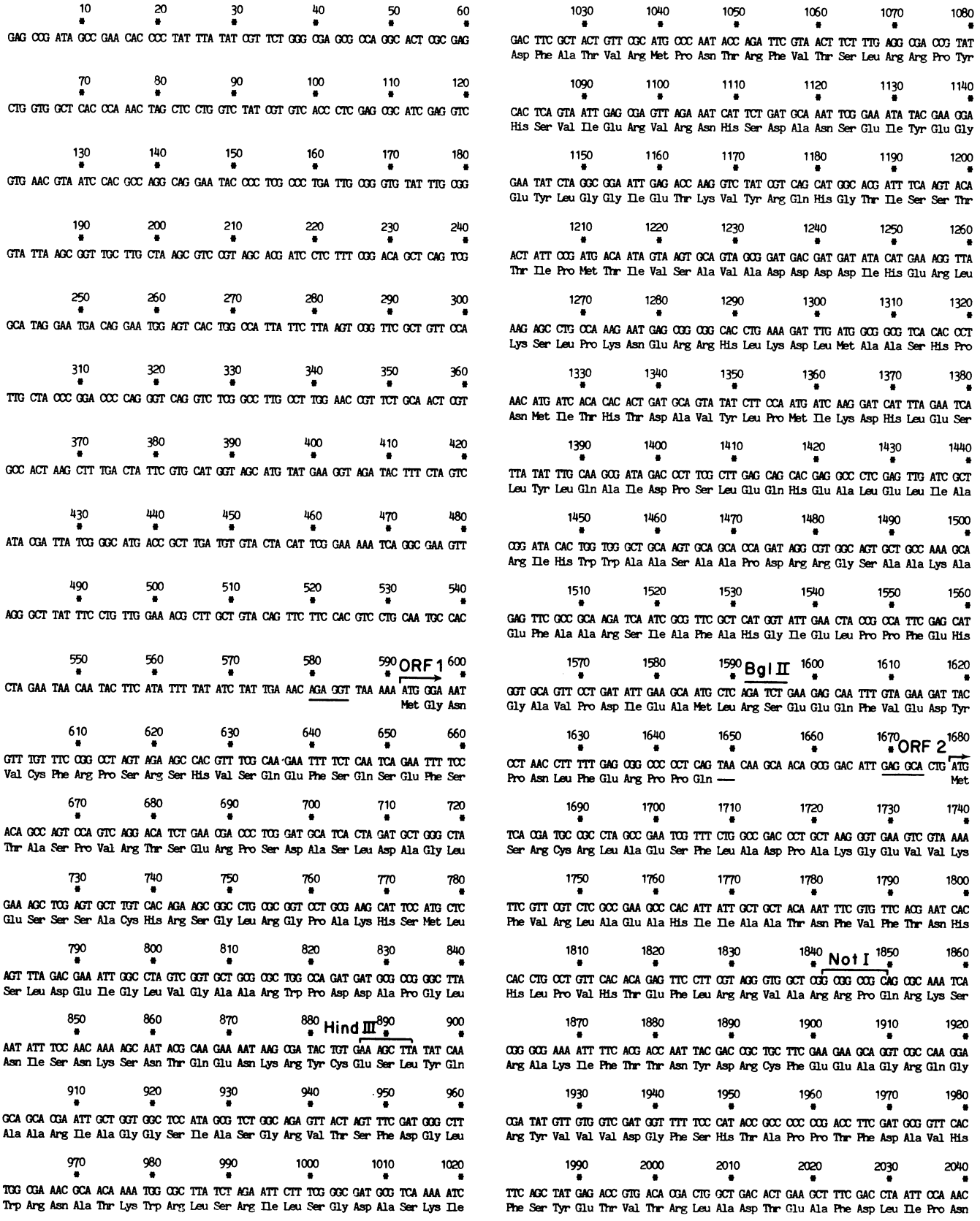
Sequencing of the *avrC* gene. A 2.7-kb *Pst*I fragment of *P. syringae* pv. *glycinea* race 0 DNA from pPSG0101, previously shown to confer the *avrC* phenotype (21), was sequenced. This DNA fragment contained two ORFs arranged in tandem (Fig. 2 and 3). No other ORFs of significant length were detected in either strand.

