INFECTION AND IMMUNITY, Aug. 1990, p. 2651–2658
0019-9567/90/082651-08$02.00/0

Vol. 58, No. 8

Genetic and DNA Sequence Analysis of the *Salmonella typhimurium* Virulence Plasmid Gene Encoding the 28,000-Molecular-Weight Protein

PAUL A. GULIG* AND VINCE A. CHIODO

Department of Immunology and Medical Microbiology, University of Florida College of Medicine, Gainesville, Florida 32610

Received 2 February 1990/Accepted 14 May 1990

We have confirmed that the 28,000-molecular-weight (28K) protein encoded by the *virA* gene of the 90-kilobase *Salmonella typhimurium* virulence plasmid is a virulence factor. It was previously shown that a Tn*5* insertion, *vir-22*::Tn*5*, located in the virulence plasmid greatly attenuated virulence for mice and inhibited the production of a 28K protein (P. A. Gulig and R. Curtiss III, Infect. Immun. 56:3262–3271, 1988). Plasmid pYA426 fully complemented *vir-22*::Tn*5* to virulence by increasing splenic infection after oral inoculation and encoded the 28K protein. To identify the virulence gene(s) of pYA426 mutated by *vir-22*::Tn*5*, we constructed nested deletions in pYA426 and examined deletion derivatives for their abilities to complement *vir-22*::Tn*5*. Only derivatives still producing the 28K protein complemented *vir-22*::Tn*5*. Furthermore, the smallest complementing derivative encoded only the 28K protein, as determined by DNA sequence analysis. Therefore, the 28K protein is sufficient for complementation of the attenuating mutation *vir-22*::Tn*5* and must be the virulence factor inhibited by the insertion. We determined the nucleotide sequence of the 1.2-kilobase *Bam*HI-*Eco*RI fragment encoding the 28K protein and identified the structural gene, *virA*. A 723-base-pair open reading frame which encodes a peptide with a molecular weight of 27,572 was found.

The role of virulence plasmids in the pathogenesis of infection with several *Salmonella* species and serotypes has been the subject of considerable research during this decade (for a review, see reference 12a). A virulence plasmid is necessary for salmonellae to cause infection beyond the Peyer's patches of the intestines to the mesenteric lymph nodes and spleen after oral inoculation of mice (13, 26). Although the contribution of the plasmid to virulence at the cellular or molecular levels has not been determined, the locations of several virulence genes and homologous regions of different plasmids have been identified. Deletion and transposon insertion mutagenesis (2, 3, 14, 23, 30) and cloning (14, 24) have been used to identify regions associated with virulence. Through restriction mapping of attenuating mutations and cloned virulence sequences, a common virulence region has been identified (2, 3, 12a, 14, 23, 24, 28, 30, 36). Most recently, DNA sequence analysis has been reported for virulence plasmid genes (25, 32).

Gulig and Curtiss (14) used cloning of the virulence plasmid into plasmid-cured *Salmonella typhimurium* to select for recombinant plasmids that encoded critical virulence genes. From virulence-conferring clones, a consensus virulence region of 14 kilobases (kb) was identified. This region coincided with virulence regions identified by others (12a). A Tn*5* insertion, *vir-22*::Tn*5*, in the virulence region greatly attenuated virulence in mice in terms of splenic infection after oral inoculation and inhibited the production of a 28,000-molecular-weight (28K) protein. Plasmid pYA422 contained the 3.2-kb *Bam*HI fragment corresponding to the site of insertion of *vir-22*::Tn*5*, expressed only the 28K protein, and conferred virulence to plasmid-cured *S. typhimurium* in terms of 50% lethal dose. These results suggested that the 28K protein was an important plasmid-encoded virulence factor. Plasmid pYA426 contained the same 3.2-kb *Bam*HI fragment, also produced the 28K protein, and fully complemented *vir-22*::Tn*5*. However, the identity of the virulence gene(s) mutated by *vir-22*::Tn*5* was left in question because of the large size of the *Bam*HI fragment and the possibility that *vir-22*::Tn*5* exerted multigenic (i.e., polar) effects.

In this report, we demonstrate that the 28K protein alone is sufficient to complement *vir-22*::Tn*5* and therefore must be the virulence factor inhibited by this mutation. The DNA sequence of the gene *virA* encoding the 28K protein was determined. The nucleotide sequence data reported in this paper will appear in the EMBL, GenBank, and DDBJ nucleotide sequence data bases under the accession number M33927.

(A preliminary report of these data was presented at the 1989 meeting of the American Society for Microbiology [P. A. Gulig and V. A. Chiodo, Abstr. Annu. Meet. Am. Soc. Microbiol. 1989, B217, p. 66].)

MATERIALS AND METHODS

Bacterial strains. The bacterial strains used in this study are described in Table 1. The growth medium was L broth or L agar (20). Antibiotics were used at the following concentrations (in micrograms per milliliter): ampicillin, 100; chloramphenicol, 30; kanamycin, 40; nalidixic acid, 50; and tetracycline, 25.

Genetic techniques. Transformation was performed as described by Humphreys et al. (17). Plasmid extraction was done by the procedure of Birnboim (7). Enzymes for manipulation of DNA were purchased from Bethesda Research Laboratories, Inc., Gaithersburg, Md., or Promega Corp.,

* Corresponding author.

