

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 94720²

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.

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.,

* Corresponding author.

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

Bacterial strains, plasmids, or phage	Description	Reference or source
Bacteria		
<i>P. syringae</i> pv. <i>glycinea</i> race 4 rif ^r amp		13
<i>E. coli</i>		
DH5 α	<i>endA1 hsdR17</i> (r _K ⁻ m _K ⁺) F ⁻ <i>supE44 thi-1 λ-recA1 gyrA relA1 ϕ80dlacZΔM15 Δ(lacZYA-argF)U169</i>	BRL ^a
JM-109	<i>recA1 Δ(lac pro) endA1 gyrA96 thi-1 hsdR17 supE44 relA1 F' traD36 proAB⁺ lacI^qZΔM15</i>	23
MV1193	<i>Δ(lac-proAB) thi rpsL endA sbcB15 hspR4 Δ(srl-recA) 306::Tn10 (Tet^r) [F':traD36 proAB lacI^qZΔM15]</i>	Messing, unpublished data
HB101	F ⁻ <i>hsdS20 (hsdR hsdM) recA13 ara-14 proA2 lacY1 galK2 rpsL20 (Str^r) xyl-5 mtl-1 supE44 λ-</i>	14
BMH 71-18	<i>Δ(lac pro) F' lacI^qZΔM15-pro⁺</i>	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	<i>lacZ</i> fusion protein vector	19
pPSG0002	2.2-kb <i>Pst</i> I fragment cloned from race 0 of <i>P. syringae</i> pv. <i>glycinea</i> containing the <i>avrB</i> gene	21
pPSG0101	2.7-kb <i>Pst</i> I fragment cloned from race 0 of <i>P. syringae</i> pv. <i>glycinea</i> containing the <i>avrC</i> gene	21
pAVRB1	1,287-bp <i>Bgl</i> II fragment from pPSG0002 cloned in opposite orientation to promoter in pUC118	This study
pAVRB2	Same as pAVRB1 except in downstream orientation to vector <i>lac</i> promoter	This study
pAVRCp	2.7-kb <i>Pst</i> I insert from pPSG0101 cloned in <i>Pst</i> I site of pUC119 downstream of vector promoter	This study
pAVRC1	Sequencing deletion from pPSG0101 including 91 bp 5' to start codon of <i>avrC</i> gene and continuing to downstream <i>Pst</i> I site of pPSG0101	This study
pAVRC4	1.2-kb fragment from 91 bp upstream of ATG start codon to <i>Not</i> 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 <i>Sma</i> 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	<i>Hind</i> III fragment from pAVRC4 cloned into pUR292 in antiorientation, leading to truncated fusion protein	This study
pAVRC12	<i>Hind</i> III fragment from pAVRC4 cloned into pUR292 in orientation leading to <i>lacZ-avrC</i> fusion protein	This study
pAVRC16	Fragment from leftward <i>Pst</i> I site of pPSG0101 to unique <i>Not</i> I site, containing <i>avrC</i> ; cloned downstream of promoter in pUC129	This study
pAVRC19	ORF 2 of pPSG0101 from <i>Bgl</i> II site at base 1591 to <i>Bgl</i> II site at base 2494, cloned into <i>Bam</i> HI site of pUC119 (opposite orientation to vector promoter)	This study
pAVRC20	Same construct as pAVRC19 except <i>Bgl</i> 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 μ g of ampicillin with or without 1 mM isopropyl- β -D-thiogalactopyranoside (IPTG). Bacterial cells were pelleted by centrifugation for 5 min at 5,000 $\times$ g and suspended in 2 ml of water per 15 ml of original culture; 50 μ l of this suspension was mixed with 50 μ l of 2.5 $\times$ Laemmli sample solution (10). The mixture was boiled for 5 min, and samples were electrophoresed on 10% acrylamide gels which

