JOURNAL OF BACTERIOLOGY, June 1997, p. 3670–3675 Vol. 179, No. 11
0021-9193/97/$04.00+0

Plasmid Maintenance Functions of the Large Virulence Plasmid of *Shigella flexneri*

LYNDSAY RADNEDGE, MICHAEL A. DAVIS,† BRENDA YOUNGREN AND STUART J. AUSTIN*

Gene Regulation and Chromosome Biology Laboratory, ABL-Basic Research Program, NCI-Frederick Cancer Research and Development Center, Frederick, Maryland 21702-1201

Received 5 December 1996/Accepted 28 March 1997

The large virulence plasmid pMYSH6000 of *Shigella flexneri* contains a replicon and a plasmid maintenance stability determinant (Stb) on adjacent *Sal*I fragments. The presence of a RepFIIA replicon on the *Sal*I C fragment was confirmed, and the complete sequence of the adjacent *Sal*I O fragment was determined. It shows homology to part of the transfer (*tra*) operon of the F plasmid. Stb stabilizes a partition-defective P1 miniplasmid in *Escherichia coli*. A 1.1-kb region containing a homolog of the F *trbH* gene was sufficient to confer stability. However, the *trbH* open reading frame could be interrupted without impairing stability. Deletion analysis implicated the involvement of two small open reading frames, STBORF1 and STBORF2, that fully overlap *trbH* in the opposite direction. These open reading frames are closely related to the *vagC* and *vagD* genes of the *Salmonella dublin* virulence plasmid and to open reading frame pairs in the F *trbH* region and in the chromosomes of *Dichelobacter nodosus* and *Haemophilus influenzae*. Stb appears to promote better-than-random distribution of plasmid copies and is a plasmid incompatibility determinant. The F homolog does not itself confer stability but exerts incompatibility against the activity of the Stb system. Stb is likely to encode either an active partition system or a postsegregational killing system. It shows little similarity to previously studied plasmid stability loci, but the genetic organization of STBORF1 and STBORF2 resembles that of postsegregational killing mechanisms.

The ability of *Shigella* and enteroinvasive *Escherichia coli* species to invade epithelial cells and cause enteric disease is dependent on the presence of one of a family of large, low-copy-number plasmids collectively termed pINV (15, 29). The pINV plasmid of *Shigella flexneri* 2a YSH6000 is pMYSH6000 (29). It is about 230 kb, or roughly 5% the size of the host chromosome. Makino et al. (15) identified two adjacent *Sal*I fragments of the plasmid that are involved in plasmid maintenance. The 17-kb *Sal*I C fragment contains the replicon which can promote autonomous replication. This region hybridizes to a DNA probe from a RepFIIA (IncFII) replicon (30). Miniplasmids driven by the *Sal*I C fragment were unstably maintained, but improved stability was obtained when the adjacent 5.7-kb *Sal*I O fragment was included. Thus, *Sal*I-C contains a replicon, and *Sal*I-O contains an ancillary stability function termed Stb (15). In this study, we subcloned and sequenced the relevant regions, confirmed the identity of the replicon, and identified the minimal stability region. We describe the properties of these systems in plasmid constructs replicating in *E. coli.*

MATERIALS AND METHODS

Media and chemicals. Media, reagents, enzymes, buffers, and chemicals were as previously described (1). Enzymes were obtained from New England Biolabs (Beverly, Mass.) and Boehringer Mannheim (Indianapolis, Ind.) and used under conditions recommended by the manufacturers. M63-minimal glucose plates and liquid medium (21) contained 0.2% glucose, 5 μg of thiamine per ml, and 50 μg each of methionine, tryptophan, and leucine per ml. Ampicillin was used at 100 μg/ml, chloramphenicol was used at 10 μg/ml (6 μg/ml in M63), and spectinomycin was used at 25 μg/ml. Cloning methods were as described by Sambrook et al. (27).