ORF 1 (Fig. 2) comprised 1,056 bp and encoded a protein with 352 amino acids and a calculated molecular weight of 39,145. This ORF had a GC content of 47% overall and 40% in position 3. Thus, there was no tendency for the strong position 3 GC bias frequently observed in *Pseudomonas* genes. The start codon is preceded by a sequence which would be expected to function as a ribosome-binding site (Fig. 2). A sequence similar to the *E. coli* −10 promoter box was present at base 542, but the corresponding −35 box was not seen. The amino-terminal sequence of the protein product showed no relationship to known bacterial signal peptide secretion sequences, and no significant stretches of hydrophobic amino acids were detected.

ORF 2 (Fig. 2) was composed of 720 bases and encoded a protein of 240 amino acids with a calculated weight of 27,483. This coding region had an overall GC content of 53% and a position 3 GC content of 56%. A sequence which weakly resembles a ribosome-binding site occurred 5′ to the start codon of ORF 2 (Fig. 2). The protein product lacked a recognizable signal peptide sequence and had no detectable amino acid homology with the protein encoded by ORF 1. ORF 2 lies 27 bp downstream of the termination codon of ORF 1 (Fig. 2), raising the possibility that the two genes may be cotranscribed or indeed cotranslated via readthrough of the ORF 1 stop codon.

To determine whether ORF 1 or 2 conferred the *avrC* phenotype, various subclones were assessed for the production of a hypersensitive response when conjugated into race 4 of *P. syringae* pv. *glycinea*. ORF 1, located between bases 592 and 1647, conferred an HR when all of the 5′ DNA or only 91 bp of upstream DNA were retained and the reading frame was cloned in either orientation in shuttle plasmids pRK415 or pDSK519. These HRs were observed only on cultivars previously shown to be incompatible with the *avrC* gene (21). Furthermore, the HRs conferred by ORF 1 in either orientation developed relatively slowly (40 to 48 h after inoculation) and were less intense than those which occurred in response to *avrB*, consistent with earlier observations with larger cloned DNA fragments (21). ORF 2, occurring between bases 1678 and 2397, did not yield an HR in soybean leaves when subcloned downstream from a vector *lac* promoter and introduced into race 4 of *P. syringae* pv. *glycinea* (Fig. 3). Thus, ORF 1 in Fig. 3 was sufficient to condition an HR and therefore constitutes the *avrC* gene.

Homology between *avrB* and *avrC* genes. Computer analyses of the translated *avrB* and *avrC* protein sequences disclosed considerable homology (Fig. 4). An overall amino acid identity of 46% was observed, but homology as high as 80% was noted in the carboxyl ends of the proteins. This region also contains three small hydrophobic domains, but overall it is hydrophilic like the remainder of the proteins.