Pst I	10	20	30	40	50	60		1030	1040	1050	1060	1070	1080
CT GCA GCT GTT GCA CAG GTA TTT GAC GTG CCG GCA GCT CTG GTT GCC GCA GGT GTA GTG TTT								GAT ACT GAA TAT CTG AGA GAT CGA GTA GCG CAT CTG AGA ACA GAG TTG GCG GCA AAA GCT					
								Asp Thr Gln Tyr Leu Arg Asp Arg Val Ala His Leu Arg Thr Glu Leu Gly Ala Lys Ala					
	70	80	90	100	110	120		1090	1100	1110	1120	1130	1140
GAC CCG CCA GGA TCG AGG TGC CCG AGG CCA GCA TTT TGG TGA CCG ACT CCA CTT CGA TGT								CTT AAA CAG CAT TTG CAG AGA TAC AAT CCT GAT AGA ATA GAC CAT ACG AAC GCT TCC TAT					
								Leu Lys Gln His Leu Gln Arg Tyr Asn Pro Asp Arg Ile Asp His Thr Asn Ala Ser Tyr					
	130	140	150	160	170	180		1150	1160	1170	1180	1190	1200
CGC GTA AGC CCG CCT CCT CAA GGC TAC GCG ATC GTA AGA ACG GGC GAT TTG CAG CCG CAG								TTA CCG ATA ATA AAA GAT CAT TTA AAT GAT CTT TAC AGA CAA GCA ATA TCT TCC GAT TTA					
								Leu Pro Ile Ile Lys Asp His Leu Asn Asp Leu Tyr Arg Gln Ala Ile Ser Ser Asp Leu					
	190	200	210	220	230	240	Bgl I	1210	1220	1230	1240	1250	1260
GCT GCC CTG TTC ATG ACC AAC ATC GGT ATG GCG GTG ATG CCG TAC AGG CCG GCT TTA TGG								AGC CAA GGC GAA TTG ATA AGC CTG ATA GCG CGT ACC CAT TGG TGG GCT GCG AGT GCA ATG					
								Ser Gln Ala Glu Leu Ile Ser Leu Ile Ala Arg Thr His Trp Trp Ala Ala Ser Ala Met					
	250	260	270	280	290	300		1270	1280	1290	1300	1310	1320
GCC TAG TTA ACA ACA GCT TGT CCG ACG GAA GCG TTT CTT GAG CCG CAG TTA TCG ATG TAA								CCT GAC CAA AGA GGT AGT GCT GCT AAG GCG GAG TTT GCA GCT AGA GCG ATA GCT AGC GCA					
								Pro Asp Gln Arg Gly Ser Ala Ala Lys Ala Glu Phe Ala Ala Arg Ala Ile Ala Ser Ala					
	310	320	330	340	350	360		1330	1340	1350	1360	1370	1380
CCG CCG CCG CAT AGT CAA TTG CCA ATT TTT GCG CAC TCA CGT GGA ACC TAA TTC AGG GTA								CAT GGT ATA GAG CTC CCG CCT TTT CGA AAT GGT AAC GTT TCG GAT ATA GAA GCG ATG CTC					
								His Gly Ile Glu Leu Pro Pro Phe Arg Asn Gly Asn Val Ser Asp Ile Glu Ala Met Leu					
	370	380	390	400	410	420		1390	1400	1410	1420	1430	1440
AAT GCC ACA CAG CTC AAG CAA ACT ACA CAG CAC AAC ATA TTA GCG TTT ATG TGG TGG TTT								Eco RI	Hind III				
								AGC GGA GAG GAG GAA TTC GTA GAA AAA TAC AGA AGC TTG CTA GAT TCT GAT TGC TTT TAA					
								Ser Gly Glu Glu Glu Phe Val Glu Lys Tyr Arg Ser Leu Leu Asp Ser Asp Cys Phe					
	430	440	450	460	470	480		1450	1460	1470	1480	1490	1500
AAC ATA CTT AAG TGT GTT GGC ATT TAA TGT ACA GCC AAA ACG AGG TAA TTA TTC ATG GGC								ATG TAT AGG ACC CTC ATA GTT TCC ATA CAG AGC AAT CCG CCG CAA ATA GGC ACC CTC TGC					
								Met Gly					
	490	500	510	520	530	540		1510	1520	1530	1540	1550	1560
TGC GTC TCG TCA AAA AGC ACC ACA GTG CTT TCT CCA CAG ACA TCT TTT AAT GAA GGC TCC								Bgl I					