TABLE 1. Bacterial strains

Strain	Genotype	Comments or reference
S. typhimurium SR-11		
χ3306	*gyrA1816* pStSR100$^+$	Virulence plasmid positive, nalidixic acid resistant, virulent (13)
χ3337	*gyrA1816* pStSR100$^-$	Virulence plasmid-negative derivative of χ3306, nalidixic acid resistant (13)
χ3589	*gyrA1816* pStSR100*vir-22*::Tn*5*	*vir-22*::Tn*5* derivative of χ3306, nalidixic acid resistant, kanamycin resistant, attenuated (14)
S. typhimurium LT2-Z		
χ3477	*hsdL6* Δ(*galE-uvrB*)*1005 flaA66 rpsL20 xyl-404 lamB*$^+$ (*E. coli*) Δ(*zja*::Tn*10*) *hsdSA29*	14
E. coli K-12		
LE392	*lacY1 glnV44 galK2 tyrT58 metB1 hsdR514 trpR55*	22
χ925	*thr-1 ara-13 leu-6 azi-8 tonA2 lacY1 minA1 glnV44 gal-6 minB2 rpsL135 xyl-7 mtl-2 thi-1*	Minicell producer (12)

Madison, Wis., and used in accordance with manufacturer instructions.

Deletion mutagenesis of pYA426. Unidirectional deletions were generated in pYA426 with exonuclease III (Bethesda Research Laboratories) exactly as described previously (16). Briefly, pYA426 was linearized with *Sph*I to generate a 3′ cohesive end within the tetracycline resistance gene of pACYC184 and was then digested with *Xho*I to create a 5′ cohesive end within the right end of the insert sequence (Fig. 1). Exonuclease III was added to begin digestion into the 5′ end of the *Xho*I site. At various times thereafter, the reaction was stopped, and the resulting single-stranded DNA was digested with S1 nuclease and filled in with the Klenow fragment of DNA polymerase I. The plasmid DNA was circularized with T4 DNA ligase and transformed into *Escherichia coli* LE392. Clones were examined by minilysate analysis (7). To confirm that digestion had proceeded only into the insert sequence at the *Xho*I site, we digested the plasmids with *Sal*I, which cuts 89 base pairs beyond the *Sph*I site (Fig. 1). Deletion derivatives were further mapped with *Bam*HI, *Eco*RI, and *Sal*I.

To create a plasmid containing precisely the 1.2-kb *Bam*HI-*Eco*RI fragment in the same orientation in pACYC184 as pYA426, we constructed pGTR127 as follows. pYA426 was digested with *Eco*RI, and the cohesive ends were filled in with the Klenow fragment of DNA polymerase I. DNA was digested with *Bam*HI, and the 1.2-kb *Bam*HI-blunt-ended fragment was extracted from a

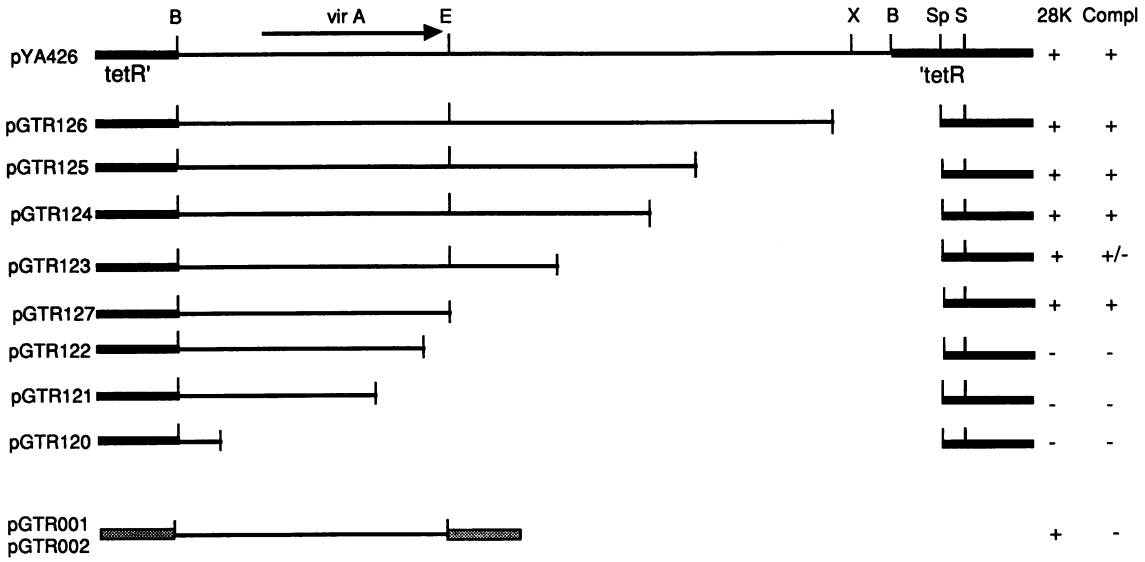

FIG. 1. Deletion derivatives of pYA426 and inserts of pGTR001 and pGTR002. The *vir-22*::Tn*5*-complementing plasmid pYA426 was digested with *Xho*I and *Sph*I, resulting in 5′ and 3′ cohesive ends, respectively. Exonuclease III digestion resulted in the generation of deletions into the *Xho*I site. The relevant portion of pYA426 is shown as follows: the solid bars represent the tetracycline resistance gene of pACYC184, and the solid line represents the 3.2-kb *Bam*HI insert from the virulence plasmid. The sequences remaining in the deletion derivatives are represented by solid lines. The stippled bars for pGTR001 and pGTR002 represent pUC19 and pUC18 sequences, respectively. Abbreviations used for restriction sites: B, *Bam*HI; E, *Eco*RI; X, *Xho*I; Sp, *Sph*I; S, *Sal*I. The column labeled 28K shows the production of the 28K protein in *E. coli* minicells (summarized from Fig. 2). The column labeled Compl shows the complementation of *vir-22*::Tn*5* (summarized from Table 2); +/−, intermediate complementation relative to positive and negative controls.

preparative agarose gel (4). pACYC184 was digested with *Sal*I, and the cohesive ends were filled in with the Klenow fragment of DNA polymerase I. The *Sal*I-blunt-ended DNA was digested with *Bam*HI and ligated with the 1.2-kb fragment from pYA426.