Bacterial strains. The *E. coli* strains used were BR825 (*trp polA*::Tn*10*) (17), BR2846 (*supE44 hsdR17 recA1 endA1 gyrA96 thi relA* Δ*lacU169*), which was derived from DH5 (9), and N100 (6).

Plasmids. The mini-P1 test vector pALA136, containing the replication regions of P1 and pBR322 with a chloramphenicol resistance cassette, has been previously described (18). Two plasmids were kindly supplied by S. Makino and M. Yoshikawa: pCS1367, containing the *Sal*I C and O fragments of the virulence plasmid pMYSH6000 from *S. flexneri* 2a YSH6000 in pBR322, and pMYSH6609, containing the *Sal*I O fragment alone in pBR322 (15). The *Sal*I O fragment was inserted into the *Sal*I site of pALA136 to make pALA1184. Plasmid pALA1194 was made by deleting the *Eco*RV-*Eco*RV fragment of pALA1184 and inserting an oligonucleotide containing sites for *Eco*RV, *Kpn*I, and *Bam*HI. Plasmids pALA1132, pALA1133, and pALA2500 through pALA2506 are deletion derivatives of pALA1194 generated by using the double-stranded nested deletion kit as directed by the manufacturer (Pharmacia Biotech, Piscataway, N.J.). Plasmid pALA1101 was constructed as follows. pALA801 (8) was cut with *Hin*dIII, and the overhanging ends were filled in with Klenow polymerase. The DNA was ligated to the linker sequence CGGATCCG, thus converting the *Hin*dIII site to a *Bam*HI site. Plasmid pALA1557 was made by inserting the *Bam*HI-*Bam*HI fragment of pALA1101 into the *Bam*HI site of pALA136. Plasmid pALA1580 was made by inserting the *Eco*RV-*Sph*I fragment of pKI288 (16) into the same sites of pALA136. The F segment contains the F *trbH* gene and 766 bp of upstream and 220 bp of downstream F sequences. Plasmid pALA1581 was made by inserting the *Eco*RV-*Bam*HI fragment of pKI288 into the same sites of pALA136. In addition to the *trbH* open reading frame, it has 766 bp of upstream and 1,762 bp of downstream F sequences. Plasmid pALA1196 was made by ligating the *Sma*I-*Sal*I backbone of pALA1133 with the *Sma*I-*Sal*I fragment of pALA1132. The pMYSH6000-derived sequences of pALA1196 were amplified by PCR. The primers used contained *Bam*HI sites, and the resulting fragment was cut with *Bam*HI and inserted into the *Bam*HI site of pALA136 to yield plasmid pALA2515. Plasmid pAL1573 is identical to pALA2515 except for a single base change that creates an amber codon in pMYSH6000 *trbH*. The amber mutation was created by using a modified strand overlap extension PCR using *Bam*HI primers and two complementary mutagenic primers (25). The λ-P1 miniplasmid λ-P1:5RCm was as described (3).

Partition tests. Partition tests were performed as described by Martin et al. (18). Briefly, the mini-P1 derivatives to be tested were transformed into BR825, selecting on L plates with chloramphenicol or M63-glucose-chloramphenicol plates. Twelve transformant colonies were purified by restreaking once on the same medium, and then one colony from each streak was restreaked on L or M63-glucose plates. Growth of these colonies provided the 25 generations of nonselective growth during which plasmid loss occurs. To determine the proportion of cells in these colonies which still contained the mini-P1 plasmid, the colonies were restreaked again on L or M63-glucose plates, and eight colonies from each of the 12 restreaks were picked to L and L-chloramphenicol plates or

* Corresponding author. Phone: (301) 846-1266. Fax: (301) 846-6988. E-mail: AUSTIN@NCIFCRF.GOV.

† Present address: Department of Microbiology and Immunology, Uniformed Services University of the Health Sciences, Bethesda, MD 20814-4799.