ORF 1
Met Gly Asn
Hind III
Bgl II
ORF 2
Met
Not I

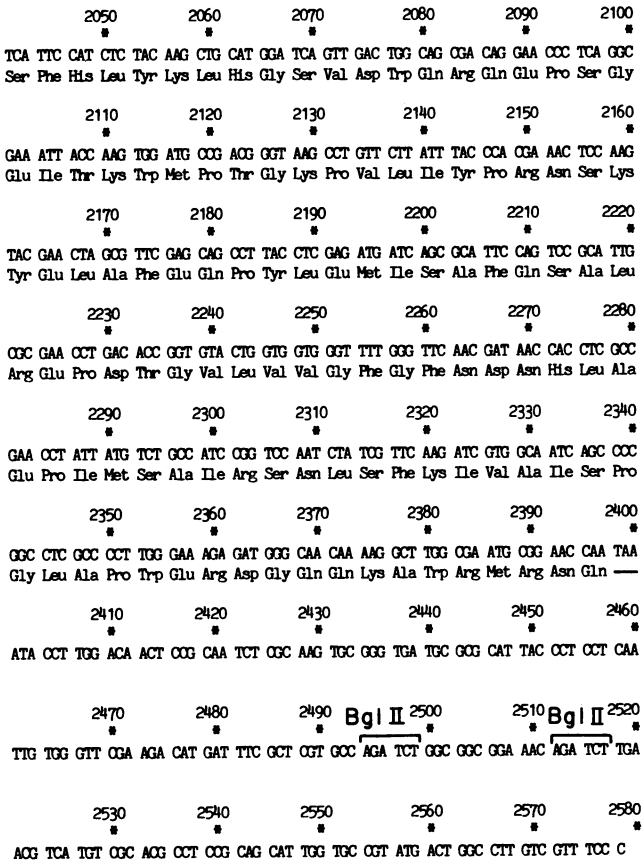


FIG. 2. Sequence of the *avrC* gene (ORF 1) and its deduced protein sequence. The translated amino acid sequence of ORF 2 initiating at position 1678 is also shown. Selected restriction endonuclease sites are shown, and Shine-Dalgarno sequences are underlined.

The *avrB* and *avrC* gene products can be aligned colinearly and adjusted for extra amino acids at the amino-terminal end of *avrC*, an 18-bp insertion near the 5′ end of *avrB* and a single amino acid deletion in the middle of the *avrC* protein.

Expression of *avrC* gene and ORF 2. *E. coli* JM-109 cells containing the various constructs shown in Fig. 3 were grown in the presence or absence of IPTG, and whole cells were electrophoresed on sodium dodecyl sulfate-polyacrylamide gels. Cells containing ORF 1 (Fig. 3, *avrC*) produced a new protein product which had a molecular weight of 39,000 (Fig. 5), in agreement with the molecular weight predicted for the *avrC* protein from DNA sequence data. The intensity of this band was stronger when cells were supplied with IPTG to activate the vector *lac* promoter (Fig. 5, lane I) than when they were uninduced (Fig. 5, lane J). One construct, pAVRC2, resulted in the gradual appearance of a slightly smaller protein as *E. coli* cells were sequentially subcultured (Fig. 5, lane G). However, the comparable construct, pAVRC4 (Fig. 5, lanes I and O; Table 1), did not show similar signs of instability.

A construct in which the *avrC* ORF, from the internal *Hin*dIII site (Fig. 2) to the *avrC* translational stop, was fused to the 3′ end of the *lacZ* gene resulted, as predicted, in a fusion protein which was ca. 27 kDa larger than β-galactosidase (Fig. 5, lane D). The control construct into which the *avrC* DNA was inserted in the opposite orientation resulted in a fusion protein only slightly larger than native β-galactosidase (Fig. 5, lane C).

E. coli cells carrying clone pAVRCp yielded a relatively weak band at 39 kDa but a strong band at 29 kDa, which was presumed to result from ORF 2 (Fig. 5, lanes E and L). Cloning of ORF 2 in the absence of ORF 1 confirmed this suspicion, since pAVRC20, in which the ORF was oriented downstream of the vector *lac* promoter, yielded the same 29-kDa band (Fig. 5, lane N), but the opposite orientation, in pAVRC19, did not (Fig. 5, lane M). Thus, ORF 2 appears to be functional in *E. coli* whether cloned singly or in the presence of the intact ORF 1. The 5-bp spacing between the putative Shine-Dalgarno sequence and the translational start site of ORF 2 therefore appears functional in *E. coli* cells.

Characterization of *avrC* protein overproduced in *E. coli* cells. Substantial amounts of the *avrC* protein and the *avrC-lacZ* fusion protein were produced in *E. coli* cells which were supplied IPTG to induce the vector *lac* promoters (Fig. 5). However, fractionation of *E. coli* cells showed that the native *avrC* protein and the *lacZ* fusion proteins were exclusively present in the cytoplasmic fraction and were pelleted upon low-speed centrifugation of spheroplast or whole-cell lysates (data not shown). It is therefore assumed that the overproduced proteins are deposited in paracrystalline inclusion bodies, as is often the case with overproduced foreign proteins in *E. coli* (15). Low-speed-centrifugation pellets containing the *avrC* protein were readily solubilized by stirring at room temperature for 10 min with 0.01 M Tris hydrochloride (pH 7.5) containing 2 M urea and 0.1% octyl glucoside. The reagents could be removed by exhaustive dialysis against 0.01 M Tris hydrochloride (pH 7.5), and the *avrC* protein would remain soluble, yielding water-clear solutions up to 10 mg/ml. Sodium dodecyl sulfate gels showed that the *avrC* protein constituted ca. 90% of the total protein present in these preparations (data not shown). Dialysis of the solubilized *avrC* protein against water, however, resulted in significant precipitation.

