

Characterization of Virulence Genes of Enteroinvasive *Escherichia coli* by *TnphoA* Mutagenesis: Identification of *invX*, a Gene Required for Entry into HEp-2 Cells

RU-CHING HSIA,¹ PAM L. C. SMALL,² AND PATRIK M. BAVOIL^{1*}

Department of Microbiology and Immunology, University of Rochester, Rochester, New York 14642,¹
and Rocky Mountain Laboratories, Hamilton, Montana 59840²

Received 15 January 1993/Accepted 21 May 1993

While enteroinvasive *Escherichia coli* (EIEC) and shigellae are genotypically nearly identical, a difference has been reported in the infective dose to humans: EIEC is 10,000-fold less infectious than shigellae. A possible basis for this difference lies in the inherent invasiveness of these bacteria toward epithelial cells. Thus, despite the high degree of homology between the invasion plasmids of EIEC and shigellae, substantial differences in genetic organization and/or sequence may exist. We have undertaken a systematic genetic analysis of the EIEC plasmid pSF204, using transposon mutagenesis. Congo red-negative *TnphoA* insertion mutants ($Pcr^- PhoA^-$) and *TnphoA* fusion mutants ($PhoA^+$) were isolated and screened for the ability to invade cultured HEp-2 cells. Most invasion-negative (Inv^-) mutations mapped to a 30-kb segment of the invasion plasmid, including homologs of the *Shigella flexneri* *ipa*, *mxi*, and *spa* genes. $Inv^- PhoA^+$ fusions in the EIEC *ipaC*, *mxiG*, *mxiJ*, *mxiM*, and *mxiD* homologs and in a proposed new gene, named *invX*, located downstream of the *spa* region were identified and characterized. This analysis indicates the presence of the *ipaC*, *mxiG*, *mxiJ*, *mxiM*, *mxiD*, and *invX* gene products in the EIEC cell envelope and demonstrates a strict requirement for these genetic loci in invasion. Overall, our results suggest a high degree of genetic, structural, and functional homology between the EIEC and *S. flexneri* large invasion plasmids.

Enteroinvasive *Escherichia coli* (EIEC) comprises a group of organisms known to cause a dysentery-like enteritis in humans, similar to that caused by *Shigella* species. Experiments conducted with both chimpanzees and human volunteers by DuPont et al. (13) demonstrated that EIEC has the capacity to invade intestinal epithelial cells, multiply intracellularly, and produce a disease indistinguishable from shigellosis. Genetically, EIEC pathogenesis has been shown to be very similar to that of *Shigella* species in that it is multifactorial and requires the concerted expression of genes present both on the chromosome and on a 140-MDa plasmid (for a review, see reference 16). Expression of plasmid genes has been shown to be required and sufficient for the bacterium to invade epithelial cells in vitro (36). In *Shigella flexneri*, these genes constitute the *vir* regulon (16). The *vir* regulon includes a 30-kb DNA stretch encompassing the transcriptionally divergent *ipa* and *mxi-spa* genes (16). The *ipaBCD* genes encode immunodominant surface-associated proteins proposed to act as adhesins and/or invasins (7, 8, 26, 38, 44) and are conserved in EIEC (17, 18). Among the activities also attributed to products of the *ipa* locus is a contact hemolysin (IpaB) implicated in escape from the phagosome (20, 37). Further upstream of the *ipa* genes lie the *mxi* and *spa* genes, which are required for surface presentation and/or secretion of the Ipa and other proteins into the medium (3-6, 45). The *mxiA* gene and the *mxiB* gene (also named *spa47*) and other genes of the *mxi-spa* region have been proposed to function as a specialized export machinery for the Ipa proteins (5, 6, 45) which, despite the lack of signal sequences, are found at the cell surface (44). An additional gene, *icsB*, which lies promoter proximal within the *ipa* operon, and the unlinked *icsA* (*virG*) and *virK* genes are

required for intra- and intercellular spread (2, 9, 24, 30). The whole *vir* regulon is regulated by two transcriptional activators, the *virB* (*ipaR*) and *virF* gene products, which function in a cascade (1, 11, 32). Repression of the plasmid *vir* regulon in response to low temperature (30°C) is mediated by the chromosomal *virR* (*osmZ*) gene product (12, 21, 22, 28), a histone-like protein which itself directly modulates the expression of the plasmid transcriptional activator *virB* gene (43). Other chromosomally encoded virulence factors include the *kcpA* gene product necessary for cell-to-cell spread (31, 46) and for a positive Sereny test (14), the O side chain of lipopolysaccharide (35), and products of the *arg-mtl*-linked *iuc* locus, which encodes the aerobactin siderophore system required for intracellular iron acquisition (23).

Although bacillary dysentery due to EIEC is clinically indistinguishable from that caused by *Shigella* species (41), studies with human volunteers place the EIEC infectious dose several orders of magnitude higher than that of *Shigella* species. Whereas fewer than 500 *Shigella* organisms are sufficient to cause disease in healthy volunteers (13), more than 10^6 EIEC organisms are required. It has been suggested that this difference in infective dose may be partially due to the fact that *Shigella* species are better able to survive transit through the acidic stomach than are EIEC bacteria (15). Alternatively, shigellae may be more successful at entering into colonic cells than are EIEC bacteria. The latter view is consistent with observed differential rates of invasion in vitro (8a). Several investigations have shown that all plasmid-mediated and chromosomal invasion determinants identified in *Shigella* species are also present in EIEC (16). However, this evidence is derived from DNA sequence homologies determined by Southern blot analysis or from protein similarities identified by immunoblot rather than from detailed comparisons of DNA or protein sequences. There is considerable heterogeneity, both inter- and in-

* Corresponding author. Electronic mail address: PBAV@db1.cc.rochester.edu.