TABLE 1. Plasmid stability determinants assayed by using a P1 miniplasmid[a]

Plasmid	Stability locus	% Retention (25 generations)	
		L broth	Minimal medium
pALA136	None	4	4
pALA1184	pMYSH6000 *Sal*I O fragment	83	90
pALA1557	P1 *par*	80	93
pALA2515	pMYSH6000 1,127-bp minimal stability region	95	97
pALA1573	pMYSH6000 1,127-bp amber mutant	97	96
pALA1196	pMYSH6000 1,127-bp minimal stability region	92	99
pALA1580	F *trbH Eco*RV-*Sph*I	2	<2
pALA1581	F *trbH Eco*RV-*Bam*HI	3	NT

[a] The plasmids listed were introduced into the *polA* strain BR825, and partition tests were done as described in Materials and Methods. Retention refers to the percentage of cells containing the test plasmid after 25 generations of unselected growth. pALA2515 carries the minimal stability region on a *Bam*HI fragment; pALA1196 carries the minimal stability region on an *Eco*RV-*Sph*I fragment. NT, not tested.

to M63-glucose plates without and with chloramphenicol. The proportion of colonies growing both on the antibiotic-free and chloramphenicol medium was taken as the proportion of cells retaining the plasmid in the nonselectively grown colonies.

Incompatibility tests. Strain BR825 was transformed with plasmid pCS1367, selecting for ampicillin resistance. Competent cells were prepared and transformed with pALA136 or its derivatives, selecting for chloramphenicol-resistant colonies at 37°C. Eight colonies were streaked for single-colony growth on L agar with chloramphenicol. Ten colonies from each resultant streak were picked with sterile toothpicks and stabbed sequentially to L agar with ampicillin and L agar with chloramphenicol. The proportion of chloramphenicol-resistant colonies that were ampicillin sensitive was taken as the proportion of cells that had lost the resident pCS1367 plasmid in the approximately 25 generations of growth of the initial colonies formed on transformation.

DNA sequencing and analysis. The sequence of the *Sal*I O fragment was determined by cycle sequencing, using a Prism Ready Reaction Dyedeoxy Terminator kit in conjunction with an Applied Biosystems model 373A DNA sequencer. DNA from plasmid pCS1367 or pMYSH6609 was used as a template. Sequence information was obtained from both strands, with each segment of the sequence obtained from at least two different primers. The DNA sequence was assembled by using Sequencher 3.0 (Gene Codes Corp., Ann Arbor, Mich.). Sequence analysis was carried out by using the Genetics Computer Group (Madison, Wis.) package, version 8.0. The sequence of the F *trbH* region was obtained as described above, using plasmid pKI288 DNA (16) as the template.

Determination of average plasmid copy number. The average copy number of pALA136 and its derivatives in strain BR825 was found by comparison with the P1 miniplasmid λ-P1:5RCm, whose copy number has previously been determined (2, 3). The method used is similar to that described previously (2). Briefly, a culture of cells to be tested was grown to a fixed optical density (optical density at 600 nm of 0.2). An aliquot of each harvested culture was streaked on the appropriate nonselective agar, and the proportion of chloramphenicol-resistant colonies that resulted was determined by stabbing onto plates with and without chloramphenicol. The test cultures were mixed in equal volume with an equivalent culture of N100 λ-P1:5RCm cells grown in M63-minimal glucose medium. The mixed cultures were harvested by centrifugation, and the plasmid DNA was extracted by the alkaline lysis procedure, digested with endonuclease *Eco*RI, and displayed by gel electrophoresis and ethidium bromide staining (27). Photographed bands were quantitated with an optical scanner, and the peaks were assigned to λ-P1:5RCm or pALA136 or its derivative by comparing their positions with those derived from digests of DNA containing the relevant plasmids alone. By plotting the intensities of the λ-P1:5RCm fragment peaks against their positions in the gel, a standard curve was obtained and was used to correct the intensity of each test peak to compensate for differences in the sizes of the fragments. The molar ratios of the test and λ-P1:5RCm plasmids were thus obtained. The number of cells present was determined by counting a fixed volume of culture under the microscope, using a Petroff-Hauser counting chamber. Given the copy number per average cell of λ-P1:5RCm obtained under identical conditions (2), the copy number of pALA136 and its derivatives per average cell was determined. The average copy numbers were then corrected for the proportion of plasmid-containing cells present to give the average copy number of the test plasmid in those cells that contained a plasmid.

