Transcriptional Organization and Function of Invasion Genes within *Salmonella enterica* Serovar Typhimurium Pathogenicity Island 1, Including the prgH, prgl, prgJ, prgK, orgA, orgB, and orgC Genes

JOANNA R. KLEIN, THOMAS F. FAHLEN, AND BRADLEY D. JONES*

Department of Microbiology, University of Iowa School of Medicine, Iowa City, Iowa 52242-1109

Received 20 December 1999/Returned for modification 18 February 2000/Accepted 24 March 2000

*Corresponding author. Mailing address: Department of Microbiology, University of Iowa School of Medicine, Iowa City, Iowa 52242-1109. Phone: (319) 353-5457. Fax: (319) 335-9006. E-mail: bjones@blue.weeg.uio.edu.

*Salmonella enterica* serovar Typhimurium initiates infection of a host by inducing its own uptake into specialized M cells which reside within the epithelium overlaying Peyer’s patches. Entry of *Salmonella* into intestinal epithelial cells is dependent upon invasion genes that are clustered together in *Salmonella* pathogenicity island 1 (SPI-1). Upon contact between serovar Typhimurium and epithelial cells targeted for bacterial internalization, bacterial proteins are injected into the host cell through a type III secretion system that leads to internalization of the bacteria. Previous work has established that the prgH, -J, -I, and -K and orgA genes reside in SPI-1, and the products of these genes are predicted to be components of the invasion secretion apparatus. We report that an error in the published orgA DNA sequence has been identified so that this region encodes two small genes rather than a single large open reading frame. These genes have been designated orgA and orgB. Additionally, an opening reading frame downstream of orgB, which we have designated orgC, has been identified and partially characterized. Previously published work has indicated that the prgH, -I, -J, and -K genes are transcribed from a promoter distinct from that used by the gene immediately downstream, orgA. Here, we present experiments indicating that orgA expression is driven by the prgH promoter. In addition, using reverse transcriptase PCR analysis, we have found that this polyclonal message extends downstream of prgH to include a total of 10 genes. To more fully characterize this invasion operon, we demonstrate that the prgH, prgl, prgJ, prgK, orgA, and orgB genes are each required for invasion and secretion, while orgC is not essential for the invasive phenotype.

*Salmonella* infections are an important health problem in both the developing and developed world (9, 32, 34). Pathogenic *Salmonella* species cause infections that range in severity from self-limiting gastroenteritis to life-threatening systemic dissemination (45). After entry into a host, the bacteria establish infection by attaching to and invading specialized M cells associated with Peyer’s patches in the small intestine (5, 23, 26). Following M-cell invasion and destruction, host-restricted *Salmonella* species cause localized destruction of the intestinal epithelium (gastroenteritis). In contrast, passage of host-adapted *Salmonella* species through M cells allows rapid dissemination to the mesenteric lymph nodes and then to the liver and spleen, where unchecked growth causes death (25).

A critical determinant in the development of *Salmonella* disease is the ability of the bacteria to invade cells. *Salmonella enterica* serovar Typhimurium mutants that are unable to invade tissue culture cells are defective in their ability to invade and destroy M cells (26, 43). This defect severely limits the ability of the bacteria to initiate infection and reduces their virulence in mice (14, 24, 43). Genes required for *Salmonella* internalization into mammalian cells have been identified (4, 7, 14, 16, 17, 20, 22, 24, 27–29, 31, 39) and have been shown to reside on *Salmonella* pathogenicity island 1 (SPI-1) at centisome 63 (38) as well as two genes, sopE and sigD (sopB), that reside outside of SPI-1 (21, 47, 48). The invasion mechanism of *Salmonella* relies on a type III secretion system that secretes effector proteins into host cells targeted for invasion (6, 7, 20, 22, 28, 29, 33). Intracellular effector proteins transmit a signal to the cell which induces a rearrangement of the host cell cytoskeleton that results in bacterial uptake (12, 13, 16).

Four genes that encode SPI-1 secretion apparatus proteins are prgH, -I, -J, and -K. The prgH gene was first identified as a phoP-repressed gene (37) that was required for serovar Typhimurium invasion of tissue culture cells and virulence in an animal model by both the oral and intraperitoneal routes of infection (4). Subsequent work with the prgH gene identified a cluster of four genes (prgH, -I, -J, and -K) that were cotranscribed (42). Comparative analysis of the predicted products of the prgH, -I, -J, and -K genes revealed similarities to *Shigella flexneri* and *Yersinia enterocolitica* proteins that are essential for protein secretion via a type III mechanism (42). Northern blotting indicated that the gene downstream of the prgH operon, orgA, did not reside on the same mRNA transcript nor did phoP appear to regulate levels of the orgA transcript, in contrast to the prgH transcript. However, work by another group indicated that an orgA::lacZY fusion is repressed by a phoP(Con) serovar strain of Typhimurium (3).