TABLE 1. Characterization of Pcr and PhoA⁺ Tn_{phoA} insertion mutants^a

Name	Phenotype		Expression of Ipa antigen			
	Pcr	Inv	B	C	D	A
Pcr mutants						
PB141	var	+/-	+	+	+	+
PB143	-	-	+	+	+	+
PB148	var	+/-	+	+	+	+
PB151	var	-	+	+	+	+
PB158	-	-	+	+	+	+
PB169	-	-	+	+	+	+
PhoA⁺ mutants						
PB120	-	-	+	+	+	+
PB124	-	-	+	+	+	+
PB182	-	-	+	+	+	+
PB200	-	-	+	+	+	+
PB223	-	-	+	+	+	+
PB225	-	-	+	+	+	+
PB228	-	-	+	+	+	+
PB233	-	-	+	+	+	+
PB234	var	-	+	+	+	+
PB255	-	-	+	+	+	+
PB257	-	-	+	-	-	-
PB265	-	-	+	+	+	+

^a Pcr phenotype, Inv phenotype, and expression of Ipa antigens were determined as described in Materials and Methods.

traspecies, in the restriction endonuclease patterns of virulence plasmids isolated from *S. flexneri* and EIEC. There are also differences in the apparent molecular weights of the Ipa antigens (8a, 18). These differences may reflect functional differences relevant to the ability to invade colonic cells, and better invasion may translate into lower infectious dose.

In this work, we have undertaken a molecular comparison of invasion determinants in EIEC and *S. flexneri* invasion plasmids. We have found that a high degree of genetic and structural similarity exists between the two species. While minor differences have been found, a more extensive analysis will be needed before we can assign to these differences a role in infective dose.

MATERIAL AND METHODS

Bacterial strains and plasmids. The isolation of plasmid pSF204 from an EIEC clinical isolate and the properties of its transconjugant HB101(pSF204) have been described elsewhere (39, 40). The mutant strains used in this study are listed in Table 1. SM10λpir(pRT733) (42) was obtained from R. Taylor, University of Tennessee.

Media and chemicals. Bacteria were routinely grown in L broth (LB). For determination of the Congo red colony phenotype (Pcr), cells were plated on Trypticase soy agar supplemented with 0.003% dye (TS-CR) as described previously (27). For isolation of Tn_{phoA} fusions, the chromogenic substrate 5-bromo-4-chloro-3-indolyl phosphate, *p*-toluidine salt (XP; Bachem), was included at 20 μg/ml. When appropriate, antibiotics were added in the following concentrations: ampicillin, 50 μg/ml; kanamycin, 40 μg/ml; and streptomycin, 25 μg/ml. HEp-2 cells were maintained at 37°C with 5% CO₂ in RPMI 1640 (Irvine Biochemical) supplemented with 5 mM glutamine (GIBCO) and 5% fetal calf serum (GIBCO).

Isolation of Tn_{phoA} insertion mutants. HB101(pSF204): Tn_{phoA} mutants were obtained by mating HB101(pSF204) with SM10λpir(pRT733) and screened directly on LB-strep-

tomycin-kanamycin-XP for PhoA activity or alternatively on TS-CR-ampicillin-kanamycin-streptomycin for Pcr phenotype. All mutants were then screened for PhoA activity, for the ability to bind Congo red, for in vitro invasion, and, in the case of active PhoA fusions, for regulation of PhoA activity by temperature.

Characterization of insertions. Presence of an insertion on the invasion plasmid was determined by examination of BamHI digests of plasmid DNA. Southern blot analysis for mapping of the plasmid Tn_{phoA} insertions was carried out by using a *phoA*-specific oligonucleotide (5'-AATATCGCCTGAGC) and inserts from two pSF204 (fragment E; see below) subclones described earlier, pRJ6 and pPS15 (39). For further analysis of the Tn_{phoA} fusion joints, pKS⁺ (Stratagene) subclones were obtained by using pSF204 restricted with BamHI or, when necessary, EcoRV and kanamycin resistance and/or PhoA⁺ for selection and screening.

Nucleotide sequence analysis. Single-strand sequence analysis was performed at the University of Rochester DNA Microchemistry Core Facility by using Tn_{phoA}-specific primers and the KS and SK primers on an automated sequencer (Applied Biosystems). The nucleotide sequences of both complementary strands of *invX* were obtained by the dideoxy method (33), using double-stranded purified plasmid DNA as the template, obtained after subcloning in Bluescript plasmids. Sequence homology searches were carried out by using software from the University of Wisconsin Genetics Computer Group. Homology searches for PB182 and PB233 were kindly performed by Claude Parsot (Institut Pasteur, Paris, France).

Other methods. The gentamicin protection assay was performed as described previously (40). Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblot analyses were carried out as described elsewhere (10, 25). The Ipa-reactive monkey serum and the rabbit-raised PhoA-specific antiserum were kind gifts from T. L. Hale (Walter Reed Army Institute of Research, Washington, D.C.) and D. Low (University of Utah), respectively.

Nucleotide sequence accession number. The nucleotide sequence of *invX* has been deposited in GenBank under accession number L18941.