Infiltration of *avrC* protein and *lacZ-avrC* fusion protein into soybean leaves. The *avrC* and *avrC-lacZ* fusion proteins were solubilized, dialyzed against 5 mM Tris hydrochloride (pH 7.5), and infiltrated into primary leaves of the soybean cultivars Harosoy, Norchief, Acme, Centennial, and Flambeau. Neither, however, elicited a visibly detectable HR in soybean leaves of any cultivar infiltrated with up to 5 mg of total protein per ml. Similarly, infiltration of intact *E. coli* cells which were overproducing the *avrC* protein did not lead to a detectable HR in any of the standard soybean differential cultivars. Experiments in which the dialyzed *avrC* protein was mixed with living race 4 cells of *P. syringae* pv. *glycinea* resulted in typical compatible plant reactions with no signs of an HR observed at any time after inoculation. Thus, the *avrC* protein does not appear to possess elicitor activity in soybean leaves.

DISCUSSION

Two avirulence genes from race 0 of *P. syringae* pv. *glycinea*, *avrB* and *avrC*, were sequenced and found to encode single protein products of 36 and 39 kDa, respectively, as determined by DNA sequence data and expression in *E. coli*. The two genes can be readily distinguished by the particular cultivars on which they elicit the HR and by the distinctive HRs which they condition; *avrB* confers a more rapid and necrotic HR than *avrC*. Despite these differences, the two coding regions share considerable amino acid identity (Fig. 4). Since both genes were cloned from race 0 of *P. syringae* pv. *glycinea* (21), it is possible that one of them originated by duplication and modification of the other.

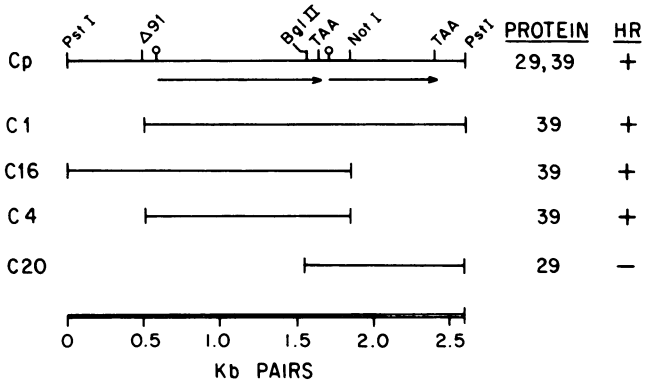


FIG. 3. Various subclones of the insert DNA of pPSG0101, carrying *avrC*. Fragments were recloned into pUC119, tested for protein synthesis in *E. coli* by gel electrophoresis, and shuttled to pRK415 or pDSK519. The resulting plasmids were transferred to race 4 of *P. syringae* pv. *glycinea* and tested for their ability to elicit an HR on the differential soybean cultivars. Subclones were prepared by using the noted restriction sites or an exonuclease-S1 sequencing deletion (Δ91) (Table 1). Abbreviations: Cp, pAVRCp; C1, pAVRC1; C16, pAVRC16; C4, pAVRC4; and C20, pAVRC20.

Significantly, both genes were previously observed to possess repeated DNA sequences in the flanking 3′ and 5′ regions (21). Furthermore, while *avrB* is located on the race 0 chromosome, *avrC* appears to occur on a plasmid (unpublished data). Coupled with the relatively low GC content of the two genes, especially in position 3, these observations raise the possibility that *avrB* and *avrC* were relatively recently introduced into *P. syringae* pv. *glycinea* from another source.