								GTG ATG CCG ATG CCG CGT TGA GGC TGG GGA AAC GGT TTT TTC CAC GTA AGC GTC CCG CCA					
								Cys Val Ser Ser Lys Ser Thr Thr Val Leu Ser Pro Gln Thr Ser Phe Asn Glu Ala Ser					
	550	560	570	580	590	600		1570	1580	1590	1600	1610	1620
CGT ACG TCT TTC AGA GCA CTC CCG GCG CCA TCG CAA AGA CAA TTG GAG GTC TAT GAT CAA								AGT CCG CTT CCG AAA CTT GAT GGC CAT TCA GCG GGT GAG TTC AGG TCG CCG AGG GTG CTG					
								Arg Thr Ser Phe Arg Ala Leu Pro Gly Pro Ser Gln Arg Gln Leu Glu Val Tyr Asp Gln					
	610	620	630	640	650	660		1630	1640	1650	1660	1670	1680
TGC TTA ATT GGT GCA GCG CCG TGG OCT GAC GAT TCC AGT AAG TCG AAT ACG OCT GAA AAC								GCG CCG TGG TAT CCG GAG TTG AAG CCG ACG GCA ACG CAC TGT TCA ATA TCG TGC CCG CAA					
								Cys Leu Ile Gly Ala Ala Arg Tyr Pro Asp Ser Ser Lys Ser Asn Thr Pro Glu Asn					
	670	680	690	700	710	720		1690	1700	1710	1720	1730	1740
AGG GCA TAT TGT CAG AGC ATG TAC AAC TCA ATT CCG TCT GCT GGA GAT GAA ATT TCC AGA								CGA TGA AGT ACT TTG TGA CCG GCA AGG TGG GTT GTT CGA GTG AGG GCG ATC AAC CTC AGG					
								Arg Ala Tyr Cys Gln Ser Met Tyr Asn Ser Ile Arg Ser Ala Gly Asp Glu Ile Ser Arg					
	730	740	750	760	770	780		1750	1760	1770	1780	1790	1800
GGT GGA ATC ACA TCT TTT GAG GAA CTA TGG GCG CCG GCA ACT GAA TGG CGA CTT TCA AAG								GCG GGA TTT CCG GGA TTT CGA CTG CCG CTG TGG AAT TGC ACG ACC CGT ACC CCG CAA ACC					
								Gly Gly Ile Thr Ser Phe Glu Glu Leu Trp Gly Arg Ala Thr Glu Trp Arg Leu Ser Lys					
	790	800	810	820	830	840		1810	1820	1830	1840	1850	1860
TTG CAG AGA GGA GAG CCG TTA TAC TCT GCA TTC GCG TCC GAA AGG ACG TCC GAT ACA GAC								GTT GAA AGT GCG AGT GCG GCG AAG GCA GAA CTT GCG ACT TGG CCG GCT GGT ACT GGA TCT					
								Leu Gln Arg Gly Glu Pro Leu Tyr Ser Ala Phe Ala Ser Glu Arg Thr Ser Asp Thr Asp					
	850	860	870	880	890	900		1870	1880	1890	1900	1910	1920
GCA GTA ACC CCT CTG GTC AAA CCT TAC AAG TCT GTC CTT GCG AGA GTC GTT GAT CAC GAG								TCA ACA GCG CTT GCT GCG CCG TAT CCG ACT CAA GCG TGA ACA GCG ACT GAA TGT GTT TTA					
								Ala Val Thr Pro Leu Val Lys Pro Tyr Lys Ser Val Leu Ala Arg Val Val Asp His Glu					
	910	920	930	940	950	960		1930	1940	1950	1960	1970	1980
GAT GCG CAC GAT GAA ATA ATG CAG GAC AAT TTG TTT GCG GAT CTG AAT GTT AAA GTA TAT								CTG GGT TTG GCG AGC AAT GGA ACC AAA AAC CAA CGT AAG CCG TAA TCT AGC TCC GGT CAT					
								Asp Ala His Asp Glu Ile Met Gln Asp Asn Leu Phe Gly Asp Leu Asn Val Lys Val Tyr					
	970	980	990	1000	1010	1020	Nru I	1990	2000	2010	2020	2030	2040
CGC CAA ACA GCA TAC CTC CAT GGA AAT GTT ATT CCA CTT AAC ACT TTT CCG CTC GCG ACA								AGA TGG TGG TCT GGA TTA CGT AGG GCG AGC AGC TCA CCT CCG GCG ACC TCA TCC GGT ACC					
								Arg Gln Thr Ala Tyr Leu His Gly Asn Val Ile Pro Leu Asn Thr Phe Arg Val Ala Thr					