RESULTS

Isolation of Pcr mutants affected in in vitro invasion. The ability of *S. flexneri* colonies to adsorb the basic dye Congo red has been correlated to virulence (27). A total of 39 Tn_{phoA} insertion mutants (PB135 to PB173) affected in the ability to bind Congo red, half of which carried the mutation on the large invasion plasmid pSF204, were isolated. Of these, six mutants were retained for detailed study (Table 1). Two different Pcr phenotypes could be distinguished. Negative (Pcr⁻) mutant colonies were undistinguishable from HB101 colonies on Congo red medium, and variant (Pcr^{var}) mutant colonies were different from both HB101 and HB101(pSF204) in either tint or color intensity; e.g., some colonies were as dark red as the parent colonies but of a different red, and others were clearly less dark than the parent colonies but not as clear as HB101 colonies. It should be noted here that while determination of the Pcr phenotype is inherently subjective, as it is affected by colony size, light, and depth of the growth medium, as well as the eyes of the reader, all Pcr phenotype evaluations were carried out in a blind fashion with regard to invasion phenotype and mapping information. None of the mutants tested were PhoA⁺ on

TABLE 2. Molecular characterization of *TnphoA* insertions and *TnphoA* fusion products

Strain	<i>Shigella</i> gene/codon no.	PhoA activity ^a (U)		Size (kDa) of fusion product	
		30°C	37°C	Predicted	Observed
HB101(pSF204)	NA ^b		2.5		
PB184(pSF204) ^c	NA	2,023	1,412		48
Group 1					
PB257	<i>ipaC</i> /121	6.7	23.5	61	66
Group 2					
PB124	<i>mxiD</i> /422	5.8	33.5	92	85
PB120	<i>mxiD</i> /210	12.4	100	71	66
PB265	<i>mxiM</i> /21	8.7	79.0	49	49
PB200	<i>mxiJ</i> /60	ND ^d	85.5	53	52
PB225	<i>mxiJ</i> /20	10.3	38.0	49	49
PB233	<i>mxiG</i> /279	3.7	39.5	78	77
PB182	<i>mxiG</i> /164	7.5	72.5	66	66
Group 3					
PB255	<i>invX</i> /36	ND	11.5	50.5	52
PB141	<i>spa</i> 24/203	ND	0.07		
PB151	<i>invX</i> /44	ND	0.5		

^a Determined as previously described (29).

^b NA, not applicable.

^c PB184 is a chromosomal mutant derivative of HB101(pSF204) which overexpresses endogenous PhoA (38a).

^d ND, not detectable.

indicator media, indicating that functional fusions were not generated by this unselective approach. These mutants all showed a decreased ability to invade HEP-2 cells, confirming the tight coupling between the ability to invade and the ability to bind Congo red. Two different invasion phenotypes (Inv) could be distinguished in addition to the parent Inv⁺ phenotype: noninvasive (Inv⁻; less than 10% of Inv⁺ survivors) and intermediately invasive (Inv^{+/-}; between 10 and 50% of Inv⁺ survivors). All of the Pcr⁻ or Pcr^{var} mutants considered here expressed a complete set of Ipa antigens (see below).

Isolation of PhoA⁺ *TnphoA* insertion mutants altered in invasion ability in vitro. Of 120 PhoA⁺ mutants isolated, 99 were tested for the ability to invade cultured HEP-2 cells. Thirty-five mutants were found to be either Inv⁻ (26 mutants) or Inv^{+/-} (9 mutants). Three PhoA⁺ mutants were hyperinvasive in vitro and will be described elsewhere (38a). Most of the Inv⁻ and Inv^{+/-} PhoA⁺ mutants (77%) carried the mutation on the invasion plasmid. Twelve of these were retained for further characterization (Table 1). These mutants expressed PhoA activity in a temperature-regulated manner (12 of 12) (Table 2 and data not shown) and a full set of Ipa antigens (11 of 12) (Table 1) and were defective in the ability to bind Congo red (12 of 12), indicating an imbalance in the mutational frequency of the Congo red genes versus the *ipa* genes with use of *TnphoA*. Overall, these findings are consistent with at least two distinct plasmid genetic determinants within the Congo red genes, one of which is required for Ipa expression.

Mapping and transcriptional orientation of *TnphoA* insertion mutations in the invasion plasmid. Restriction of pSF204 with *Bam*HI generates 14 large fragments, A through N, with estimated molecular sizes ranging from 39.5 to 2.8 kb. All plasmid *TnphoA* insertion mutations with altered in vitro invasion ability or Congo red phenotypes mapped to the contiguous A (39.5-kb) and E (27-kb) fragments (Fig. 1). Relative positions of *TnphoA* insertions on each fragment were determined by integrating information obtained from *Bam*HI restriction analysis (*TnphoA* carries a single *Bam*HI site within its sequence) and Southern blot analysis using

phoA-specific and plasmid-specific probes. The PhoA⁺ and Pcr⁻ PhoA⁻ insertion mutations considered here define a stretch of approximately 30 kb which is essential for invasion of HEP-2 cells. Within this segment, three clusters of PhoA⁺ fusions identify genetic loci encoding proteins which are exported beyond the cytoplasmic membrane: group 1, characterized by a single isolate, PB257; group 2, comprising PB124; and group 3, comprising PB234 and PB255 (Fig. 1). In contrast, Pcr⁻ PhoA⁻ insertions are less clustered. Furthermore, two of these mutations (*TnphoA*¹⁴³ and *TnphoA*¹⁶⁹) map to an area located between group 1 and group 2 mutations where PhoA⁺ mutations were not found, suggesting that these mutations may lie in genes encoding proteins which are not exported beyond the cytoplasmic membrane yet are required for entry. Another PhoA⁻ mutation (*TnphoA*¹⁴⁸) carries a deletion internal to fragment E, which does not affect the level of Ipa expression but causes a Pcr^{var} colony phenotype. All PhoA⁺ and PhoA⁻ mutations