The orgA gene was identified by using a screen to identify oxygen-regulated genes that were required for serovar Typhimurium invasion (24). A serovar Typhimurium orgA::Tn5lacZY mutant is noninvasive and has reduced virulence for mice following oral infection. Other work has shown that this mutation prevents the invasion and destruction of M cells and has a general defect in secretion of invasion effector proteins (43). In addition, orgA is similar to the mxiK gene in *Shigella* (1), a putative component of the type III secretion system in that pathogen. The orgA gene was originally believed to encode a
protein of 412 amino acids. We have recently identified a sequencing error in the orgA gene which, when corrected, reveals that this region actually encodes two open reading frames (ORFs) which we have named orgA and orgB.

In an effort to more fully characterize the prgH, -I, -J, and -K and orgA and -B genes, we have performed work to determine the transcriptional organization of these genes. In addition, the role of each of these genes in type III-mediated secretion and tissue culture invasion has been assessed, and our findings are presented below.

**MATERIALS AND METHODS**

**Bacterial strains and growth conditions.** The bacterial strains and plasmids used are listed in Table 1. Bacteria were grown in Lennox broth (LB) or Mueller-Hinton broth (DIFCO). LB agar (Gibco/BRL) or MacConkey lactose agar (Gibco/BRL) were used where indicated. Antibiotics were added at the following concentrations: ampicillin, 100 μg/ml; kanamycin, 25 μg/ml; tetracycline, 5 mM.

**TABLE 1. Bacterial strains and plasmids used in this study**

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Genotype or phenotype</th>
<th>Reference or source</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> DH12S</td>
<td>mcrA Δ(mrr-hsdRMS-merBC) F’ lacI513 lacZΔM15</td>
<td>Gibco-BRL</td>
</tr>
<tr>
<td>Serovar Typhimurium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SL1344</td>
<td>Mouse virulent strain</td>
<td>49</td>
</tr>
<tr>
<td>BJ53</td>
<td>EE251 derivative of orgA::Tn5lacZY</td>
<td>23</td>
</tr>
<tr>
<td>BJ66</td>
<td>SL1344 derivative of orgA::Tn5lacZY</td>
<td>43</td>
</tr>
<tr>
<td>EE251</td>
<td>rpsL isolate of SL4012</td>
<td>31</td>
</tr>
<tr>
<td>EE566</td>
<td>SL1344, prgH::Tn5lacZY, Tet'</td>
<td>3</td>
</tr>
<tr>
<td>BJ749</td>
<td>EE251, prgH::Tn5, Kan', Tet'</td>
<td>11</td>
</tr>
<tr>
<td>JK1</td>
<td>SL1344, prgK::aphT, Amp', Kan'</td>
<td>This work</td>
</tr>
<tr>
<td>JK9</td>
<td>SL1344, prgK::aphT, Amp', Kan'</td>
<td>This work</td>
</tr>
<tr>
<td>JK10</td>
<td>SL1344, prgK::aphT, Amp', Kan'</td>
<td>This work</td>
</tr>
<tr>
<td>JK11</td>
<td>SL1344, prgH mutation, Kan'</td>
<td>This work</td>
</tr>
<tr>
<td>JK17</td>
<td>SL1344, prgK::aphT, Amp', Kan'</td>
<td>This work</td>
</tr>
<tr>
<td>JK23</td>
<td>SL1344, orgB::Tn5, Kan'</td>
<td>This work</td>
</tr>
<tr>
<td>TF78</td>
<td>SL1344, orgC::Tn5, Kan'</td>
<td>This work</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plasmids</th>
<th>Genotype or phenotype</th>
<th>Reference or source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBDJ116</td>
<td>prgH, -I, -J, -K, orgA::Tn5lacZY, Amp', Tet'</td>
<td>This work</td>
</tr>
<tr>
<td>pBDJ129</td>
<td>Variable-copy-number plasmid, Cm'</td>
<td>This work</td>
</tr>
<tr>
<td>pBDJ134</td>
<td>prgK, orgA, -B, Amp'</td>
<td>This work</td>
</tr>
<tr>
<td>pBDJ135</td>
<td>prgK, orgA, Amp'</td>
<td>This work</td>
</tr>
<tr>
<td>pBDJ142</td>
<td>prgH, -I, -J, -K, orgA::Tn5lacZY, Tet', Cm'</td>
<td>This work</td>
</tr>
<tr>
<td>pBDJ143</td>
<td>orgA and -B, Amp'</td>
<td>This work</td>
</tr>
<tr>
<td>pBDJ149</td>
<td>prgH, -I, -J, -K, orgA::Tn5lacZY, Tet', Cm'</td>
<td>This work</td>
</tr>
<tr>
<td>pBDJ200</td>
<td>Suicide plasmid, Amp'</td>
<td>This work</td>
</tr>
<tr>
<td>pBDJ200K</td>
<td>Kan' derivative of pBDJ200</td>
<td>This work</td>
</tr>
<tr>
<td>pBDJ233</td>
<td>prgH, -I, Amp', Tet'</td>
<td>11</td>
</tr>
<tr>
<td>pJK001</td>
<td>prgH, -I, -J, -K, orgA, -B, Amp', Kan'</td>
<td>11</td>
</tr>
<tr>
<td>pJK007</td>
<td>prgH, -I, -J, -K, orgA, Amp'</td>
<td>This work</td>
</tr>
<tr>
<td>pJK003</td>
<td>prgH, -I, -J, -K, orgB, Amp', Kan'</td>
<td>This work</td>
</tr>
<tr>
<td>pACL016</td>
<td>pGEX-2T (Pharmacia) derivative, Spe', Amp'</td>
<td>This work</td>
</tr>
<tr>
<td>pUC4K-D</td>
<td>pUC4K derivative with Kan' cassette lacking transcriptional terminator</td>
<td>This work</td>
</tr>
</tbody>
</table>