Nucleotide sequence accession number. The complete sequence of the pMYSH6000 *Sal*I O fragment is available from GenBank as accession no. U82621.

RESULTS

Stabilization of a low-copy-number test plasmid by the *Sal*I O fragment of the *Shigella* virulence plasmid. Plasmid pALA136 is a vector that contains two functional replicons, one derived from pBR322 and the other derived from P1. In wild-type strains, it replicates stably at high copy number via the pBR322 origin. In a *polA* strain, the pBR322 origin is inactive and the plasmid replicates at low copy number from the P1 origin (18). As the plasmid lacks the P1 *par* locus required for stable maintenance at low copy number, it is unstable in the *polA* strain in the absence of selection (18) (Table 1). Introduction of the *Sal*I O fragment of pMYSH6000 into pALA136 gave rise to a plasmid (pALA1184) that is much more stable under the same conditions (Table 1). Thus, in addition to conferring stability on plasmids driven by the pMYSH6000 replicon (15), the Stb function of the *Sal*I O fragment is capable of stabilizing a P1 replicon. The *Sal*I-O stabilizing function is therefore not replicon specific. The degree of stability conferred on pALA136 by the *Sal*I O fragment was similar to that conferred by a fragment containing the P1 partition locus, P1 *par* (pALA1557 [Table 1]).

Sequence analysis of the *Shigella Sal*I O and *Sal*I C fragments. We determined the complete sequence of the *Sal*I O fragment. Figure 1 shows a graphic representation of the sequence and depicts the regions with similarity to known genes. The fragment does not contain any region with significant homology to plasmid partition loci of known sequence. However, nearly all of the sequence elements present are homologs of known genes in other bacterial systems (Fig. 1).

The *msbB*-like open reading frame is of unknown function but shows a 70% DNA identity with the carboxy-terminal 229 amino acids of *msbB*, a nonessential chromosomal gene of *E. coli*. The *Sal*I site at the end of the sequenced fragment is within the gene, so that the *msbB* homology may continue upstream in the intact virulence plasmid. The MsbB gene

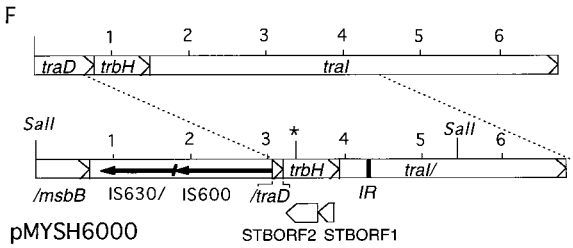

FIG. 1. A graphic representation of the F *traI* region sequence of Bradshaw et al. (5) is shown above a similar representation of the sequence of the pMYSH6000 *Sal*I O fragment and part of the adjacent *Sal*I C fragment to its right. The sequence positions (in kilobases) are shown above the maps, with bp 1 as the first base pair of the F sequence or the first base pair of the *Sal*I O fragment of pMYSH6000. The homologous regions are joined with dotted lines. The deduced amino acid sequence of the pMYSH6000 *msbB*-like open reading frame (bp 1 to 689 of the *Sal*I O fragment) is 70% identical to the carboxy-terminal 229 amino acids of *E. coli msbB* and presumably extends beyond the *Sal*I site. The regions of Tra amino acid homology (bp 3040 to 6800) show 93 to 97% identity to their F counterparts. The *traD* homolog of pMYSH6000 is truncated at bp 3040, the equivalent of F *tra* bp 1991 (5), with the amino-terminal end of the gene replaced by IS*600* sequences. The *trbH* open reading frame is maintained in pMYSH6000. The *traI* open reading frame of pMYSH6000 is interrupted by an inverted repeat sequence (*IR*) that puts the downstream portion out of frame. The pMYSH6000 *traI* homology may extend rightward beyond the region sequenced here. Slashes to the left or right of a gene or region show that the element is truncated on that side.