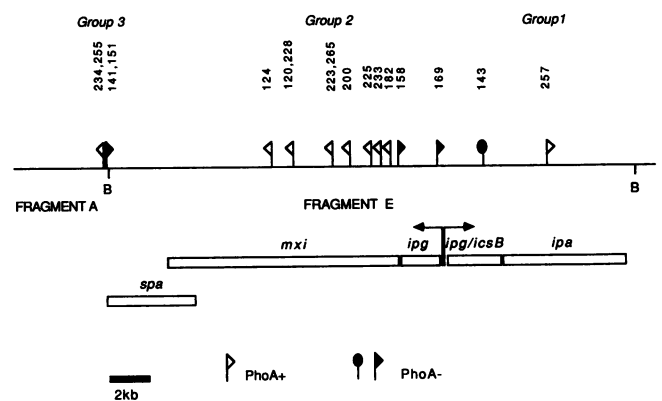


FIG. 1. Genetic map of the *spa-mxi-ipa* region of the *S. flexneri* M90T invasion plasmid pWR100 (bottom) and positions of *TnphoA* insertion mutations on the corresponding segment of the EIEC invasion plasmid derivative pSF204 (top). The three PhoA⁺ clusters are indicated.

which map between *TnphoA*¹²⁴ and *TnphoA*²⁵⁷, including group 1 and 2 mutations, were phenotypically Pcr⁻ (Table 1). In contrast, one mutation of group 3 (*TnphoA*²³⁴) and two PhoA⁻ mutations (*TnphoA*¹⁴¹ and *TnphoA*¹⁵¹) which map to the same region were phenotypically Pcr^{var}. Southern analysis of the PhoA⁺ fusions revealed two opposite transcriptional orientations, with the single fusion of group 1 running opposite those of groups 2 and 3.

Structural analysis of group 1 and group 2 *TnphoA* gene fusions and fusion products. Selected plasmid mutations were further studied with respect to the nucleotide sequence at the point of *TnphoA* insertion. Restriction fragments carrying the *TnphoA* insertion point were identified, isolated, and religated into pKS⁺, when possible using kanamycin resistance and/or PhoA activity for selection. Nucleotide sequence analysis was carried out by using a *phoA*-specific primer for *TnphoA* upstream sequences or an IS50-specific primer for downstream sequence in the case of *TnphoA*²⁵⁷. Comparison of these sequences with *Shigella* plasmid sequences revealed close to 100% homology (data not shown) and identified the *TnphoA* insertion sites (Table 2) as well as the upstream reading frame. By this analysis, *TnphoA*²⁵⁷ (group 1) lies in the EIEC *ipaC* homolog (7, 38, 44). This finding is consistent with the Inv⁻ and Pcr⁻ phenotypes of PB257 and with the cell surface association and/or excretion of the *ipaC* gene product (5, 34). Furthermore, PB257 expresses a PhoA-cross-reactive fusion product with an apparent molecular size of 66 kDa, which is in agreement with the size predicted for the *S. flexneri ipaC* gene product (Table 2). Group 2 mutations map to several *mxi* genes identified in *S. flexneri* (Table 2) (3, 4, 31a). Again, the relative map positions and transcriptional orientations of the EIEC gene homologs match those of *Shigella mxi* genes, and the observed phenotypes are consistent with those of *mxi* mutations in *Shigella* species, i.e., Inv⁻, Pcr⁻, and presence of the gene product at or near the cell surface. Apparent molecular weights of the *mxiM* (*TnphoA*²⁶⁵), *mxiJ* (*TnphoA*²²⁵ and *TnphoA*²⁰⁰), and *mxiG* (*TnphoA*¹⁸² and *TnphoA*²³³) fusion products were also in good agreement with predicted molecular weights of the *Shigella* mature gene products.

Structural analysis of a group 3 *TnphoA* gene fusion and identification of a new gene required for invasion. Surprisingly, a group 3 mutation, *TnphoA*²⁵⁵, fell within an open reading frame located beyond *spa24*, the downstream-most gene of the *mxi-spa* region of the *S. flexneri* large invasion plasmid (45). Two additional mutations selected for their Pcr^{var} phenotype, *TnphoA*¹⁴¹ and *TnphoA*¹⁵¹, also mapped to this locus. Both of these *TnphoA* insertions are in the reverse orientation from that of *TnphoA*²⁵⁵ and accordingly are PhoA⁻. Nucleotide sequence analysis downstream of the EIEC *spa24* homolog reveals a small open reading frame with two possible translational starts and a putative leader sequence, either 26 or 20 residues long, consistent with the PhoA⁺ phenotype of PB255 (Fig. 2). While it is not possible to identify the actual translational start without additional data, it is worth noting that a consensus ribosome-binding site (GGAGG) exists four nucleotides upstream of the valine codon whereas none is found in proximity to the methionine codon. Further analysis revealed that *TnphoA*¹⁴¹ lies 43 bp upstream of the termination codon of the *Shigella spa24* homolog (45), while *TnphoA*²⁵⁵ and *TnphoA*²⁵⁵ are found 125 and 148 bp downstream of the proposed GUG start of the new gene (Fig. 2). Comparison with the published *Shigella* sequence, which includes a partial sequence of the proposed gene (the first 159 nucleotides, assuming a GUG start) (45),

```

spa24                               invX
P G *                               V S D I V Y M G N K A L
TCCCTGCTTAGATGATGGAGGTAATGTGCTGATATAGTTTATATGGGTAATAAGGCTCTT 60
                                     .
Y L I L I F S L W P V G I A T V I G L T
TATTTAATCCTTATCTTTTCTTATGGCCAGTAGGTATAGCTACGGTTATGGATTAAC 120

I G L L Q T V T Q L Q E Q T L P F G I K
ATTGGTTTATTACAGACAGTGAAGTCAACTTCAAGAGCAGACACTTCTTTGGTATAAAG 180
                                     >TnphoA255
                                     <Tnpho151

L I G V S I S L L L L S G W Y G E V L L
CTTAGGTGCTCAATATCTTTGCTACTCTTTCTGGATGGTATGGTGAAGTTTATTG 240

S F C H E I M F L I K S G V *
TCTTTTGTGTCATGAATAATGTTTTTAATTAAGAGTGGGGTTGA 285