* Tet', tetracycline resistant; Amp', ampicillin resistant; Kan', kanamycin resistant; Cm', chloramphenicol resistant; Spe', spectinomycin resistant.

*5lacZY* was cloned into pACYC177 to create pBDJ116. Sequences upstream of prgH and a portion of the prgH coding region were deleted from pBDJ142 by digestion with HindIII and XhoI, resulting in plasmid pBDJ149. Plasmid pJK001 carries the prgH, -I, -J, and -K and orgA and -B genes in pACYC177 and was constructed as previously described (11). Plasmid pJK007 is a derivative of pJK001 and was created by deleting the kanamycin gene and the majority of orgB by digestion with SalI. Plasmid pJK003 is a derivative of pJK001 from which a 48-bp segment from within the orgA coding region was deleted by digestion with NarI. The plasmid pUC4K-D carries the aphT gene, encoding kanamycin resistance, with its transcriptional termination sequences deleted and was constructed as previously described (15). Insertion of this cassette into a gene creates a mutation with little, or no, downstream polar effects.

**Suicide plasmid system.** Two plasmids which are used as a chromosomal integration system and function as described below have been created in our laboratory. Plasmid pBDJ129 is a variable-copy-number plasmid which codes for chloramphenicol resistance. It contains two origins of replication, a high-copy-number F1 origin under the control of the lac promoter and a low-copy-number mini-F origin. Mini-F replication is controlled by the RepE protein which binds to DNA direct repeats in the origin to control plasmid replication and copy number (41). pBDJ200 carries the lacF allele, encodes ampicillin resistance, replicates via the high-copy-number ColE1 origin of replication, and carries a truncated repE replication gene and RepE DNA binding sites. The RepE DNA binding sites competitively bind RepE protein produced by the lower-copy-number mini-F plasmid and prevent its replication. When both plasmids are present within the same bacterium, chloramphenicol resistance can only be maintained if the mini-F plasmid integrates into the chromosome. In cases where it is necessary to subsequently remove pBDJ200 from the cell, plasmid pACL016, whose origin is incompatible with pBDJ200, can be introduced into the bacterium. This plasmid is an ampicillin-sensitive derivative of pGEX-2T and encodes spectinomycin resistance.

**RNA isolation and RT-PCR.** SL1344 RNA was isolated from 1 ml of culture grown under low-oxygen conditions by using an SV total RNA isolation kit (Promega, Madison, Wis.). RNA was treated with 10 U of DNase at 37°C for 30 min and was extracted with phenol-chloroform. The Access reverse transcriptase PCR (RT-PCR) kit (Promega) was used for generation of cDNA template and
subsequent amplification. RNA was heat denatured with 50 pmol of 3′ primer for 2 min at 94°C. Serovar Typhimurium SL1344 chromosomal DNA was used as the genomic DNA control. cDNA was synthesized after addition of avian myeloblas-
tosis virus reverse transcriptase (RT) and incubation at 37°C for 15 min at 45°C for 45 min. RT was inactivated by heating to 94°C for 3 min, and cDNA was amplified by 40 cycles of PCR (94°C, 30 s; 52 to 56°C, 1 min; 68°C, 2 min) and a 7-min extension at 68°C. The 5′ primer and T1 DNA polymerase were added during the first PCR cycle to the annealing temperature. One-fifth of this reaction was analyzed on a 0.7% agarose gel. The sequences of primers used are as follows: 1, 5′-CGTTTTCAGGTTTGCGACG-3′; 2, 5′-GGAAATTTGTCCGCGAGCGCAGCTGA-3′; 3, 5′-CAACGCGACGGGATACACGGGCTGGT-3′; 4, 5′-AGATTACCTCATTACGATGTCGCTGCTGCA-3′; 5, 5′-GGAAATTTGCGTGTTGCCAG-3′; 6, 5′-GTCGATAGCTATTACCGCAC-3′; 7, 5′-CCATTTATTGAGGAGGCATTG-3′; 8, 5′-AGATTACCTCATTACGATGTCGCTGCTGCA-3′.