As predicted by the gene-for-gene relationship, the avirulence genes thus far identified fully account for the race phenotypes of the investigated *P. syringae* pv. *glycinea* races. Thus, *avrA* determines the race 6 phenotype (17, 20), whereas *avrB* accounts for the race 1 phenotype (21). However, race 0 contains both *avrB* and *avrC* (21), and the function of both of these genes leads to the unique reactions of race 0 on the standard soybean differential set. The *avrA* gene (17) has no detectable homology to *avrB* or *avrC*, but the latter two genes and their protein products share considerable homology with each other (Fig. 4). We are currently constructing recombinants between these genes and examining phenotypes that are apparent after the introduction of the genes into race 4 of *P. syringae* pv. *glycinea* and inoculation of the differential soybean cultivars.

The *avrC* gene was subcloned on a 2.7-kb *Pst*I fragment which contains two ORFs (Fig. 2). Further subcloning (Fig. 3) showed that only ORF 1 is required for the *avrC* phenotype in *P. syringae* pv. *glycinea* when the ORF is introduced into the bacterium on a multicopy plasmid. However, it is possible that ORF 2 contributes to the avirulent phenotype in the wild-type race 0 bacterium. Southern blot analyses have indicated that ORF 2 is present in two other races of *P. syringae* pv. *glycinea* which do not contain *avrC* (21). This may suggest that ORF 2 has a function in the bacterium exclusive of the presence of *avrC*.

Both *avrB* and *avrC* contain codons such as CCT, CAA, and AAT which are not frequently used in highly expressed *E. coli* proteins (6). Whether such codon usage affects expression in *P. syringae* pv. *glycinea* is not clear. However, *avrC* was satisfactorily overexpressed in *E. coli* by using

```
                                    avrB  MGCVSSKSTTVLSPQTSFNEASRTSFRAL
                                                  :      :       :      :
avrC  MGNVCFRPSRSHVSQEFSQSEFSTASPVRTSERPSDASLDAGLESSSACHRSGL

30 PGPSQRQLEVYDQC-LIGAARWPDD------SSKSNTPENRAYCQSMYNSIRSAGDE
    ::         :   : ::::::::       : :::: ::  :: : :   : ::
55 RGPAKHSMLSLDEIGLVGAARWPDDAPGLNISNKSNTQENKRYCESLYQAARIAGGS

 80 ISRGGITSFEELWGRATEWRLSKLQRGEPLYSAFASERTSDTDAVTPLVKPYKSVL
    :  :  :::  ::  :: ::::    :       ::  :   :  :: :  :: ::
112 IASGRVTSFDGLWRNATKWRLSRILSGDASKIDFATVRMPNTRFVTSLRRPYHSVI

136 ARVVDHEDAHDEIMQDNLFGDLNVKVYRQTAYLHGNVIPLNTFRVATDTEYLRDRV
     ::  : ::  ::       :    :::::         ::         :       :
168 ERVRNHSDANSEIYEGEYLGGIETKVYRQHGTISSTTIPMTIVSAVADDDDIHERL

192 AHLRTELGAKALKQHLQRYNPDRIDHTNASYLPIIKDHLNDLYRQAISSDLSQAEL
     :         ::        :  : :: : ::: :::::  :: :::   : : :
224 KSLPKN-ERRHLKDLMAASHPNMITHTDAVYLPMIKDHLESLYLQAIDPSLEQHEA

248 ISLIARTHWWAASAMPDQRGSAAKAEFAARAIASAHGIELPPFRNGNVSDIEAMLS
      :::: :::::::: :: ::::::::::::: :: :::::::::  : : ::::::
279 LELIARIHWWAASAAPDRRGSAAKAEFAARSIAFAHGIELPPFEHGAVPDIEAMLR

304 GEEEFVEKYRSLLDSDCF 321
     :: ::: :  :
335 SEEQFVEDYPNLFERPPQ 352
```

FIG. 4. Homology of the protein products of *avrB* and *avrC*. Connecting dots denote identical amino acids.