```

FIG. 2. Nucleotide sequence of the proposed *invX* gene and predicted amino acid sequence. The positions of two possible translational starts (underlined), of the putative start of the mature protein (•), and of two *invX TnphoA* insertions are indicated.

reveals near identity with the EIEC sequence, including the proposed leader sequence and leader peptidase processing site. Three substitutions, an A→T, an A→G, and a T→C at positions 117, 129, and 150, respectively (Fig. 2), generate a single conservative amino acid change: a threonine-for-serine substitution at amino acid 32 of the *Shigella* sequence. The predicted amino acid sequence of the mature gene product, assuming a processing site located as indicated in Fig. 2, is 60 amino acids long, for a calculated molecular mass of 6,528 Da. The observation by immunoblot of a mature PhoA hybrid product of the expected molecular weight and the regulation of its expression by temperature further indicate that this gene is expressed in *E. coli* (Table 2). We conclude that these mutations (*TnphoA*²⁵⁵ and *TnphoA*¹⁵¹) define a new gene whose product is a cell envelope-associated or secreted protein which is required for invasion of HEp-2 cells. An alternative possibility is that these mutations exert a negative polar effect on unidentified downstream genes which are themselves essential for invasion. *TnphoA*¹⁴¹, because of its upstream proximity to the proposed new gene start site, may also exert its phenotype through polarity on the proposed new gene or unidentified downstream genes. In view of the phenotype of the mutant, we have designated this new gene *invX*.

***ipa* gene expression in *TnphoA* insertion mutants.** Expression of the *ipa* genes was determined by immunoblot analysis using immune serum from a monkey which had recovered from a *Shigella* infection. We found minor differences in the apparent molecular weights of antigens encoded by the EIEC plasmid, pSF204, and those encoded by the plasmid of *S. flexneri* M90T, pWR100 (data not shown). As in *S. flexneri*, expression of the Ipa homologs was repressed at 30°C. The band from HB101(pSF204) corresponding to IpaB uniquely appears as a doublet (Fig. 3). However, expression of the upper band of the doublet is present at the nonpermissive temperature (Fig. 3), suggesting that this band may actually not be a posttranslational product of *ipaB*. Instead, this band may correspond to the product of the *ipaH* gene homolog, since the IpaB and IpaH proteins have similar molecular masses in *S. flexneri* (62 kDa versus 60.8 kDa) (19). Mutant PB257 (group 1) was the only mutant exhibiting an altered Ipa pattern. In this mutant, only the IpaB homolog band is being expressed. This finding is consistent with *TnphoA*²⁵⁷ lying in the *ipaC* gene homolog of EIEC and the known genetic organization of the *ipa* locus in shigellae; i.e., the observed polarity on *ipaD* and *ipaA* suggests a similar *ipaBCDA* gene order. As in shigellae, mutations which

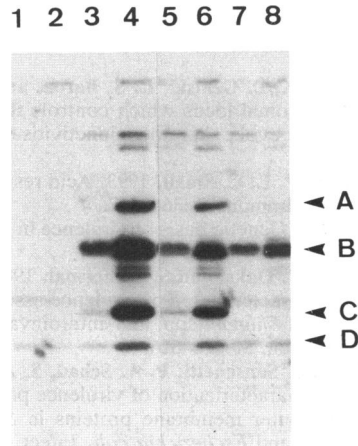


FIG. 3. Ipa antigen expression in *TnphoA* mutants. Immunoblot analysis was carried out with serum from a monkey which had recovered from shigellosis. Bacterial lysates (15 μ l) resuspended in SDS-PAGE sample buffer (200 μ l) from cultures (1.5 ml) grown to mid-log phase (optical density at 600 nm of 0.7) at 30°C (odd-numbered lanes) or 37°C (even-numbered lanes) were loaded in each lane. Lanes: 1 and 2, HB101; 3 and 4, HB101(pSF204); 5 and 6, PB182; 7 and 8, PB257. The Ipa pattern of PB182 is representative of all mutants listed in Table 1 except PB257. Positions of the IpaA (78-kDa; A), IpaB (62-kDa; B), IpaC (43-kDa; C), and IpaD (37-kDa; D) antigens are indicated.

reside in the transcriptionally divergent *mxi-spa* region of the plasmid as well as mutations in *invX* had no effect on the level of *ipa* expression in the *E. coli* background.

DISCUSSION

We have used the ability of enteroinvasive bacteria to bind Congo red and that of surface-associated proteins to form enzymatically active hybrids with alkaline phosphatase to isolate mutants of an EIEC derivative, HB101(pSF204), which are altered in the ability to invade cultured HEp-2 cells. All Pcr^- mutants isolated were less invasive than the parental strain. Expectedly, approximately half of the Pcr^- *TnphoA* insertions mapped to the large invasion plasmid, while the other half resided on the chromosome (data not shown). This is not surprising, since the Pcr^- phenotype is thought to be due to specific surface protein complexes whose structural genes lie on the large invasion plasmid and which would require chromosomally encoded proteins for synthesis, folding, secretion, and assembly. $PhoA^+$ *TnphoA* insertion mutations were unexpectedly frequently (38%) associated with loss of invasive function. This may be a reflection of the relatively high number of genes encoding secreted products which are required for invasion.