**RESULTS**

Identification of three ORFs: orgA, orgB, and orgC. While performing experiments to more thoroughly characterize the serovar Typhimurium orgA gene, several observations led us to believe that the published orgA DNA sequence might be incorrect (24). Sequence comparison of the orgA gene with the mxiK gene of Shigella revealed that mxiK aligned with the first half of orgA and an uncharacterized Shigella ORF downstream of mxiK aligned with the last half of the orgA gene. Since the downstream ORF in Shigella has not been named or assigned any function, it was unclear whether a sequencing error had occurred in the Shigella sequence or in the serovar Typhimurium orgA sequence. We subsequently used the BLAST (2) searches of our sequences were carried out against sequences found at the National Center for Biotechnology Information database (www.ncbi.nlm.nih.gov), the serovar Typhimurium genome sequenc-
ing database at Washington University (http://genome.wustl.edu/gsc/bacterial/salmonella.shtml), and the Salmonella enterica serovar Typhi sequencing data-
base at the Sanger Centre (http://www.sanger.ac.uk/Projects/S_typhi/blast_server.shtml).

Identification of three ORFs: orgA, orgB, and orgC. While performing experiments to more thoroughly characterize the serovar Typhimurium orgA gene, several observations led us to believe that the published orgA DNA sequence might be incorrect (24). Sequence comparison of the orgA gene with the mxiK gene of Shigella revealed that mxiK aligned with the first half of orgA and an uncharacterized Shigella ORF downstream of mxiK aligned with the last half of the orgA gene. Since the downstream ORF in Shigella has not been named or assigned any function, it was unclear whether a sequencing error had occurred in the Shigella sequence or in the serovar Typhimurium orgA sequence. We subsequently used the BLAST (2) searches of our sequences were carried out against sequences found at the National Center for Biotechnology Information database (www.ncbi.nlm.nih.gov), the serovar Typhimurium genome sequenc-
ing database at Washington University (http://genome.wustl.edu/gsc/bacterial/salmonella.shtml), and the Salmonella enterica serovar Typhi sequencing data-
base at the Sanger Centre (http://www.sanger.ac.uk/Projects/S_typhi/blast_server.shtml).

Identification of three ORFs: orgA, orgB, and orgC. While performing experiments to more thoroughly characterize the serovar Typhimurium orgA gene, several observations led us to believe that the published orgA DNA sequence might be incorrect (24). Sequence comparison of the orgA gene with the mxiK gene of Shigella revealed that mxiK aligned with the first half of orgA and an uncharacterized Shigella ORF downstream of mxiK aligned with the last half of the orgA gene. Since the downstream ORF in Shigella has not been named or assigned any function, it was unclear whether a sequencing error had occurred in the Shigella sequence or in the serovar Typhimurium orgA sequence. We subsequently used the BLAST (2) searches of our sequences were carried out against sequences found at the National Center for Biotechnology Information database (www.ncbi.nlm.nih.gov), the serovar Typhimurium genome sequenc-
ing database at Washington University (http://genome.wustl.edu/gsc/bacterial/salmonella.shtml), and the Salmonella enterica serovar Typhi sequencing data-
base at the Sanger Centre (http://www.sanger.ac.uk/Projects/S_typhi/blast_server.shtml).

Further analysis of this region of SPI-1 reveals the presence of another ORF, which overlaps with the orgC gene. We have named this gene orgC'. We identified orgC' through a transposon mutagenesis screen done in our laboratory to identify regulators of hiaA (11). Searches of the serovar Typhimurium sequencing database at Washington University with DNA se-
quences flanking these transposon insertions revealed that three insertions were downstream of orgB, in a previously unidenti-
fied ORF which we call orgC. The orgC gene was also found in serovar Typhi by searching the sequencing database at the Sanger Centre. The orgC sequence can be found at bp 1510 to 1995 on Stn contig 1455 at the Washington University
serovar Typhimurium sequencing database. Previous work in our laboratory has found that transposon insertions in orgC have a small effect on expression of a hilA: lacZ reporter when measured in β-Gal assays (11), but the insertions in this gene do not have a detectable effect on the invasion phenotype of serovar Typhimurium. Homology searches with the orgC gene sequence failed to reveal similarity to any known genes.