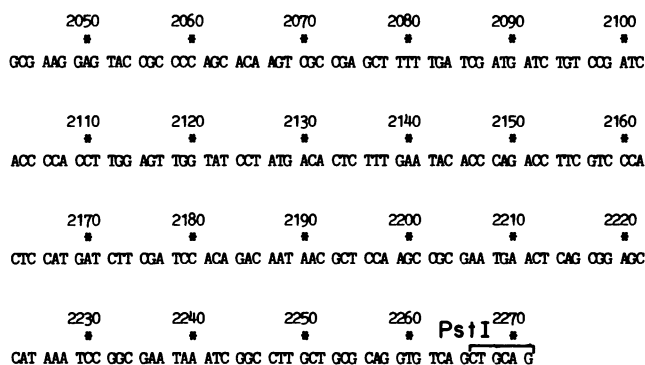


FIG. 1. Sequence of the *avrB* gene and its translated protein sequence. Selected restriction sites noted in the text are shown, and the Shine-Dalgarno box is underlined.

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 *Tn5* insertion near the *Nru*I site (Fig. 1) was observed to abolish the *avrB* phenotype (data not shown). A *Pst*I-*Hind*III 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*II 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 DH5 α 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.

10 20 30 40 50 60
 GAG CCG ATA GCC GAA CAC CCC TAT TTA TAT CGT TCT GCG CGA GCG CCA GCG ACT CCG GAG
 70 80 90 100 110 120
 CTG GTG GCT CAC CCA AAC TAG CTC CTG GTC TAT CGT GTC ACC CTC GAG CCG ATC GAG GTC
 130 140 150 160 170 180
 GTG AAC GTA ATC CAC GCG AGG CAG GAA TAC CCC TCG CCC TGA TTG CCG GTG TAT TTG CCG
 190 200 210 220 230 240
 GTA TTA AGC GGT TGC TTG CTA AGC GTC CGT AGC ACG ATC CTC TTT CCG ACA GCT CAG TCG
 250 260 270 280 290 300
 GCA TAG GAA TGA CAG GAA TGG AGT CAC TGG CCA TTA TTC TTA AGT CCG TTC GCT GTT CCA
 310 320 330 340 350 360
 TTG CTA CCC GGA CCC CAG GGT CAG GTC TGG GCG TTG CCT TGG AAC CGT TCT GCA ACT CGT
 370 380 390 400 410 420
 GGC ACT AAG CTT TGA CTA TTC GTG CAT GGT AGC ATG TAT GAA GGT AGA TAC TTT CTA GTC
 430 440 450 460 470 480
 ATA CGA TTA TGG GCG ATG ACC GCT TGA TGT GTA CTA CAT TCG GAA AAA TCA GCG GAA GTT
 490 500 510 520 530 540
 AGG GCT TAT TTC CTG TTG GAA ACG CTT GCT GTA CAG TTC TTC CAC GTC CTG CAA TGC CAC
 550 560 570 580 590 600
 CTA GAA TAA CAA TAC TTC ATA TTT TAT ATC TAT TGA AAC AGA GGT TAA AAA ATG GGA AAT
 Met Gly Asn
 610 620 630 640 650 660
 GTT TGT TTC CCG CCT AGT AGA AGC CAC GTT TGG CAA GAA TTT TCT CAA TCA GAA TTT TCC
 Val Cys Phe Arg Pro Ser Arg Ser His Val Ser Gln Glu Phe Ser Gln Ser Glu Phe Ser
 670 680 690 700 710 720
 ACA GCC AGT CCA GTC AGG ACA TCT GAA CGA CCC TGG GAT GCA TCA CTA GAT GCT GCG CTA
 Thr Ala Ser Pro Val Arg Thr Ser Glu Arg Pro Ser Asp Ala Ser Leu Asp Ala Gly Leu
 730 740 750 760 770 780
 GAA AGC TGG AGT GCT TGT CAC AGA AGC GGC CTG CCG GGT CCT GCG AAG CAT TCC ATG CTC
 Glu Ser Ser Ser Ala Cys His Arg Ser Gly Leu Arg Gly Pro Ala Lys His Ser Met Leu
 790 800 810 820 830 840
 AGT TTA GAC GAA ATT GGC CTA GTC GGT GCT GCG CCG TGG CCA GAT GAT GCG CCG GCG TTA
 Ser Leu Asp Glu Ile Gly Leu Val Gly Ala Ala Arg Trp Pro Asp Asp Ala Pro Gly Leu
 850 860 870 880 890 900
 AAT ATT TCC AAC AAA AGC AAT ACG CAA GAA AAT AAG CGA TAC TGT GAA AGC TTA TAT CAA
 Asn Ile Ser Asn Lys Ser Asn Thr Gln Glu Asn Lys Arg Tyr Cys Glu Ser Leu Tyr Gln
 910 920 930 940 950 960
 GCA GCA CGA ATT GCT GGT GGC TCC ATA GCG TCT GGC AGA GTT ACT AGT TTC GAT GCG CTT
 Ala Ala Arg Ile Ala Gly Gly Ser Ile Ala Ser Gly Arg Val Thr Ser Phe Asp Gly Leu
 970 980 990 1000 1010 1020
 TGG CGA AAC CGA ACA AAA TGG CCG TTA TCT AGA ATT CTT TGG GCG GAT GCG TCA AAA ATC
 Trp Arg Asn Ala Thr Lys Trp Arg Leu Ser Arg Ile Leu Ser Gly Asp Ala Ser Lys Ile