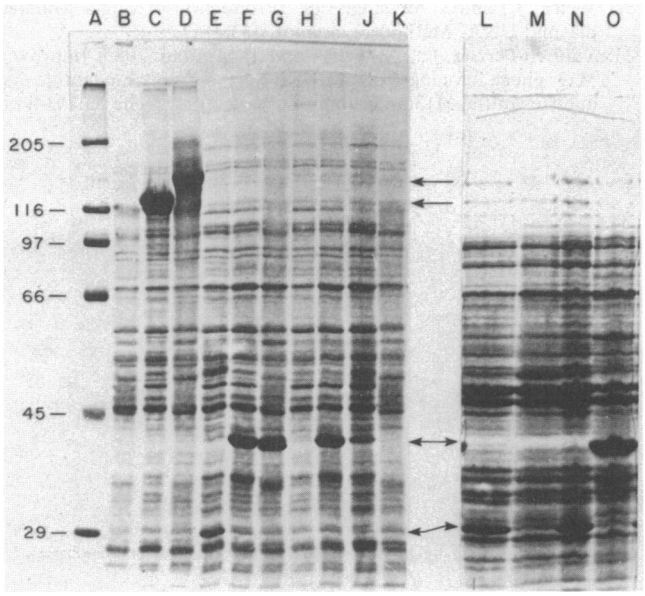


FIG. 5. Sodium dodecyl sulfate-polyacrylamide gel of *E. coli* JM-109 cells containing various plasmid constructs of the *avrC* gene (ORF 1) or ORF 2 (Fig. 2). All cells were grown on LB medium overnight at 28°C in the presence of 1 mM IPTG unless otherwise noted. Lanes: A, molecular weight standards with sizes noted on the left; B, cells with pUC119 only; C, BMH 71-18 cells containing pAVRC11 (arrow denotes β-galactosidase fusion protein only slightly larger than native β-galactosidase at 116 kDa); D, BMH 71-18 cells containing pAVRC12 (arrow denotes *lacZ-avrC* fusion protein at ca. 145 kDa); E, pAVRCp; F, pAVRC1 (arrow denotes protein at ca. 39 kDa); G, pAVRC2; H, pAVRC3; I, pAVRC4; J, cells with clone pAVRC4 but not supplied with IPTG; K, pUC119 only; L, pAVRCp; M, pAVRC19; N, pAVRC20 showing protein band at 29 kDa as denoted by the arrow; O, pAVRC4.

transcriptional fusions of the *avrC* gene behind the vector *lac* promoter of pUC119 (Fig. 5). Most of the protein, however, occurred in insoluble aggregates which were pelleted from cell lysates. These were assumed to represent the inclusion bodies frequently formed by overexpressed proteins in *E. coli* (15).

We do not know the functions of the *avrB* and *avrC* gene products in *P. syringae* pv. *glycinea* cells, and database searching with programs available through the BIONET resource has failed to locate known proteins with significant homology with the *avrB* or *avrC* proteins. As previously observed with the *avrA* protein product (17), the proteins encoded by *avrB* and *avrC* do not contain recognizable signal sequences or significant hydrophobic regions. Infiltration of suspensions of the inclusions or soluble preparations of the *avrC* protein prepared with urea-octyl glucoside did not result in a detectable HR in any of the soybean differential cultivars. The *avrC* protein product may, therefore, not function as an elicitor of the soybean HR per se. This possibility must be considered with caution, however, since expression of the gene in *P. syringae* pv. *glycinea* and *E. coli* cells may vary in some way, or solubilization of the *avrC* protein in our experiments may have resulted in structural changes that abolished elicitor activity.

ACKNOWLEDGMENTS

The research was supported by National Science Foundation grants DCB-8502132 and DCB-8709867 to N.T.K. and U.S. Department of Agriculture grant 86-CRCR-1-2086 to B.S. Some of the sequence data were analyzed with help from the BIONET facility, whose funding is provided by Public Health Service grant U41 RR-01685-02 from the Biomedical Research Technology Program, Division of Research Resources, National Institutes of Health.

Soybean seeds were generously supplied by R.I. Buzzell.