Abilities of plasmids to complement *vir-22*::Tn*5*. Plasmids were transformed into *S. typhimurium* χ3589, which contains the *vir-22*::Tn*5* insertion in the virulence plasmid, with *S. typhimurium* χ3477 as an intermediate host strain for plasmids constructed in *E. coli* (14). Mice were infected orally with 2.6×10^8 to 7.4×10^8 CFU as described previously (13), and splenic CFU were enumerated 5 days later. Spleens were homogenized in 10 ml of phosphate-buffered saline containing 0.1% (wt/vol) gelatin (10) with a Stomacher (model 80; Tekmar Co., Cincinnati, Ohio), and dilutions were plated on media containing appropriate antibiotics. The complementing plasmid, pYA426, and the vector, pACYC184, were included as positive and negative controls, respectively.

DNA sequence analysis. The nucleotide sequence of plasmid DNA was determined on both strands by the dideoxy nucleotide chain termination reaction with double-stranded DNA (Sequenase; US Biochemical Corp., Cleveland, Ohio) and [α-^{35}S]dATP (specific activity, >1,000 Ci/mmol; Amersham Corp., Arlington Heights, Ill.). The 1.2-kb *Bam*HI-*Eco*RI fragment was subcloned from pYA422 into pUC19 and pUC18 to yield plasmids pGTR001 and pGTR002, respectively (Fig. 1). In addition to the "forward M13 primer" supplied by the manufacturer, oligonucleotides were synthesized by the University of Florida Interdisciplinary Center for Biotechnology Research DNA Core Laboratory for use as "walking primers." A "reverse M13 primer" (ATGT GTGGAATTGTGAGCGG) was kindly provided by Henry Baker. Deletions were constructed in pGTR001 with exonuclease III as described above or with restriction enzyme digestion. DNA fragments were also subcloned into pUC19 for sequence analysis with the forward and reverse primers.

To sequence the junctions of exonuclease III-generated deletions of pYA426, a primer was synthesized complementary to the bases immediately adjacent to the unique *Sal*I site of pACYC184 (31) (GGGCATCGGTCGACG). An additional primer (CGTTACCATGTTAGG) based on the sequence of IS*50* of Tn*5* (1) was constructed so that the junctions of previously isolated Tn*5* insertions in virulence plasmid sequences could be identified. pYA403 mutants containing different Tn*5* insertions (14) were digested with *Bam*HI, and the resulting fragments were subcloned into pUC19, selecting for kanamycin resistance. In this manner, the half of Tn*5* containing the kanamycin resistance gene (5) and the flanking virulence plasmid sequence to the proximal *Bam*HI site were isolated. The Tn*5* primer was used with the subclone in DNA sequencing. Similarly, the junctions of Tn*5* insertions in pYA422 (14) were sequenced by deleting one-half of the Tn*5* element and the flanking virulence plasmid sequence by digestion with either *Sal*I or *Hin*dIII and religation. The single remaining Tn*5*-virulence plasmid junction was sequenced with the Tn*5* primer.

The DNA sequences were analyzed with programs of the Genetics Computer Group of the University of Wisconsin (11) on a VAX computer at the Institute for Food and Agricultural Science of the University of Florida. Both strands were examined for open reading frames and restriction enzyme sites. The amino acid sequence of the 28K protein was analyzed for hydrophobicity-hydrophilicity with the Kyte-Doolittle algorithm (19) and for secondary structure with the Chou-Fasman algorithm (9) by use of the Genetics Computer Group programs Peptidestructure and Plotstructure (18). Nucleic acid and amino acid homologies with the GenBank and EMBL data bases were examined with the FastN, FastP, and TFastA programs (21, 27).

Analysis of plasmid-encoded proteins with *E. coli* minicells. Plasmid-encoded proteins were examined with *E. coli* minicells as described previously (14), with modifications as noted. Static overnight cultures of *E. coli* K-12 χ925 containing the appropriate plasmid were grown in minicell-labeling medium (14), which consisted of medium M9 (22) with all of the amino acids except methionine and containing thiamine hydrochloride. On the next day, these cultures were diluted 1:20 into 75 ml of prewarmed minicell-labeling medium and grown with shaking at 37°C to the late log phase (optical density at 600 nm, approximately 0.4). The use of log-phase cultures replaced the previous use of stationary-phase cultures harvested after overnight growth with shaking. Minicells were isolated by differential centrifugation and two successive sucrose gradient centrifugation steps instead of a single sucrose gradient centrifugation step. Plasmid-encoded proteins were labeled by incubation of minicells in minicell-labeling medium containing [^{35}S]methionine (specific activity, 1,000 Ci/mmol; Amersham). Proteins were resolved by sodium dodecyl sulfate-polyacrylamide gel electrophoresis, and radiolabeled proteins were identified by fluorography with 1 M sodium salicylate (8) in the place of En^{3}Hance (New England Nuclear Corp., Boston, Mass.).

RESULTS

Identification of the minimal DNA sequence of pYA426 required to complement *vir-22*::Tn*5*. To identify the virulence plasmid gene(s) encoded by pYA426 corresponding to that mutated by *vir-22*::Tn*5* present in the virulence plasmid of *S. typhimurium* χ3589, we constructed deletion derivatives of pYA426 to identify the minimal sequence necessary for complementation. It was previously reported that *vir-22*::Tn*5*, which inhibits the production of the 28K protein, mapped near the *Bam*HI site of a 1.2-kb *Bam*HI-*Eco*RI fragment comprising the left end of the insert sequence of pYA426 (14). Furthermore, *vir-26*::Tn*5* and *vir-27*::Tn*5*, which do not affect the production of the 28K protein, mapped to the right of the *Eco*RI site (14). These facts suggested that the 28K protein was encoded in the left sequences of pYA426, possibly within the *Bam*HI-*Eco*RI fragment. To examine this possibility, we subcloned the fragment into pUC19 and pUC18, yielding pGTR001 and pGTR002, respectively (Fig. 1). The subclones, which possessed the *Bam*HI-*Eco*RI fragment in opposite orientations relative to the *lacZ* promoter, expressed the 28K protein in *E. coli* minicells (data not shown), confirming the location of the structural gene for the 28K protein and the presence of the promoter for the gene within the fragment. pGTR001 and pGTR002 were also used for DNA sequence analysis.