1030 1040 1050 1060 1070 1080
 GAC TTC GCT ACT GTT CCG ATG CCC AAT ACC AGA TTC GTA ACT TCT TTG AGG CGA CCG TAT
 Asp Phe Ala Thr Val Arg Met Pro Asn Thr Arg Phe Val Thr Ser Leu Arg Arg Pro Tyr
 1090 1100 1110 1120 1130 1140
 CAC TCA GTA ATT GAG CGA GTT AGA AAT CAT TCT GAT GCA AAT TGG GAA ATA TAC GAA GGA
 His Ser Val Ile Glu Arg Val Arg Asn His Ser Asp Ala Asn Ser Glu Ile Tyr Glu Gly
 1150 1160 1170 1180 1190 1200
 GAA TAT CTA GCG GGA ATT GAG ACC AAG GTC TAT CGT CAG CAT GCG ACG ATT TCA AGT ACA
 Glu Tyr Leu Gly Gly Ile Glu Thr Lys Val Tyr Arg Gln His Gly Thr Ile Ser Ser Thr
 1210 1220 1230 1240 1250 1260
 ACT ATT CCG ATG ACA ATA GTA AGT GCA CCG GAT GAC GAT GAT ATA CAT GAA AGG TTA
 Thr Ile Pro Met Thr Ile Val Ser Ala Val Ala Asp Asp Asp Ile His Glu Arg Leu
 1270 1280 1290 1300 1310 1320
 AAG AGC CTG CCA AAG AAT GAG CCG CCG CAC CTG AAA GAT TTG ATG GCG GCG TCA CAC CCT
 Lys Ser Leu Pro Lys Asn Glu Arg Arg His Leu Lys Asp Leu Met Ala Ala Ser His Pro
 1330 1340 1350 1360 1370 1380
 AAC ATG ATC ACA CAC ACT GAT GCA GTA TAT CTT CCA ATG ATC AAG GAT CAT TTA GAA TCA
 Asn Met Ile Thr His Thr Asp Ala Val Tyr Leu Pro Met Ile Lys Asp His Leu Glu Ser
 1390 1400 1410 1420 1430 1440
 TTA TAT TTG CAA GCG ATA GAC CCT TGG CTT GAG CAG CAC GAG GCG CTC GAG TTG ATC GGT
 Leu Tyr Leu Gln Ala Ile Asp Pro Ser Leu Glu Gln His Glu Ala Leu Glu Leu Ile Ala
 1450 1460 1470 1480 1490 1500
 CCG ATA CAC TGG TGG GCT GCA AGT GCA GCA CCA GAT AGG CGT GCG AGT GCT GCG AAA CCA
 Arg Ile His Trp Trp Ala Ala Ser Ala Ala Pro Asp Arg Arg Gly Ser Ala Ala Lys Ala
 1510 1520 1530 1540 1550 1560
 GAG TTC GCG CGA AGA TCA ATC CCG TTC CCT CAT GGT ATT GAA CTA CCG CCA TTC GAG CAT