Deletions and insertions in the promoter and ORF of prgH affect transcription of an orgA:: lacZ reporter. It was previously reported that the serovar Typhimurium invasion genes prgH, -I, -J, and -K are transcribed by a promoter upstream of the prgH gene and that the downstream gene, orgA, is transcribed by a separate promoter (42). During our efforts to characterize the orgA and orgB genes, we sought to identify the promoter elements that are responsible for transcription of the orgA and orgB genes by performing a series of deletion experiments. Initial experiments were performed with a low-copy-number plasmid, pBDJ142, which carries the prgH promoter; the prgH, -I, -J, and -K genes; and the S′ portion of the orgA gene transcriptionally fused to lacZ. The orgA:: lacZ fusion on this plasmid has been shown to be regulated in a fashion similar to that observed when the orgA:: lacZ fusion resides on the Salmonella chromosome, indicating that the org promoter and regulatory sequences are intact on plasmid pBDJ142 (24). As part of the experiments performed to localize the orgA and -B promoter sequences, a HindIII-XhoI deletion was made in pBDJ142 that removes all sequences upstream of the prgH gene as well as a portion of the prgH coding region. To our surprise, the orgA:: lacZ fusion on the plasmid with this deletion could no longer be induced by low-oxygen conditions and expressed β-Gal at levels comparable to the parent plasmid when grown in repressing (high-oxygen) conditions. This result indicated that the HindIII-XhoI deletion had disrupted sequences important for orgA expression (Fig. 1). These results led us to investigate whether the orgA and orgB genes might, in fact, be part of the prgH transcriptional unit. Experiments were performed to determine whether a transposon insertion in the upstream prgH gene has a polar effect on orgA transcription. The orgA:: Tn5lacZY (Tet′) insertion of BJ53 was moved by P22 transduction into serovar Typhimurium BJ749 prgH::Tn5 (Kan′). Ninety-six kanamycin-resistant, tetracycline-resistant transductants were picked, and the expression of the orgA:: Tn5lacZY reporter of each strain was examined on MacConkey lactose agar plates. Colonies that normally express the orgA::Tn5lacZY have a white periphery with a red center (fish-eye) on MacConkey lactose agar plates due to oxygen regulation of orgA expression. Conversion of the fish eye phenotype to a white phenotype on MacConkey agar was observed for 93 of 96 kanamycin-resistant, tetracycline-resistant transductants. For comparison, transduction of the orgA:: Tn5lacZY fusion to serovar Typhimurium SL1344 reproducibly gave a fish-eye phenotype in the vast majority of transductants (data not shown). Quantitative assays of β-Gal activity produced by a representative kanamycin-resistant, tetracycline-resistant isolate were performed. Consistent with the MacConkey agar phenotype, β-Gal levels produced by the orgA:: Tn5lacZY fusion in a strain also carrying a prgH::Tn5 insertion (strain JK1) were significantly reduced under both high- and low-oxygen growth conditions compared to the control strain BJ53, which lacks the prgH:: Tn5 insertion (Table 2). To determine whether the reduced expression of the orgA::Tn5lacZY fusion was possibly due to loss of function of an activator in the prgH operon, plasmid pJK001, encoding prgH, -I, -J, and -K and orgA and -B was introduced into strain JK1. Expression of the orgA:: Tn5lacZY fusion was the same (uninduced) under high- and low-oxygen conditions (Table 1), suggesting that there was a cis effect of prgH::Tn5 on orgA expression.

The orgB gene is required for complementation of a polar serovar Typhimurium prgH:: Tn5lacZY mutant. Another method that was used to identify genes transcribed from the prgH promoter was determining what genes were required to complement a noninvasive polar prgH mutant back to wild-type invasiveness. Serovar Typhimurium E656, which carries a polar Tn5lacZY insertion in prgH, was transformed with either plasmid pJK001 or pJK007. Plasmid pJK007 carries the coding sequences of prgH, -I, -J, and -K and orgA but not orgB, while

![FIG. 1. Deletions in the prgH promoter disrupt expression and regulation of a Tn5lacZY transcriptional fusion in the orgA gene. Plasmid pBDJ149 is a HindIII-XhoI deletion derivative of pBDJ142. Serovar Typhimurium strains carrying either pBDJ142 or pBDJ149 were grown in high or low concentrations of oxygen before being assayed for β-Gal activity. Values are the means ± standard deviations from one experiment performed in triplicate and are representative of several experiments.