Analysis of the Pcr^- and $PhoA^+$ *TnphoA* plasmid mutants identified a 30-kb DNA segment which is required for invasion (Fig. 1). These data confirm those obtained earlier by using *Tn5* mutagenesis and complementation analysis of the same plasmid (39). Pcr^- $PhoA^-$ mutations were relatively more spread out than the $PhoA^+$ mutations, which were clustered in three distinct areas. This finding implies that plasmid genes encoding both envelope-associated proteins and cytoplasmic or inner membrane proteins are required for invasion, as has been observed in *Shigella* species. Overall, our results indicate a high degree of structural and functional homology between *Shigella* species and

EIEC in regions of the plasmid corresponding to the genes of the *ipa*, *mxi*, and *spa* loci (3, 4, 7, 38, 44, 45). Analysis of the EIEC plasmid mutations revealed near identity of the EIEC plasmid pSF204 to the *S. flexneri* plasmid pWR100 with regard to gene order, gene size, intergenic distances, transcriptional orientation, and regulation of gene expression by temperature. Furthermore, $PhoA^+$ fusions were selectively found within genes already known to encode secreted products in *S. flexneri*, namely, *ipaC* (7, 38, 44), *mxiD* (4), *mxiM* (3), *mxiJ* (3), and *mxiG* (31a). Finally, comparisons of the molecular weights of the *PhoA* hybrid polypeptide observed in the EIEC mutants with those predicted for *S. flexneri* (Table 2) and of the single-strand nucleotide sequences at the fusion sites (data not shown) revealed no significant differences between the EIEC and *Shigella* gene products considered here.

Analysis of three *TnphoA* mutations located at one end of the 30-kb plasmid segment identified a new, previously undescribed gene, *invX*, located immediately downstream of the *spa24* homolog. Comparison of the nucleotide sequence of the proposed *invX* with a stretch of *Shigella* sequence downstream of *spa24* (45) reveals near identity and conservation of the proposed reading frame. Furthermore, the primary structure of the 26 amino-terminal amino acid residues in both EIEC and *Shigella* species suggests that this region may function as a leader peptide during secretion of the protein product (Fig. 2). This is consistent with the detection by immunoblot of a *PhoA*-cross-reactive hybrid polypeptide with an apparent molecular weight of 52,000 in PB255. Taken together, these data support the existence of *invX* in both species and, with the caveat of possible polar effects toward other undetected downstream genes, suggest a role for the *invX* gene product in invasion. It is tempting to hypothesize a function for *invX* akin to that of the *mxi* and *spa* genes in view of the presence of *invX* in the same operon. In preliminary experiments, excretion of the Ipa antigens was shown to be similarly negatively affected in *mxiD*, *mxiM*, *mxiJ*, *mxiG*, *spa24*, and *invX* mutants (22a). It may nevertheless be significant that two of three mutations in the proposed *invX* gene and one immediately upstream exhibited Pcr^{var} phenotypes (Table 2). This observation distinguishes mutations in this locus from mutations in the *mxi* genes, which all produced a Pcr^- phenotype indistinguishable from that of plasmid-free HB101. Thus, *invX* mutants, in contrast to *mxi* mutants, may retain some ability to bind Congo red via unique surface properties.

In conclusion, within the limitations of our analysis and notwithstanding the possible role of point mutations, no extensive structural difference could be observed between some of the essential invasion plasmid genes of EIEC and *S. flexneri*. While this finding supports the earlier observation that the EIEC and *Shigella* plasmids are structurally and functionally equivalent, it does not solve the question of differential infectivity between the two species. In this respect, the findings that minor structural differences exist equally in all sequences analyzed in this study (data not shown) and that small molecular weight differences in the Ipa antigens can also be observed suggest that the difference between EIEC and *Shigella* infectivity may not be the result of a single quantum loss or gain of an invasion determinant. Rather, the observed difference in infectivity may result from the cumulative effects of a series of small changes in many genes, including chromosomal genes and genes of the plasmid *vir* regulon.

ACKNOWLEDGMENTS

We thank C. Parsot for performing some of the homology searches and A. Allaoui, C. Parsot, and P. Sansonetti (Institut Pasteur, Paris, France) for communicating unpublished results. We also are grateful to R. Taylor (University of Tennessee), D. Low (University of Utah), and T. L. C. Hale (Walter Reed Army Institute of Research) for gifts of strains and reagents.

This work was supported in part by National Institutes of Health grants AI26280 to P.M.B. and AI26216 to P.L.C.S.

ADDENDUM IN PROOF

A recent article (C. Sasakawa, K. Komatsu, T. Tobe, T. Suzuki, and M. Yoshikawa, *J. Bacteriol.* 175:2334–2346, 1993) reports the nucleotide sequence analysis of eight genes in region 5 of the plasmid of *S. flexneri* 2a; one of these genes, the one labeled *spa9* (ORF7), corresponds to *invX* of EIEC.