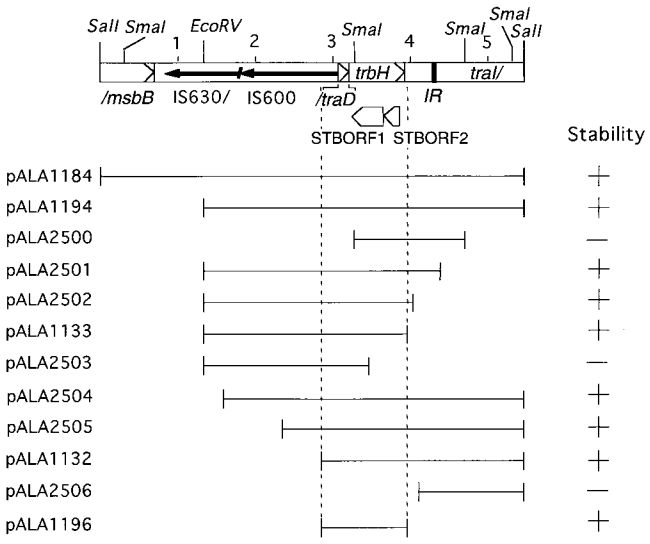

FIG. 2. Stability of deletion derivatives of a P1 miniplasmid carrying the *Sal*I O fragment. The map above shows the loci present in the starting plasmid. The bars below show the *Sal*I-O material remaining in the deletion derivatives. The endpoints of most of the exonuclease-generated deletions are approximate, as they were estimated from the apparent sizes of restriction fragments determined by gel electrophoresis. The deletion endpoints of pALA1132, pALA1133, and pALA1196 were precisely determined by DNA sequencing. The symbols − and + denote derivatives giving less than 10% and more than 80% retention in 25 generations in the stability tests as described in Materials and Methods. *IR*, inverted repeat.

product is thought to play a role in outer membrane structure or function (12).

Immediately downstream of the *msbB* homolog, there are two insertion sequences, IS*630* and IS*600*. Both are common insertion sequences in *Shigella* species (19). IS*600* appears to be inserted within the IS*630* sequence. IS*600* flanks a sequence which is almost identical to the 3′ end of the *traD* gene of the F plasmid transfer operon (5). In the F plasmid, *traD* appears in the transfer operon immediately upstream of *trbH* and *traI*. The pMYSH6000 *trbH* homolog is intact, and its predicted amino acid sequence is 95% identical to that of F *trbH*. The sequence homology between *Shigella* and F continues through the *traI* gene, whose 5′ portion extends into the adjacent *Sal*I C fragment. The homology to F *tra*I is interrupted by a deletion of 7 bp which causes the remaining downstream portion of the gene to be out of frame. This deletion creates a perfect 17-bp inverted repeat with a 4-bp loop in the *Shigella* sequence (Fig. 1).

We have obtained sequence data for approximately 2 kb of the central region of the 17-kb *Sal*I C fragment (data not shown). It includes elements of a replicon nearly identical to that of the IncFII (RepFIIA) plasmid NR1/R100 (26). The sequenced portion includes the 261-bp sequence identified by Silva et al. (30) as encoding a small regulatory RNA of the RepFIIA type. NR1/R100 has a *tra* region that closely resembles that of F and the *Sal*I-O *tra* region, except that it lacks a *trbH* homolog (33). The relative positions of the replicon and the *tra*I sequences within the *Sal*I C fragment are the same as those of NR1/R100 (data not shown), suggesting that the organization of the NR1/R100 region spanning the last part of the *tra* operon and the replicon is conserved in pMYSH6000.