 Glu Phe Ala Ala Arg Ser Ile Ala Phe Ala His Gly Ile Glu Leu Pro Pro Phe Glu His
 1570 1580 1590 1600 1610 1620
 GGT GCA GTT CCT GAT ATT GAA GCA ATG CTC AGA TCT GAA GAG CAA TTT GTA GAA GAT TAC
 Gly Ala Val Pro Asp Ile Glu Ala Met Leu Arg Ser Glu Glu Gln Phe Val Glu Asp Tyr
 1630 1640 1650 1660 1670 1680
 CCT AAC CTT TTT GAG CCG CCC CCT CAG TAA CAA GCA ACA GCG GAC ATT GAG GCA CTG ATG
 Pro Asn Leu Phe Glu Arg Pro Pro Gln ———
 1690 1700 1710 1720 1730 1740
 TCA CGA TGC CCG CTA GCG GAA TGG TTT CTG GCG GAC CCT GCT AAG GGT GAA GTC GTA AAA
 Ser Arg Cys Arg Leu Ala Glu Ser Phe Leu Ala Asp Pro Ala Lys Gly Glu Val Val Lys
 1750 1760 1770 1780 1790 1800
 TTC GTT CGT CTC GCG GAA GCG CAC ATT ATT GCT GCT ACA AAT TTC GTG TTC ACG AAT CAC
 Phe Val Arg Leu Ala Glu Ala His Ile Ile Ala Ala Thr Asn Phe Val Phe Thr Asn His
 1810 1820 1830 1840 1850 1860
 CAC CTG CCT GTT CAC ACA GAG TTC CTT CGT AGG GTG GCT CCG CCG CCG CAG CCG AAA TCA
 His Leu Pro Val His Thr Glu Phe Leu Arg Arg Val Ala Arg Arg Pro Gln Arg Lys Ser
 1870 1880 1890 1900 1910 1920
 CCG GCG AAA ATT TTC ACG ACC AAT TAC GAC CCG TGC TTC GAA GAA GCA GGT CCG CAA CGA
 Arg Ala Lys Ile Phe Thr Thr Asn Tyr Asp Arg Cys Phe Glu Glu Ala Gly Arg Gln Gly
 1930 1940 1950 1960 1970 1980
 CGA TAT GTT GTG GTC GAT GGT TTT TCC CAT AOC GCG CCC CCG AOC TTC GAT GCG GTT CAC
 Arg Tyr Val Val Val Asp Gly Phe Ser His Thr Ala Pro Pro Thr Phe Asp Ala Val His
 1990 2000 2010 2020 2030 2040
 TTC AGC TAT GAG ACC GTG ACA CGA CTG GCT GAC ACT GAA GCT TTC GAC CTA ATT CCA AAC
 Phe Ser Tyr Glu Thr Val Thr Arg Leu Ala Asp Thr Glu Ala Phe Asp Leu Ile Pro Asn