Because the gene encoding the 28K protein was located within the left 1.2 kb of the 3.2-kb insert of pYA426, we constructed nested deletions into the insert from the right end. Exonuclease III was used to unidirectionally digest into the 5′ end of *Xho*I left after *Xho*I-*Sph*I digestion. We isolated deletions with remaining insert sizes of 3.0 kb, well outside the region encoding the 28K protein, to 0.2 kb, well within the region encoding the 28K protein (Fig. 1). The deletion derivatives were mapped by restriction enzyme digestion and transformed into *S. typhimurium* χ3589. Complementation of *vir-22*::Tn*5* was examined by determining levels of splenic infection after oral inoculation of mice (Table 2).

TABLE 2. Complementation of *vir-22*::Tn*5* by deletion derivatives of pYA426[a]

Strain	Log_{10} CFU/spleen (mean ± SD)	*P* value compared with:	
		χ3589 (pACYC184)	χ3589 (pYA426)
χ3306	5.78 ± 0.42	<0.001	NS
χ3337	3.11 ± 0.67	NS	<0.002
χ3589 (pACYC184)	2.54 ± 0.36		<0.001
χ3589 (pYA426)	5.60 ± 0.60	<0.0005	
χ3589 (pGTR120)	2.56 ± 0.38	NS	<0.001
χ3589 (pGTR121)	2.75 ± 0.57	NS	<0.001
χ3589 (pGTR122)	3.33 ± 1.03	NS	<0.001
χ3589 (pGTR127)	4.90 ± 0.66	<0.001	NS
χ3589 (pGTR123)	3.83 ± 0.35	<0.005	<0.005
χ3589 (pGTR124)	5.75 ± 1.41	<0.005	NS
χ3589 (pGTR125)	6.13 ± 1.68	<0.01	NS
χ3589 (pGTR126)	5.20 ± 0.78	<0.001	NS

[a] Groups of four mice were inoculated orally with 2.6×10^8 to 7.4×10^8 CFU of an *S. typhimurium* strain. Five days later, spleens were homogenized and splenic infection was quantitated. *P* values are for the two-tailed Student *t* test. NS, Not significant ($P > 0.05$).

Controls included wild-type χ3306, plasmid-cured χ3337, complementation of *vir-22*::Tn*5* with pYA426, and lack of complementation by pACYC184. Plasmids pGTR124, pGTR125, and pGTR126, in which the deletions did not extend into the 1.2-kb *Bam*HI-*Eco*RI fragment, fully complemented *vir-22*::Tn*5*. In contrast, plasmids pGTR120, pGTR121, and pGTR122, in which the deletions extended into the *Bam*HI-*Eco*RI fragment, did not complement *vir-22*::Tn*5*. The *Bam*HI-*Eco*RI subclone pGTR127 contained the shortest insert sequence capable of complementing *vir-22*::Tn*5*. All of the plasmids which complemented *vir-22*::Tn*5* produced the 28K protein, whereas the noncomplementing plasmids did not produce the 28K protein (Fig. 2). An exception was pGTR123, which consistently yielded intermediate complementation, although the deletion did not extend into the 1.2-kb *Bam*HI-*Eco*RI fragment and the 28K protein was produced. The *P* value was less than 0.005 for splenic CFU of χ3589 (pGTR123) when compared with both the positive and negative controls for complementation.

DNA sequence analysis of *virA* encoding the 28K protein. We determined the nucleotide sequence of the 1.2-kb *Bam*HI-*Eco*RI fragment with pGTR001, deletion derivatives, subclones, and synthetic oligonucleotide primers. Only a single complete open reading frame consisting of 723 base pairs was identified from nucleotides 393 to 1115 from the *Bam*HI site (Fig. 3). This open reading frame encodes a peptide with a deduced molecular weight of 27,572, in close agreement with the size of the 28K protein estimated from sodium dodecyl sulfate-polyacrylamide gel electrophoresis. On the basis of the correlation of the open reading frame with deletion analysis of pYA426 and expression of the 28K protein (Fig. 2 and 3), we determined that the 723-base-pair open reading frame encodes the 28K protein, and we named the gene *virA*. An open reading frame extends into the *Bam*HI-*Eco*RI fragment from bases 1 through 108. This open reading frame represents the 3′ end of a gene, *mkaA*, encoding a 70K protein identified by Taira and Rhen (32). Because *virA* was the only gene needed to complement *vir-22*::Tn*5*, the 28K protein must be a virulence factor encoded by the *S. typhimurium* virulence plasmid.

A consensus ribosome-binding site (AGGAG) was identified 8 base pairs upstream of the ATG start codon, which was located at base pair 393 from the *Bam*HI site (Fig. 3).