REFERENCES

- Adler, B., C. Sasakawa, T. Tobe, S. Makino, K. Komatsu, and M. Yoshikawa. 1989. A dual transcriptional activation system for the 230 kb plasmid genes coding for virulence-associated antigens of *Shigella flexneri*. *Mol. Microbiol.* 3:627–635.
- Allaoui, A., J. Mounier, M. C. Prevost, P. J. Sansonetti, and C. Parsot. 1992. *icsB*: a *Shigella flexneri* virulence gene necessary for the lysis of protrusions during intercellular spread. *Mol. Microbiol.* 6:1605–1616.
- Allaoui, A., P. J. Sansonetti, and C. Parsot. 1992. MxiJ, a lipoprotein involved in secretion of *Shigella* Ipa invasins, is homologous to YscJ, a secretion factor of the *Yersinia* Yop proteins. *J. Bacteriol.* 174:7661–7669.
- Allaoui, A., P. J. Sansonetti, and P. J. Parsot. 1993. MxiD, an outer membrane protein necessary for the secretion of the *Shigella flexneri* Ipa invasins. *Mol. Microbiol.* 7:59–68.
- Andrews, G. P., A. E. Hromockyj, C. Coker, and A. T. Maurelli. 1991. Two novel virulence loci, *mxiA* and *mxiB*, in *Shigella flexneri* 2a facilitate excretion of invasion plasmid antigens. *Infect. Immun.* 59:1997–2005.
- Andrews, G. P., and A. T. Maurelli. 1992. *mxiA* of *Shigella flexneri* 2a, which facilitates export of invasion plasmid antigens, encodes a homolog of the low-calcium-response protein, LcrD, of *Yersinia pestis*. *Infect. Immun.* 60:3287–3295.
- Baudry, B., M. Kaczorek, and P. J. Sansonetti. 1988. Nucleotide sequence of the invasion plasmid antigen B and C genes (*ipaB* and *ipaC*) of *Shigella flexneri*. *Microb. Pathog.* 4:345–357.
- Baudry, B., A. T. Maurelli, P. Clerc, J. C. Sadoff, and P. J. Sansonetti. 1987. Localization of plasmid loci necessary for the entry of *Shigella flexneri* into HeLa cells, and characterization of one locus encoding four immunogenic polypeptides. *J. Gen. Microbiol.* 133:3403–3413.
- Bavoil, P., and P. L. C. Small. Unpublished data.
- Bernardini, M. L., J. Mounier, H. D'Hauteville, M. Coquis Rondon, and P. J. Sansonetti. 1989. Identification of *icsA*, a plasmid locus of *Shigella flexneri* that governs bacterial intra- and intercellular spread through interaction with F-actin. *Proc. Natl. Acad. Sci. USA* 86:3867–3871.
- Burnette, W. N. 1981. "Western blotting": electrophoretic transfer of proteins from sodium dodecyl phosphate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal. Biochem.* 112:195–203.
- Buyse, J. M., M. Venkatesan, J. A. Mills, and E. V. Oaks. 1990. Molecular characterization of a trans-acting, positive effector (*IpaR*) of invasion plasmid antigen synthesis in *Shigella flexneri* serotype 5. *Microb. Pathog.* 8:197–211.
- Dorman, C. J., N. N. Bhriain, and C. F. Higgins. 1990. DNA Supercoiling and environmental regulation of virulence gene expression in *Shigella flexneri*. *Nature (London)* 344:789–792.
- DuPont, H. L., S. B. Formal, R. B. Hornick, M. J. Snyder, J. P. Libonati, D. J. Sheahan, E. H. LaBrec, and J. P. Kalas. 1971. Pathogenesis of *Escherichia coli* diarrhea. *N. Engl. J. Med.* 285:1–9.
- Formal, S. B., Jr., P. Gemski, L. S. Baron, and E. H. LaBrec. 1971. A chromosomal locus which controls the ability of *Shigella flexneri* to evoke keratoconjunctivitis. *Infect. Immun.* 3:73–79.
- Gorden, J., and P. L. C. Small. 1993. Acid resistance in enteric bacteria. *Infect. Immun.* 61:364–367.
- Hale, T. L. 1991. Genetic basis of virulence in *Shigella* species. *Microbiol. Rev.* 55:206–224.
- Hale, T. L., E. V. Oaks, and S. B. Formal. 1985. Identification and antigenic characterization of virulence-associated, plasmid-coded proteins of *Shigella* spp. and enteroinvasive *Escherichia coli*. *Infect. Immun.* 50:620–626.
- Hale, T. L., P. J. Sansonetti, P. A. Schad, S. Austin, and S. B. Formal. 1983. Characterization of virulence plasmids and plasmid-associated outer membrane proteins in *Shigella flexneri*, *Shigella sonnei*, and *Escherichia coli*. *Infect. Immun.* 40:340–350.
- Hartman, A. B., M. Venkatesan, E. V. Oaks, and J. M. Buyse. 1990. Sequence and molecular characterization of a multicopy invasion plasmid antigen gene, *ipaH*, of *Shigella flexneri*. *J. Bacteriol.* 172:1905–1915.
- High, N., J. Mounier, M. C. Prevost, and P. J. Sansonetti. 1992. *IpaB* of *Shigella flexneri* causes entry into epithelial cells and escape from the phagocytic vacuole. *EMBO J.* 11:1991–1999.
- Hromockyj, A. E., and A. T. Maurelli. 1989. Identification of an *Escherichia coli* gene homologous to *virR*, a regulator of *Shigella* virulence. *J. Bacteriol.* 171:2879–2881.
- Hromockyj, A. E., S. C. Tucker, and A. T. Maurelli. 1992. Temperature regulation of *Shigella* virulence: identification of the repressor gene *virR*, an analogue of *hns*, and partial complementation by tyrosyl transfer RNA (tRNA^{Tyr}). *Mol. Microbiol.* 6:2113–2124.
- Hsia, R.-C., and P. Bavoil. Unpublished data.
- Lawlor, K., and S. M. Payne. 1984. Aerobactin genes in *Shigella* spp. *J. Bacteriol.* 160:266–272.
- Lett, M. C., C. Sasakawa, N. Okada, T. Sakai, S. Makino, M. Yamada, K. Komatsu, and M. Yoshikawa. 1989. *virG*, a plasmid-coded virulence gene of *Shigella flexneri*: identification of the *VirG* protein and determination of the complete coding sequence. *J. Bacteriol.* 171:353–359.
- Lugtenberg, B., J. Meijers, R. Peters, P. van der Hoek, and L. van Alphen. 1975. Electrophoretic resolution of the "major outer membrane protein" of *Escherichia coli* K12 into four bands. *FEBS Lett.* 58:254–259.
- Maurelli, A. T., B. Baudry, H. d'Hauteville, T. L. Hale, and P. J. Sansonetti. 1985. Cloning of plasmid DNA sequences involved in invasion of HeLa cells by *Shigella flexneri*. *Infect. Immun.* 49:164–171.
- Maurelli, A. T., B. Blackmon, and R. Curtiss III. 1984. Loss of pigmentation in *Shigella flexneri* 2a is correlated with loss of virulence and virulence-associated plasmid. *Infect. Immun.* 43:397–401.
- Maurelli, A. T., and P. J. Sansonetti. 1988. Identification of a chromosomal gene controlling temperature-regulated expression of *Shigella* virulence. *Proc. Natl. Acad. Sci. USA* 85:2820–2824.
- Michaelis, S., H. Inouye, D. Oliver, and J. Beckwith. 1983. Mutations that alter the signal sequence of alkaline phosphatase in *Escherichia coli*. *J. Bacteriol.* 154:366–374.
- Nakata, N., C. Sasakawa, N. Okada, T. Tobe, I. Fukuda, T. Suzuki, K. Komatsu, and M. Yoshikawa. 1992. Identification and characterization of *virK*, a virulence-associated large plasmid gene essential for intercellular spreading of *Shigella flexneri*. *Mol. Microbiol.* 6:2387–2395.
- Pal, T., J. W. Newland, B. D. Tall, S. B. Formal, and T. L. Hale. 1989. Intracellular spread of *Shigella flexneri* associated with the *kcpA* locus and a 140-kilodalton protein. *Infect. Immun.* 57:477–486.
- Parsot, C., and P. J. Sansonetti. Personal communication.
- Sakai, T., C. Sasakawa, S. Makino, and M. Yoshikawa. 1986.