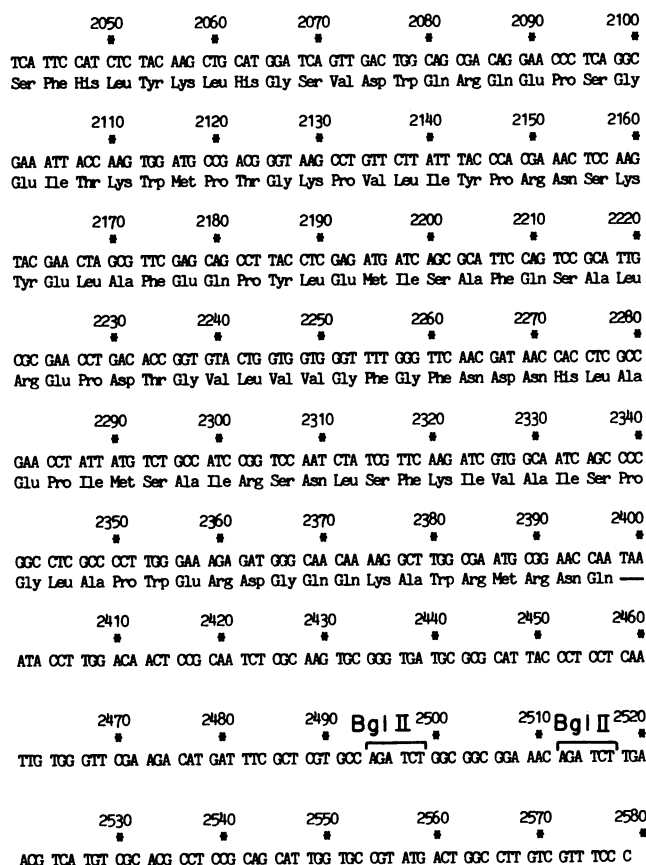


FIG. 2. Sequence of the *avrC* gene (ORF 1) and its deduced protein sequence. The translated amino acid sequence of ORF 2 initiating at position 1678 is also shown. Selected restriction endonuclease sites are shown, and Shine-Dalgarno sequences are underlined.

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 *Hind*III 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.

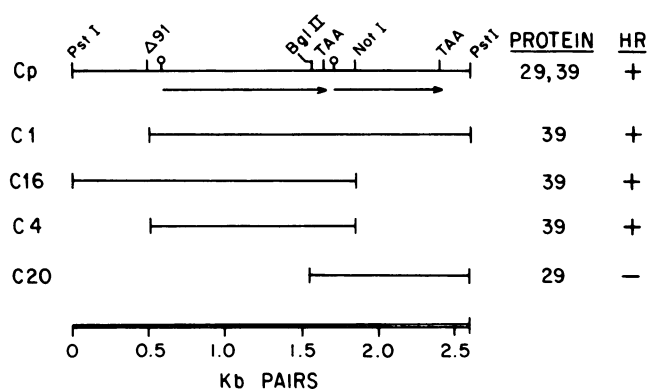


FIG. 3. Various subclones of the insert DNA of pPSG0101, carrying *avrC*. Fragments were recloned into pUC119, tested for protein synthesis in *E. coli* by gel electrophoresis, and shuttled to pRK415 or pDSK519. The resulting plasmids were transferred to race 4 of *P. syringae* pv. *glycinea* and tested 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 ($\Delta 91$) (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 *PstI* 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