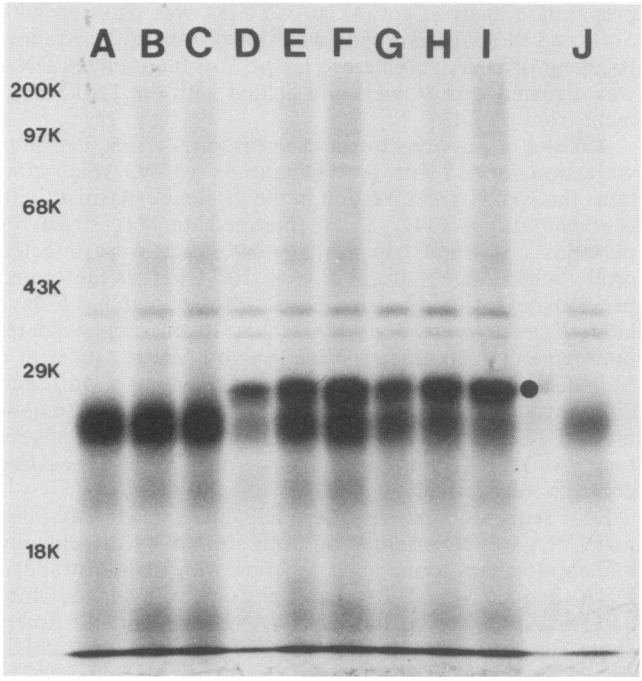

FIG. 2. Expression of the 28K protein by pYA426 and deletion derivatives in *E. coli* minicells. pYA426, deletion derivatives, and the vector pACYC184 were transformed into minicell-producing *E. coli* χ925. Plasmid-encoded proteins were radiolabeled with [^{35}S]methionine as described in Materials and Methods and resolved by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (10% [wt/vol] polyacrylamide). Lanes: A, pGTR120; B, pGTR121; C, pGTR122; D, pGTR127; E, pGTR123; F, pGTR124; G, pGTR125; H, pGTR126; I, pYA426; J, pACYC184. ●, 28K protein. The migration of ^{14}C-labeled molecular weight markers (Bethesda Research Laboratories) is shown on the left.

The only consensus promoter identified was 97 base pairs from *Bam*HI and was identified by a −10 region (AATAAT) at base 81 and a −35 region (CTGACA) at base 59. *vir-22*::Tn*5* was located at base 170 from *Bam*HI, between the putative promoter and start codon. Since *vir-22*::Tn*5* does not fall within the structural gene for the 28K protein, it is a polar mutation, giving credence to the above-stated assignment of −35 and −10 sequences identifying the promoter at base 97. The DNA and amino acid sequences of *virA* were examined for homology to sequences of the GenBank and EMBL data bases with the TFastA, FastN, and FastP programs (21, 27), and no obviously homologous genes or peptides were identified. The deduced amino acid sequence was further examined with the Kyte-Doolittle algorithm for hydropathy (19) and the Chou-Fasman algorithm for secondary structure (9). The 28K protein appears to be highly hydrophilic, with no stretches of hydrophobic sequences of more than five to six residues. No leaderlike sequence was apparent at the amino-terminal end. Therefore, the 28K protein is probably not an integral membrane protein, nor is it likely to be transported across the cytoplasmic membrane. The pI of the 28K protein was calculated to be 6.96 with the Isoelectric program of the Genetics Computer Group (11).

DNA sequence analysis of previously described Tn*5* insertions. As part of the initial characterization of the *S. typhimurium* SR-11 virulence plasmid, four Tn*5* insertions (*vir-1*::Tn*5*, *vir-3*::Tn*5*, *vir-4*::Tn*5*, and *vir-7*::Tn*5*) were iso-