LITERATURE CITED

1. **Cross, J. E., B. W. Kennedy, J. W. Lambert, and R. L. Cooper.** 1961. Pathogenic races of the bacterial blight pathogen of soybeans, *Pseudomonas glycinea*. Plant Dis. Rep. **50**:557–560.
2. **Crouse, G. F., A. Frischauf, and H. Lehrach.** 1983. An integrated and simplified approach to cloning into plasmids and single-stranded phages. Methods Enzymol. **101**:78–89.
3. **Dean, C., P. van den Elzen, S. Tamaki, P. Dunsmuir, and J. Bedbrook.** 1985. Differential expression of the eight genes of the petunia ribulose diphosphate carboxylase small subunit multigene family. EMBO J. **4**:3055–3061.
4. **Ditta, G., S. Stanfield, D. Corbin, and D. R. Helinski.** 1980. Broad host range cloning system for Gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. Proc. Natl. Acad. Sci. USA **27**:7347–7351.
5. **Ellingboe, A. E.** 1984. Genetics of host-parasite relations: an essay. Adv. Plant Pathol. **2**:131–151.
6. **Goug, M., and C. Gautier.** 1982. Codon usage in bacteria: correlation with gene expressivity. Nucleic Acids Res. **10**:7055–7074.
7. **Keen, N. T.** 1982. Specific recognition in gene-for-gene host-parasite systems. Adv. Plant Pathol. **1**:35–82.
8. **Keen, N. T., and M. J. Holliday.** 1982. Recognition of bacterial pathogens by plants, p. 179–217. *In* M. S. Mount and G. Lacy (ed.), Phytopathogenic procaryotes, vol. 2. Academic Press, Inc., New York.
9. **Keen, N. T., and S. Tamaki.** 1986. Structure of two pectate lyase genes from *Erwinia chrysanthemi* EC16 and their high-level expression in *Escherichia coli*. J. Bacteriol. **168**:595–606.
10. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
11. **Lane, L. C.** 1978. A simple method for stabilizing protein sulfhydryl groups during SDS-gel electrophoresis. Anal. Biochem. **86**:655–665.
12. **Lederberg, E. M.** 1986. Plasmid prefix designations registered by the plasmid reference center. Plasmid **15**:57–92.
13. **Long, M., P. Barton-Willis, B. J. Staskawicz, D. Dahlbeck, and N. T. Keen.** 1985. Further studies on the relationship between glyceollin accumulation and the resistance of soybean leaves to *Pseudomonas syringae* pv. *glycinea*. Phytopathology **75**:235–239.
14. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
15. **Masui, Y., T. Mizuno, and M. Inouye.** 1984. Novel high-level expression cloning vehicles: 10^4-fold amplification of *Escherichia coli* minor protein. Bio/Technology **2**:81–85.
16. **Mukherjee, D., J. W. Lambert, R. L. Cooper, and B. W. Kennedy.** 1966. Inheritance of resistance to bacterial blight in soybeans. Crop Sci. **6**:324–326.
17. **Napoli, C., and B. J. Staskawicz.** 1987. Molecular characterization and nucleic acid sequence of an avirulence gene from race 6 of *Pseudomonas syringae* pv. *glycinea*. J. Bacteriol. **169**:572–578.
18. **Pustell, J., and F. C. Kafatos.** 1984. A convenient and adaptable package of computer programs for DNA and protein sequence management, analysis and homology determination. Nucleic Acids Res. **12**:643–655.
19. **Rüther, U., and B. Müller-Hill.** 1983. Easy identification of cDNA clones. EMBO J. **2**:1791–1794.
20. **Staskawicz, B. J., D. Dahlbeck, and N. T. Keen.** 1984. Cloned avirulence gene of *Pseudomonas syringae* pv. *glycinea* determines race-specific incompatibility on *Glycine max*. Proc. Natl.

Acad. Sci. USA **81**:6024–6028.

21. **Staskawicz, B., D. Dahlbeck, N. Keen, and C. Napoli.** 1987. Molecular characterization of cloned avirulence genes from race 0 and race 1 of *Pseudomonas syringae* pv. *glycinea*. J. Bacteriol. **169**:5789–5794.
22. **Vieira, J., and J. Messing.** 1987. Production of single-stranded plasmid DNA. Methods Enzymol. **153**:3–11.
23. **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.