- DNA sequence and product analysis of the *virF* locus responsible for Congo red binding and cell invasion in *Shigella flexneri* 2a. *Infect. Immun.* **54**:395–402.
33. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
 34. Sankaran, K., V. Ramachandran, Y. V. Subrahmanyam, S. Rajarathnam, S. Elango, and R. K. Roy. 1989. Congo red-mediated regulation of levels of *Shigella flexneri* 2a membrane proteins. *Infect. Immun.* **57**:2364–2371.
 35. Sansonetti, P. J., T. L. Hale, G. J. Dammin, C. Kapfer, H. H. Collins, Jr., and S. B. Formal. 1983. Alterations in the pathogenicity of *Escherichia coli* K-12 after transfer of plasmid and chromosomal genes from *Shigella flexneri*. *Infect. Immun.* **39**:1392–1402.
 36. Sansonetti, P. J., D. J. Kopecko, and S. B. Formal. 1982. Involvement of a plasmid in the invasive ability of *Shigella flexneri*. *Infect. Immun.* **35**:852–860.
 37. Sansonetti, P. J., A. Ryter, P. Clerc, A. T. Maurelli, and J. Mounier. 1986. Multiplication of *Shigella flexneri* within HeLa cells: lysis of the phagocytic vacuole and plasmid-mediated contact hemolysis. *Infect. Immun.* **51**:461–469.
 38. Sasakawa, C., B. Adler, T. Tobe, N. Okada, S. Nagai, K. Komatsu, and M. Yoshikawa. 1989. Functional organization and nucleotide sequence of virulence region-2 on the large virulence plasmid in *Shigella flexneri* 2a. *Mol. Microbiol.* **3**:1191–1201.
 - 38a. Sinai, A., and P. Bavoll. Submitted for publication.
 39. Small, P. L., and S. Falkow. 1988. Identification of regions on a 230-kilobase plasmid from enteroinvasive *Escherichia coli* that are required for entry into HEP-2 cells. *Infect. Immun.* **56**:225–229.
 40. Small, P. L., R. R. Isberg, and S. Falkow. 1987. Comparison of the ability of enteroinvasive *Escherichia coli*, *Salmonella typhimurium*, *Yersinia pseudotuberculosis*, and *Yersinia enterocolitica* to enter and replicate within HEP-2 cells. *Infect. Immun.* **55**:1674–1679.
 41. Taylor, D. N., P. Echeverria, O. Sethabutr, C. Pitarangsi, U. Leksomboon, N. R. Blacklow, B. Rowe, R. Gross, and J. Cross. 1988. Clinical and microbiologic features of *Shigella* and enteroinvasive *Escherichia coli* infections detected by DNA hybridization. *J. Clin. Microbiol.* **26**:1362–1366.
 42. Taylor, R. K., C. Manoil, and J. J. Mekalanos. 1989. Broad-host-range vectors for delivery of *TnphoA*: use in genetic analysis of secreted virulence determinants of *Vibrio cholerae*. *J. Bacteriol.* **171**:1870–1878.
 43. Tobe, T., S. Nagai, N. Okada, B. Adler, M. Yoshikawa, and C. Sasakawa. 1991. Temperature-regulated expression of invasion genes in *Shigella flexneri* is controlled through the transcriptional activation of the *virB* gene on the large plasmid. *Mol. Microbiol.* **5**:887–893.
 44. Venkatesan, M. M., J. M. Buysse, and D. J. Kopecko. 1988. Characterization of invasion plasmid antigen genes (*ipaBCD*) from *Shigella flexneri*. *Proc. Natl. Acad. Sci. USA* **85**:9317–9321.
 45. Venkatesan, M. M., J. M. Buysse, and E. V. Oaks. 1992. Surface presentation of *Shigella flexneri* invasion plasmid antigens requires the products of the *spa* locus. *J. Bacteriol.* **174**:1990–2001.
 46. Yamada, M., C. Sasakawa, N. Okada, S.-I. Makino, and M. Yoshikawa. 1989. Molecular cloning and characterization of chromosomal virulence region *kcpA* of *Shigella flexneri*. *Mol. Microbiol.* **3**:207–213.