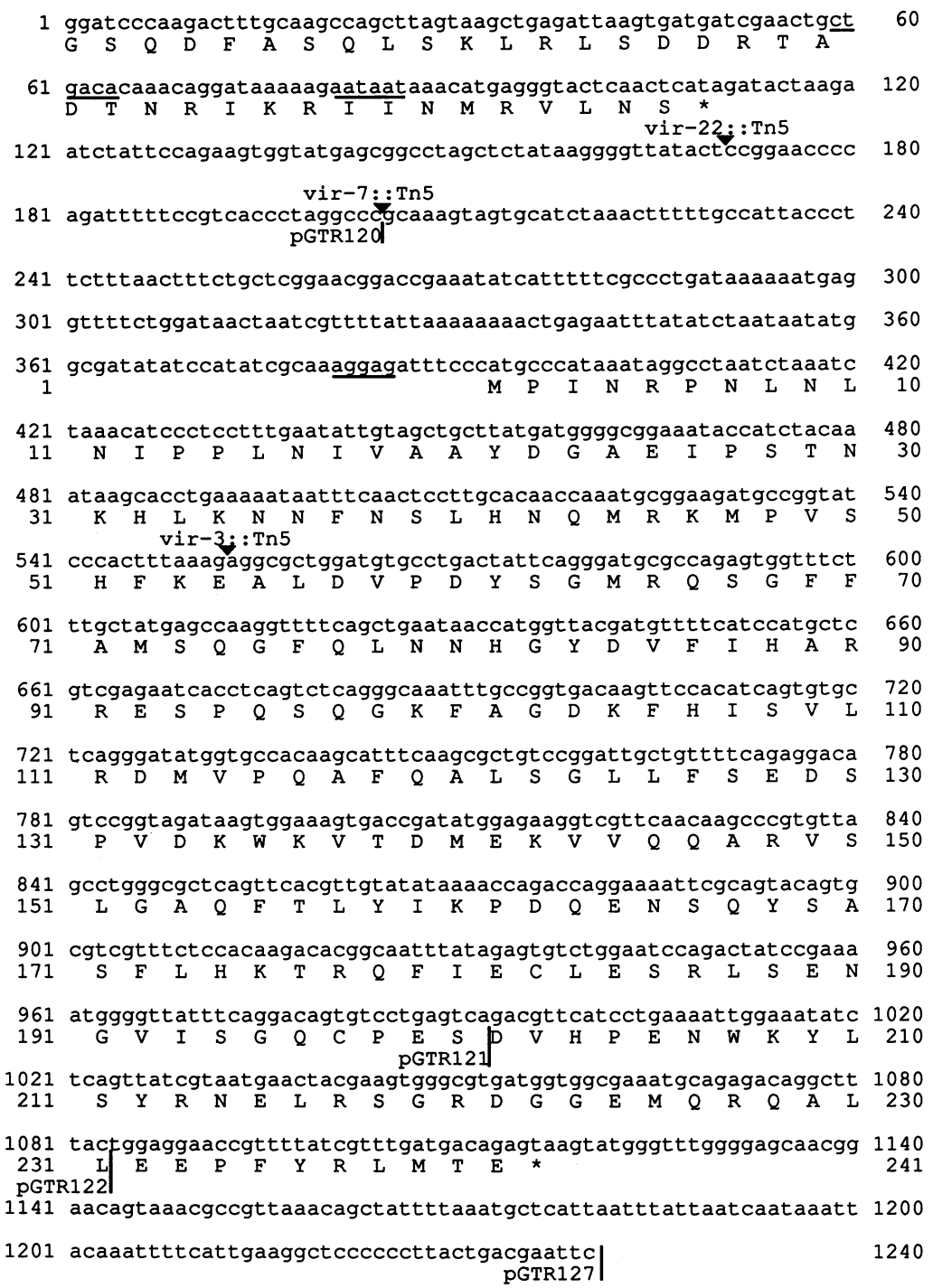

FIG. 3. DNA sequence of the 1.2-kb *Bam*HI-*Eco*RI fragment encoding the 28K protein and the deduced amino acid sequence of the 28K protein. The deduced amino acid sequences of the two major open reading frames are shown below the nucleotide sequence. *virA* encoding the 28K protein is located from nucleotides 393 to 1115. A consensus ribosome-binding site is underlined at base 381. Potential −10 and −35 regions for a promoter are underlined at bases 81 and 59, respectively. A second open reading frame from *mkaA* (32) runs from base 1 to base 108. The locations of Tn*5* insertions are shown above the DNA sequence (▼), and the sizes of the inserts for deletion derivatives are shown below the sequence (|). Stop codons for open reading frames are indicated by asterisks.

lated in the 3.2-kb *Bam*HI insert sequence of pYA422 (14). We therefore used DNA sequence analysis to identify the precise locations of two of these Tn*5* insertions originally identified in the 1.2-kb *Bam*HI-*Eco*RI fragment encoding the 28K protein. *vir-3*::Tn*5*, which was the only insertion that inhibited the production of the 28K protein, was located at base 552 of the 1.2-kb *Bam*HI-*Eco*RI fragment within the coding sequence of the 28K protein. *vir-7*::Tn*5*, which did not inhibit the production of the 28K protein, was located at base 204. This position for *vir-7*::Tn*5* represents a correction

from the originally reported location based on restriction enzyme mapping (14).

DISCUSSION

Molecular analysis of the genetics of plasmid-mediated virulence of *S. typhimurium* by Gulig and Curtiss (14) identified a Tn*5* insertion, *vir-22*::Tn*5*, which greatly attenuated virulence and inhibited the production of the 28K protein. However, pYA422, which expressed the 28K protein, did not complement *vir-22*::Tn*5* for splenic infection. When the insert sequence of pYA422 (a 3.2-kb *Bam*HI fragment) was inverted in the vector pACYC184, the resulting plasmid, pYA426, encoded the 28K protein and was able to complement *vir-22*::Tn*5* for splenic infection. This situation left some confusion as to the exact role of the 28K protein in the virulence of *S. typhimurium*. In the present study, we demonstrate that the 28K protein encoded by *virA* of pYA426 is necessary and sufficient to complement *vir-22*::Tn*5*. Therefore, the 28K protein is the virulence factor directly inhibited by *vir-22*::Tn*5* and must be a determinant for the plasmid-encoded pathogenicity of *S. typhimurium*.

An anomalous result obtained by complementation analysis of *vir-22*::Tn*5* with deletion derivatives of pYA426 was that plasmid pGTR123 incompletely complemented *vir-22*::Tn*5*, although the flanking deletion derivatives pGTR127 and pGTR124 completely complemented the Tn*5* insertion (Table 2 and Fig. 1). The lack of complete complementation could be due to the fact that the deletion of pGTR123 interrupts an open reading frame of at least 728 bases beginning 144 bases to the right of the *Eco*RI site and reading to the right (data not shown). This open reading frame is part of a locus, *virB*, that encodes a gene(s) essential for the plasmid-mediated virulence of *S. typhimurium* (P. A. Gulig, V. A. Chiodo, and A. L. Caldwell, unpublished data). Furthermore, the deletion in pGTR123 results in an in-frame fusion of the interrupted open reading frame to the tetracycline resistance gene of pACYC184. The fusion protein may be deleterious to the salmonella cells, resulting in diminished virulence. However, we have been unable to identify either the *virB* gene product or the expected fusion product from pGTR123 by using in vitro transcription-translation analysis. Completion of the determination of the DNA sequence of the *virB* region along with the elucidation of its operon structure will aid in understanding the incomplete complementation by pGTR123.

The deduced amino acid sequence of the 28K protein is highly hydrophilic, suggesting a cytoplasmic location for the protein. Such a location suggests a regulatory role for the 28K protein. *vir-22*::Tn*5* results in the lack of expression of the production of the 28K protein, but not the 29K and 32K proteins, expressed by *E. coli* minicells possessing cosmid clones that confer full virulence to plasmid-cured *S. typhimurium* (14). If the 28K protein were either a repressor or an inducer of other identified virulence plasmid genes, then quantitative or qualitative differences in the production of other proteins would be expected to result from the *vir-22*::Tn*5* mutation. However, a role for the 28K protein in the expression of chromosomally encoded genes has not been examined, and the expression of other plasmid-encoded genes possibly affected by *virA* may be dependent on chromosomal genes of *S. typhimurium* not present in *E. coli*. Interestingly, VandenBosch et al. (33) have identified a locus, *rsk*, of the virulence plasmid that may affect the expression of chromosomally encoded virulence factors. *rsk* maps approximately 10 kb from *virA*.