<table>
<thead>
<tr>
<th>Strain (genotype)</th>
<th>Plasmid (genotype)</th>
<th>β-Gal activity* (Miller units)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>High O₂</td>
</tr>
<tr>
<td>BJ53 (orgA::Tn5lacZY)</td>
<td></td>
<td>104 ± 0.9</td>
</tr>
<tr>
<td>JK1 (prgH::Tn5, orgA::Tn5lacZY)</td>
<td></td>
<td>28 ± 0.4</td>
</tr>
<tr>
<td>JK1 (prgH::Tn5, orgA::Tn5lacZY)</td>
<td>pJK001 (prgH', -I', -J' and -K' and orgA', -B')</td>
<td>32 ± 1.2</td>
</tr>
</tbody>
</table>

*Values are means ± standard deviations from one experiment performed in triplicate and are representative of several experiments.
pJK001 encodes functional genes for prgH, -I, -J, and -K and orgA and -B. The invasive phenotype of serovar Typhimurium EE656 was restored to wild-type levels when transformed with plasmid pJK001, but the strain remained noninvasive when carrying plasmid pJK007, which lacks orgB (Fig. 2). Thus, the polar prgH mutant requires complementation with orgB to regain full invasiveness. Control experiments demonstrated that the plasmids did not alter the invasion phenotype of SL1344 and that the vector backbone (pACYC177) did not alter invasion of EE656 (data not shown).

Ten genes from prgH to avrA are transcribed as a single polycistronic message. Several genes not required for the invasion mechanism of Salmonella have been identified downstream of orgB, and each is transcribed in the same direction as the prg and org genes. As described above, orgC is a gene with unknown function which overlaps the orgB ORF. The hilC-sirC-sprA gene lies 344 bp downstream of orgC, followed by sprB (384-bp intergenic region) and avrA (170-bp intergenic region). The sit operon, encoding genes involved in iron acquisition, lies downstream of avrA but is transcribed in the opposite direction.

To further demonstrate that orgA and orgB are cotranscribed with the prg genes and to determine whether any of these downstream genes are part of the prgH operon, we performed RT-PCR analysis of the region from prgH to avrA. RT-PCR is a reliable method to determine operon structure and has been used to define operon structure of genes within Salmonella pathogenicity island 2 (19) as well as in enteropathogenic Escherichia coli (35). In our experiments, primer pairs were designed to overlap with the neighboring primer pairs so that the entire transcript was covered in several steps. RT-dependent amplification was observed from primer pairs specific to the following genes: prgH-prgK, prgK-orgB, orgB-hilC, hilC-sprB, and sprB-avrA (Fig. 3), indicating that a single message extends from prgH to avrA. A 3′ primer downstream of avrA within the sitD coding region did not amplify a fragment, indicating that avrA is the last gene in the operon (Fig. 3).

Six of the genes in the prgH operon are required for invasion. Transposon insertions in prgH and orgA result in a noninvasive phenotype, demonstrating that genes within this operon are required for invasion. Since these insertions are polar on downstream genes, it was necessary to examine each gene individually to determine its role in the invasion process. Since published reports have demonstrated that the hilC gene has a subtle role in invasion and the sprB and avrA genes are not required for invasion, we only tested the role of each prg and org gene in Salmonella invasion. Similar to the genes downstream of orgC, we found that a serovar Typhimurium orgC mutant (TF78) is fully invasive for HEp-2 cells, indicating that orgC is not required for the invasive phenotype (Fig. 4). Subsequent experiments were designed and interpreted keeping this result in mind. Chromosomal mutations were made in prgH, prgI, and prgJ as described in Materials and Methods. To examine the downstream effects, if any, of these mutations, the orgA::Tn5lacZY insertion was transduced into the serovar Typhimurium strain carrying the prgH, prgI, or prgJ mutation. None of the mutations had an effect on orgA::lacZY expression compared to the strain carrying only the orgA::lacZY fusion, BJ66, indicating that the mutations had no detectable downstream effect on orgA transcription (data not shown). Each of these mutants was reduced approximately 100-fold in its ability to enter HEp-2 cells (Fig. 4). The invasion defect of each strain was restored by introducing plasmid pBDJ233 (prgH and -I) into the prgH or prgI mutant and plasmid pBDJ116 (prgH, -I, -J, and -K) into the prgJ mutant (Fig. 4). Although a downstream gene was introduced into the prgH and prgJ mutants, there are genes further downstream which are required for invasion, so the indicated plasmids were used for the complementation experiments. The primary reason for using these plasmids, and others described below, was that it was not possible to clone the individual invasion genes so that they were expressed by an exogenous promoter from a high-copy-number plasmid (ColE1 origin) or medium-copy-number plasmid (p15A origin). Therefore, we used the different plasmids that we had available to assess the function of each gene in invasion as described here. To analyze the role of the prgK gene in invasion, we examined the invasiveness of strain JK10 which has a polar mutation in prgK, affecting prgK and the downstream genes. Plasmid pBDJ143, which carries the intact orgA and orgB genes, was introduced into this strain to allow analysis of the role of the prgK gene in invasion. Genes further downstream were not included on the plasmid since they are not required for invasion, and so this strain is effectively prgK mutant. This strain was reduced 100-fold in invasiveness compared to the wild type, and invasion of JK10 was complemented by pBDJ134 (prgK, orgA, orgB) (Fig. 4). To determine the role of orgA in invasion, we looked at the noninvasive strain BJ66 (orgA::Tn5lacZY) carrying plasmid pJK003 (orgB) or plasmid pBDJ143 (orgA and orgB). BJ66 has invasion levels <1% of the wild type (24) and is complemented to wild-type levels only when both orgA and orgB are present (Fig. 4), showing that orgA is also a required invasion gene. Finally, a Tn5 insertion in orgB (JK23) results in a loss of invasion that can be complemented by plasmid pJK003 (orgB) (Fig. 4). Consistent with this result, it was previously published that BJ66 is not complemented by pBDJ135, which is lacking orgB sequences (orgA only), demonstrating that orgB is a required invasion gene (24). Control experiments demonstrated that complementing plasmids did not have an effect on invasion of
wild-type Salmonella (SL1344), and the vector backbones of the complementing plasmids did not alter invasion of the mutants (data not shown). From these data we conclude that prgH, -I, -J, and -K, in addition to orgA and orgB, are required for the serovar Typhimurium invasion phenotype.