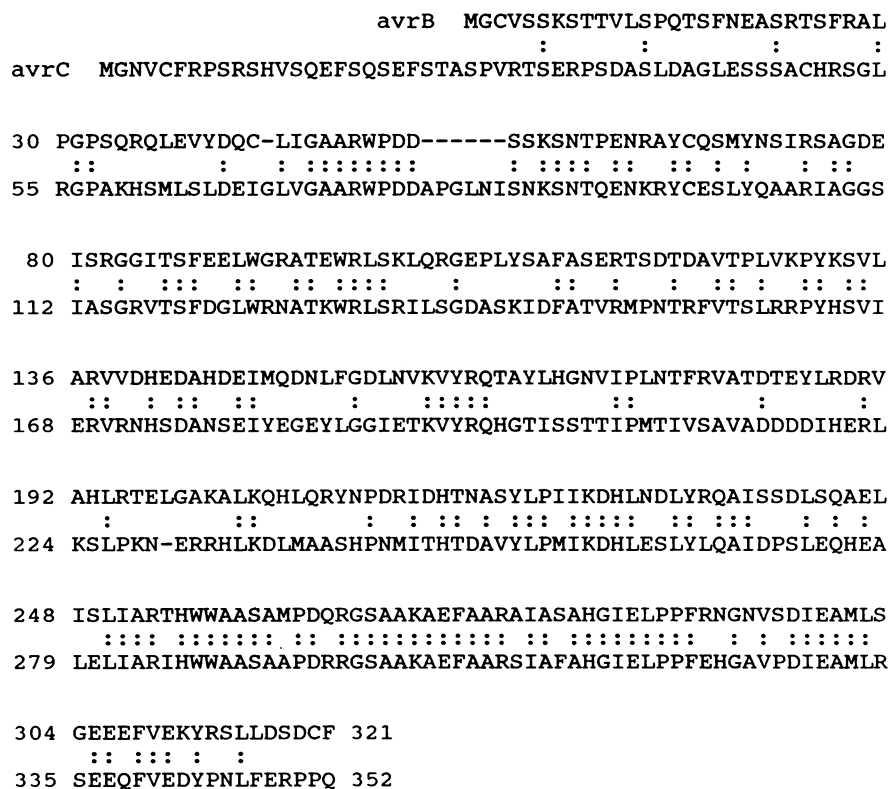


FIG. 4. Homology of the protein products of *avrB* and *avrC*. Connecting dots denote identical amino acids.

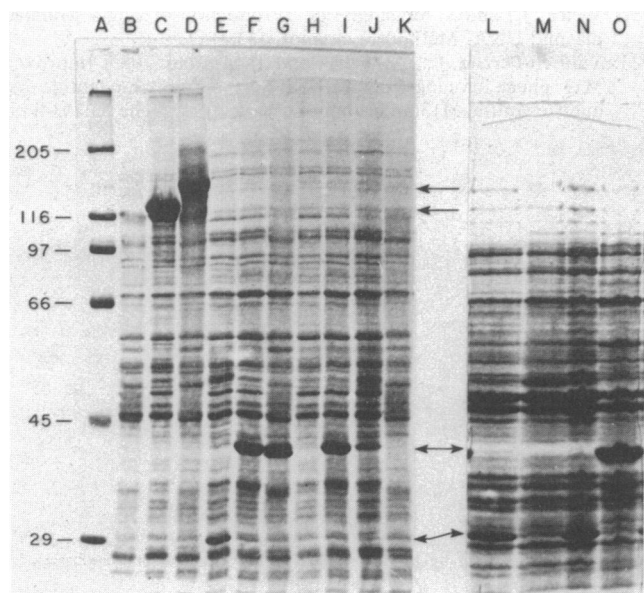


FIG. 5. Sodium dodecyl sulfate-polyacrylamide gel of *E. coli* JM-109 cells containing various plasmid constructs of the *avrC* gene (ORF 1) or ORF 2 (Fig. 2). All cells were grown on LB medium overnight at 28°C in the presence of 1 mM IPTG unless otherwise noted. Lanes: A, molecular weight standards 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. Depart-

ment of Agriculture grant 86-CR-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 multi-gene 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 glycoellin 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: 10⁴-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. Pestell, 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. R  ther, U., and B. M  ller-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.