An interesting finding from sequence analysis of the *Bam*HI-*Eco*RI fragment and the *vir-22*::Tn*5* insertion was that the only consensus promoter identified was approximately 300 bases from the start codon for the structural gene. We expected the promoter of *virA* to be contained within the 1.2-kb *Bam*HI-*Eco*RI fragment because subclones of this region expressed the 28K protein when present in both orientations in the vector plasmid (14; data not shown). Such a long leader sequence between the promoter and start codon suggests that sequences exist for the regulation of the expression of *virA*. *vir-22*::Tn*5* fell in this 300-base leader sequence. Furthermore, *vir-7*::Tn*5*, which does not inhibit the expression of the 28K protein in pYA422 (14), was located between *vir-22*::Tn*5* and the start codon. That *vir-7*::Tn*5* does not inhibit the production of the 28K protein in pYA422 suggests that Tn*5* may supply an outward promoter (6) to mediate the expression of *virA*. However, we would then expect *vir-22*::Tn*5*, which is located only 34 bases upstream of *vir-7*::Tn*5*, to similarly supply an outward promoter. *vir-7*::Tn*5* and *vir-22*::Tn*5* are inserted in opposite orientations relative to *virA*; however, outward promoters of Tn*5* are present in both termini (6). Further work needs to be done to elucidate the roles of the location and orientation of *vir-7*::Tn*5* and *vir-22*::Tn*5* in the expression of *virA*. Definitive identification of the promoter for *virA* awaits analysis of mRNA structure.

In addition to having its own promoter, *virA* may be part of the transcript of the *mkaA* gene identified and sequenced by Taira and Rhen (32), which is upstream and colinear with *virA*, ending at base 108. This hypothesis could explain why pYA422 does not complement *vir-22*::Tn*5*, although pYA422 fully encodes *virA* and a functional promoter. pYA422 lacks readthrough transcription possibly provided from *mkaA* or an upstream promoter such as the tetracycline resistance gene of pYA426 and may not sufficiently express *virA* for complementation. In contrast, *virA* of pYA426 could derive additional transcription from the tetracycline resistance gene of pACYC184. Consistent with this hypothesis, we have measured at least fivefold-higher levels of transcription of *virA* in pYA426 than in pYA422 (P. A. Gulig, unpublished data).

Others have used molecular genetic analysis to examine potential virulence genes of the *S. typhimurium* plasmid. Hackett et al. (15) have cloned a gene which confers increased serum resistance to *E. coli* K-12, and VandenBosch et al. (33–35) have identified, cloned, and sequenced a locus named *rsk* which is involved in the regulation of chromosomally encoded virulence factors. However, the direct role of these two loci in plasmid-mediated virulence has not been investigated through mutational-complementational analysis. Norel et al. (24) have also cloned regions of the virulence plasmid and have identified protein products from sequences which encode virulence functions. *virA* falls within the center of all consensus virulence regions identified by Gulig and Curtiss (14) and others (2, 3, 12a, 23, 24, 28, 30, 36). Therefore, it is likely that this gene is conserved among the virulence plasmid-containing serovars of *Salmonella* spp. Pullinger et al. (29) recently sequenced a plasmid-encoded gene encoding a 33K protein. A Tn*10* insertion, M242 (2), within this gene inhibits the virulence of *S. typhimurium*. This gene (and the 33K protein) are therefore probably involved in virulence. As stated above, Taira and Rhen (32) reported a sequence analysis of *mkaA*, a gene immediately upstream of *virA* and encoding a 70K protein.

Norel et al. (25) recently published the nucleotide sequence of a gene potentially encoding a 28K protein identi-

fied by Gulig and Curtiss (14) and Norel et al. (24). However, a biological phenotype such as virulence or a biochemical phenotype such as protein expression was not proven for the DNA sequence. The sequence of Norel et al. (25) corresponds almost exactly to that of *virA* presented here, with only 14 base-pair substitutions and nine amino acid residue differences. Additionally, the A residue at position 1044 of our sequence is missing from that of Norel et al. The resulting frameshift changes the next 20 amino acid residues and adds 4 amino acids to our peptide sequence because of the use of a different stop codon.

In summary, we have definitively identified the 28K protein encoded by *virA* of the *S. typhimurium* virulence plasmid as a virulence factor for *S. typhimurium*. We have determined the nucleotide sequence of the gene and surrounding regions of the plasmid. These results should facilitate achieving an understanding of the role of the 28K protein in plasmid-mediated virulence of *Salmonella* spp. and the determination of the regulation of the expression of *virA*.

ACKNOWLEDGMENTS

We thank Roy Curtiss III, Richard M. Lottenberg, Jorge E. Galán, Hank A. Lockman, and Steve A. Tinge for critical review of the manuscript and Henry V. Baker for advice and supplying the reverse sequencing primer.

This work was supported by a grant from the Interdisciplinary Center for Biotechnology Research of the University of Florida and by grant 89GIA81 from the American Heart Association-Florida Affiliate to P.A.G.

LITERATURE CITED

1. **Auerswald, E. A., G. Ludwig, and H. Schaller.** 1981. Structural analysis of Tn*5*. Cold Spring Harbor Symp. Quant. Biol. **45:** 107–113.
2. **Baird, G. D., E. J. Manning, and P. W. Jones.** 1985. Evidence for related virulence sequences in plasmids of *Salmonella dublin* and *Salmonella typhimurium*. J. Gen. Microbiol. **131:**1815–1823.
3. **Beninger, P. R., G. Chikami, K. Tanabe, C. Roudier, J. Fierer, and D. G. Guiney.** 1988. Physical and genetic mapping of the *Salmonella dublin* virulence plasmid pSDL2. Relationship to plasmids from other *Salmonella* strains. J. Clin. Invest. **81:** 1341–1347.
4. **Benson, S. A.** 1984. A rapid procedure for isolation of DNA fragments from agarose gels. Biotechniques **2:**66–68.
5. **Berg, D. E., and C. M. Berg.** 1983. The prokaryotic transposable element Tn*5*. Bio/Technology **1:**417–435.
6. **Berg, D. E., A. Weiss, and L. Crossland.** 1980. Polarity of Tn*5* insertion mutations in *Escherichia coli*. J. Bacteriol. **142:**439–446.
7. **Birnboim, H. C.** 1983. A rapid alkaline extraction method for the isolation of plasmid DNA. Methods Enzymol. **100:**243–255.
8. **Chamberlain, J. P.** 1979. Fluorographic detection of radioactivity in polyacrylamide gels with the water-soluble fluor, sodium salicylate. Anal. Biochem. **98:**132–135.
9. **Chou, P. Y., and G. D. Fasman.** 1978. Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. **47:**45–148.
10. **Curtiss, R., III.** 1981. Gene transfer, p. 243–265. *In* P. Gerhardt, R. G. E. Murray, R. Costilow, E. W. Nester, W. A. Wood, N. R. Krieg, and G. B. Phillips (ed.), Manual of methods for general bacteriology. American Society for Microbiology, Washington, D.C.
11. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
12. **Frazer, A. C., and R. Curtiss III.** 1975. Production, properties and utility of bacterial minicells. Curr. Top. Microbiol. Immunol. **69:**1–84.