Proteins required for invasion are also required for secretion. It has been published that polar transposon insertions in prgH and orgA result in defective secretion of proteins (24, 42), suggesting that proteins in this operon form part of the type III secretion apparatus. In our study of this operon, we have tested the requirement of each gene for secretion of invasion proteins. Bacteria were grown to late stationary phase, which has been shown to lead to the accumulation of secreted proteins in the culture supernatant, and the proteins were collected and analyzed by SDS-PAGE. The stained gel revealed that several proteins secreted from wild-type SL1344 were not secreted from strains with individual mutations in prgH, -I, -J, -K, orgA, or orgB (Fig. 5). The molecular masses of these proteins correspond to the sizes of secreted proteins previously reported: 80 kDa (SipA), 68 kDa (SipB), and 41 kDa (SipC) (28, 29, 43). No secretion defect was observed in the orgC mutant. (The bands are fainter in the orgC lane due to unequal loading.) These data, together with the invasion assay data presented above, establish that the prgH, -I, -J, and -K and orgA and -B genes each play an integral role in type III secretion, leading to Salmonella-mediated invasion of host cells.

DISCUSSION

The prgH, -I, -J, and -K and orgA and -B genes are believed to be important elements of the invasion machinery of Salmonella. Not only do polar insertions in prgH and orgA abolish invasion and secretion, it has been shown that prgH and prgK are components of a type III secretion needle structure (30). Mutations in these genes prevent formation of this needle structure. However, the effect of individual mutations in the other prg and org genes has not been established. Besides participating in the function of the type III secretion apparatus, genes in this region of the chromosome are also involved in regulation of the invasive phenotype. We have recently reported the isolation of transposon mutations which repress hilA expression (11). Many of these mutations cluster in regions upstream of prgH or near the orgC coding region. Two regulatory proteins are also encoded in this region of SPI-1. Upstream of prgH lies hilD, a derepressor of hilA (46). Downstream of orgC lies a gene with regulatory effects on SPI-1 gene expression. This gene has been independently identified by three different groups and has been named sirC (44), sprA (46), and sprA (10). Investigation of this region will lead to further understanding of the structure and function of the type III secretion apparatus and regulation of SPI-1 gene expression.

During the course of our work, it became apparent that the published serovar Typhimurium SL1344 orgA gene sequence was incorrect. Comparison of the orgA sequence to DNA sequences of this region in another serovar Typhimurium strain and in serovar Typhi, as well as to the gene homologs in Shigella, led us to resequence this region of DNA from serovar Typhimurium SL1344. That work revealed that two genes are actually encoded by this region. In addition, a new gene, orgC, which overlaps the orgB ORF was also identified. The role of this gene in Salmonella invasion is unclear since a mutation in orgC has no effect on the invasive phenotype of the strain.
Furthermore, we were unable to find homologs of the orgC gene in sequence data bank searches. However, multiple transposon insertions in orgC were identified in a screen for mutations that reduce expression of hilA (11). While the effect of these mutations on hilA expression is small, the observation suggests that orgC is a functional gene, although it is possible that the effect of these insertions is a downstream disruption of hilC. Further work will need to be done to determine its exact function.

In contrast to a published report (42), we present data acquired from several different experimental approaches that indicate orgA expression is dependent upon the prgH promoter. We observed that removing the prgH promoter from a plasmid carrying an orgA::Tn5 lacZY expression reporter significantly reduced transcription of the β-Gal reporter. We subsequently examined the effect of polar transposon insertions in prgH on orgA::Tn5 lacZY expression and found that they essentially abolished expression of the reporter construct (see Table 2). Furthermore, orgA expression in these strains does not return to normal levels when the prgH, -I, -J, and -K genes were restored to the strain on a plasmid, indicating that the transposon insertion has a cis effect on orgA expression. We also observed that complementation of the invasion defect caused by a polar prgH mutation required the orgA and orgB genes, in addition to prgH, -I, -J, and -K. Finally, RT-PCR analysis was performed, demonstrating that a single mRNA encodes sequences prgH through orgC and extends downstream to include three more genes.

Additional evidence supports our conclusion that orgA, -B, and -C are cotranscribed with prgH, -I, -J, and -K. Analysis of the prgK, orgA, orgB, and orgC ORFs reveals that these genes share coding sequence, a feature common to genes which are cotranscribed. In addition, the phoP-phoQ two-component system regulates the orgA gene in a manner identical to the prgH gene (3; B. D. Jones, unpublished observation), which is the expected result if the genes reside on the same transcript.

RT-PCR analysis suggests that the prg operon includes 10 genes: prgH, -I, -J, and -K; orgA, -B, and -C; hilC-sirC-sprA; sprB; and avrA. However, we believe that it is important to interpret this data with caution. Recent work by two laboratories (10, 44) found that the hilC-sirC-sprA and sprB genes do not require hilA for expression. These observations indicate that the hilC-sirC-sprA and sprB genes are transcribed from a promoter separate from the prgH promoter since the prgH promoter is hilA regulated (3). In light of that work, it seems
possible that our RT-PCR analysis has detected RNA transcripts that overlap in the region around the hilC-sirC-spra gene (between primers 3 and 9) and that hilC-sirC-spra, sprB, and avrA are transcribed from a promoter(s) that is separate from the prgH promoter. The hilC-sirC-spra gene is homologous to the AraC family of transcriptional regulators and has regulatory effects on SPI-1 genes. Schechter et al. identified hilC as a derepressor of hilA (46). Similarly, Rakeman et al. observed that sirC can activate SPI-1 genes in the absence of hilA (44). Eichelberg et al. found that constitutive expression of sprA increases transcription of SPI-1 genes and invasion of host cells under certain conditions (10). While a regulatory role for the HilC protein has been demonstrated, a mutant which lacks hilC is very mildly (two- to fourfold) reduced in levels of epithelial cell invasion under optimal conditions. The function of this gene may be subtle, or an assay more suitable for studying its function needs to be developed. Similarly, SprB, a member of the LuxR-UhpA family of regulatory proteins, is not required for bacterial internalization, and no specific function has been attributed to the gene or gene product (10). Recent work has shown that avrA is homologous to avirulence genes of plant pathogen and that the gene product, AvrA, is secreted through the type III secretion system of SPI-1 (18). However, to date no function in Salmonella virulence or avirulence has been described for this gene.

To provide a more-complete understanding of the role of each gene within the operon, we constructed strains with defined mutations to assess the role of each gene in the bacterial entry mechanism. The invasiveness of serovar Typhimurium strains lacking only prgH, prgl, prgJ, prgK, orgA, or orgB was reduced approximately 100-fold compared to the fully invasive serovar Typhimurium parent strain SL1344. In each instance, the loss of invasiveness of the mutant could be restored by a plasmid carrying the appropriate gene. Individual mutations in prgH, prgl, prgJ, prgK, orgA, or orgB also resulted in a general defect of Salmonella effector protein secretion. Each of these genes likely functions as an integral component of the type III invasion secretion system that transports specific effector proteins into the host cell to induce the uptake of these pathogens.

In this study, we have acquired data that clarifies the transcriptional organization of the prgH operon. In addition, we have confirmed or confirmed that six genes, prgH, -J, -K, orgA, and orgB, in this operon are required for entry of the bacteria into tissue culture cells. Future experiments in this laboratory will be aimed at determining the functions of these proteins in the type III invasion secretion mechanism.

ACKNOWLEDGMENTS

We thank Cathy Lee for providing us with serovar Typhimurium E656. J.R.K. is supported by NIH predoctoral training grant GM08629. T.F.F. is supported by a predoctoral fellowship on NIH training grant 5T32 AI 07511. This work was supported by NIH grant AI38268 and a grant from the Roy J. Carver Charitable Trust (to B.D.J.).

REFERENCES


Editor: A. D. O’Brien