12a. **Gulig, P. A.** 1990. Virulence plasmids of *Salmonella typhimurium* and other Salmonellae. Microb. Pathog. **8:**3–11.

13. **Gulig, P. A., and R. Curtiss III.** 1987. Plasmid-associated virulence of *Salmonella typhimurium*. Infect. Immun. **55:**2891–2901.
14. **Gulig, P. A., and R. Curtiss III.** 1988. Cloning and transposon insertion mutagenesis of virulence genes of the 100-kilobase plasmid of *Salmonella typhimurium*. Infect. Immun. **56:**3262–3271.
15. **Hackett, J., P. Wyk, P. Reeves, and V. Mathan.** 1987. Mediation of serum resistance in *Salmonella typhimurium* by an 11-kilodalton polypeptide encoded by the cryptic plasmid. J. Infect. Dis. **155:**540–549.
16. **Henikoff, S.** 1987. Unidirectional digestion with exonuclease III in DNA sequence analysis. Methods Enzymol. **155:**156–165.
17. **Humphreys, G. O., A. Weston, M. G. M. Brown, and J. R. Saunders.** 1979. Plasmid transformation in *Escherichia coli*, p. 254–279. *In* S. W. Glover and L. O. Butler (ed.), Transformation 1978. Cotswold Press, Oxford.
18. **Jameson, B. A., and H. Wolf.** 1988. The antigenic index: a novel algorithm for predicting determinants. Comput. Appl. Biosci. **4:**181–186.
19. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of protein. J. Mol. Biol. **157:**105–132.
20. **Lennox, E. S.** 1955. Transduction of linked genetic characters of the host by bacteriophage P1. Virology **1:**190–206.
21. **Lipman, D. J., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227:**1435–1441.
22. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual, p. 440. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. **Michiels, T., M. Y. Popoff, S. Durviaux, C. Coynault, and G. Cornelis.** 1987. A new method for the physical and genetic mapping of large plasmids: application to the localisation of the virulence determinants on the 90 kb plasmid of *Salmonella typhimurium*. Microb. Pathog. **3:**109–116.
24. **Norel, F., C. Coynault, I. Miras, D. Hermant, and M. Y. Popoff.** 1989. Cloning and expression of plasmid DNA sequences involved in *Salmonella* serotype Typhimurium virulence. Mol. Microbiol. **3:**733–743.
25. **Norel, F., M.-R. Pisano, J. Nicoli, and M. Y. Popoff.** 1989. Nucleotide sequence of the plasmid-borne virulence gene *mkfA* encoding a 28 kDa polypeptide from *Salmonella typhimurium*. Res. Microbiol. **140:**263–265.
26. **Pardon, P., M. Y. Popoff, C. Coynault, J. Marly, and I. Miras.** 1986. Virulence-associated plasmids of *Salmonella* serotype Typhimurium in experimental murine infection. Ann. Microbiol. (Paris) **137:**47–60.
27. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85:**2444–2448.
28. **Poppe, C., R. Curtiss III, P. A. Gulig, and C. L. Gyles.** 1989. Hybridization studies with a DNA probe derived from the virulence region of the 60 Mdal plasmid of *Salmonella typhimurium*. Can. J. Microbiol. **53:**378–384.
29. **Pullinger, G. D., G. D. Baird, C. M. Williamson, and A. J. Lax.** 1989. Nucleotide sequence of a plasmid gene involved in the virulence of salmonellas. Nucleic Acids Res. **17:**7983.
30. **Rhen, M., M. Virtanen, and P. H. Mäkela.** 1989. Localization by insertion mutagenesis of a virulence-associated region on the *Salmonella typhimurium* 96 kilobase pair plasmid. Microb. Pathog. **6:**153–158.
31. **Rose, R. E.** 1988. The nucleotide sequence of pACYC184. Nucleic Acids Res. **16:**355.
32. **Taira, S., and M. Rhen.** 1989. Identification and genetic analysis of *mkaA*—a gene of the *Salmonella typhimurium* virulence plasmid necessary for intracellular growth. Microb. Pathog. **7:**165–173.
33. **VandenBosch, J. L., D. R. Kurlandsky, R. Urdangaray, and G. W. Jones.** 1989. Evidence of coordinate regulation of viru-

lence in *Salmonella typhimurium* involving the *rsk* element of the 95-kilobase plasmid. Infect. Immun. **57**:2566–2568.

34. **VandenBosch, J. L., D. K. Rabert, and G. W. Jones.** 1987. Plasmid-associated resistance of *Salmonella typhimurium* to complement activated by the classical pathway. Infect. Immun. **55**:2645–2652.
35. **VandenBosch, J. L., D. K. Rabert, D. R. Kurlandsky, and G. W. Jones.** 1989. Sequence analysis of *rsk*, a portion of the 95-kilobase plasmid of *Salmonella typhimurium* associated with resistance to the bactericidal activity of serum. Infect. Immun. **57**:850–857.
36. **Williamson, C. M., G. D. Baird, and E. J. Manning.** 1988. A common virulence region on plasmids from eleven serotypes of *Salmonella*. J. Gen. Microbiol. **134**:975–982.