Deletion mapping of the minimal stabilization locus. A series of deletions from either end of the *Sal*I O fragment were made and tested to identify those sequences responsible for the Stb plasmid stabilization phenotype. Figure 2 shows that removing about 1,200 bp from the left end, including the *msbB* homolog, leaves the Stb activity intact. The resulting plasmid, pALA1194, was the starting material for the isolation of a number of exonuclease-generated deletions. The extent of the sequences remaining in the deletion derivatives and their stability phenotypes are shown in Fig. 2. The results indicate that the Stb determinant maps to a small segment of the *Sal*I O fragment that is common to stable deletion isolates pALA1133 and pALA1132 (Fig. 2). Using these isolates, we constructed pALA1196, which retains only this small segment. pALA1196 retained the Stb activity (Fig. 2). The deletion endpoints of pALA1196 were determined by DNA sequencing. The minimal Stb region contains 1,127 bp from the *Shigella* plasmid. It contains the entire pMYSH6000 *trbH* region (Fig. 2). Also included are the short carboxy-terminal portion of the *traD* open reading frame and a small section of IS*600*. On further analysis, we found that the 1,127-bp sequence has two additional open reading frames (STBORF1 and STBORF2) that would be transcribed from the opposite strand and which are completely overlapped by the *trbH* homolog (Fig. 1 and 3). These small open reading frames are also conserved in F *trbH* (Fig. 3) (11), where they have been referred to as TRAORF1 and TRAORF2 (13).

The pMYSH6000 *trbH* open reading frame is not required for Stb function. We introduced into the 1,127-bp fragment a mutation that creates a nonsense codon early in the pMYSH6000 *trbH* open reading frame which changes the 27th *trbH* codon, CAG (glutamine), to TAG (amber) (Fig. 3). The corresponding change in the STBORF2 codon (CTG to CTA) does not change the leucine (amino acid 110) in the putative gene product of STBORF2 (Fig. 3). The plasmid carrying this mutation, pALA1573, is no less stable than its wild-type equivalent (pALA2515) in the *sup*$^+$ (nonsuppressing) strain BR825 (Table 1). We conclude that if pMYSH6000 *trbH* encodes a protein product, it is not needed for Stb function. The presence of the complementary-strand STBORF1 and STBORF2 open reading frames is, however, likely to be significant. Deletion of a small carboxy-terminal portion of STBORF1 (pALA2500) abolishes Stb function, as does deletion of STBORF2 and part of STBORF1 (pALA2503 [Fig. 2]). Not only are these open reading frames highly conserved between F and pMYSH6000, but pairs of very similar open reading frames are found in the plasmids and chromosomes of disparate bacteria. Results of an alignment of the deduced amino acid sequence of the *Shigella* STBORF1-STBORF2 region with related sequences in other organisms are summarized Table 4 (see Discussion).

Stabilization tests with the F *trbH* region. Plasmid pALA1580 consists of the F *trbH* gene with 766 bp of flanking F sequences upstream and 220 bp downstream inserted into the test vector pALA136 (see Materials and Methods). The resulting plasmid was tested for stability and found to be as unstable as its pALA136 parent (Table 1). The same is true of plasmid pALA1581, which carries the same F *trbH* region with additional F sequences downstream (Table 1). Both plasmids carry the region equivalent to the 1,127-bp *Shigella* pMYSH6000 fragment in pALA1196 (Table 1). Thus, F *trbH* does not appear to encode a stabilization function similar to that of pMYSH6000 Stb. We sequenced the F *trbH* region of pKI288 used to construct both pALA1580 and pALA1581 and found no differences from the published sequence for this region of the F plasmid (5) (Fig. 3). Comparison of the DNA and amino acid sequences of the pMYSH6000 1,127-bp region and the equivalent F region (Fig. 3) shows that they are nearly identical. This close relationship includes the peptide sequences of the *trbH* open reading frames (not shown),

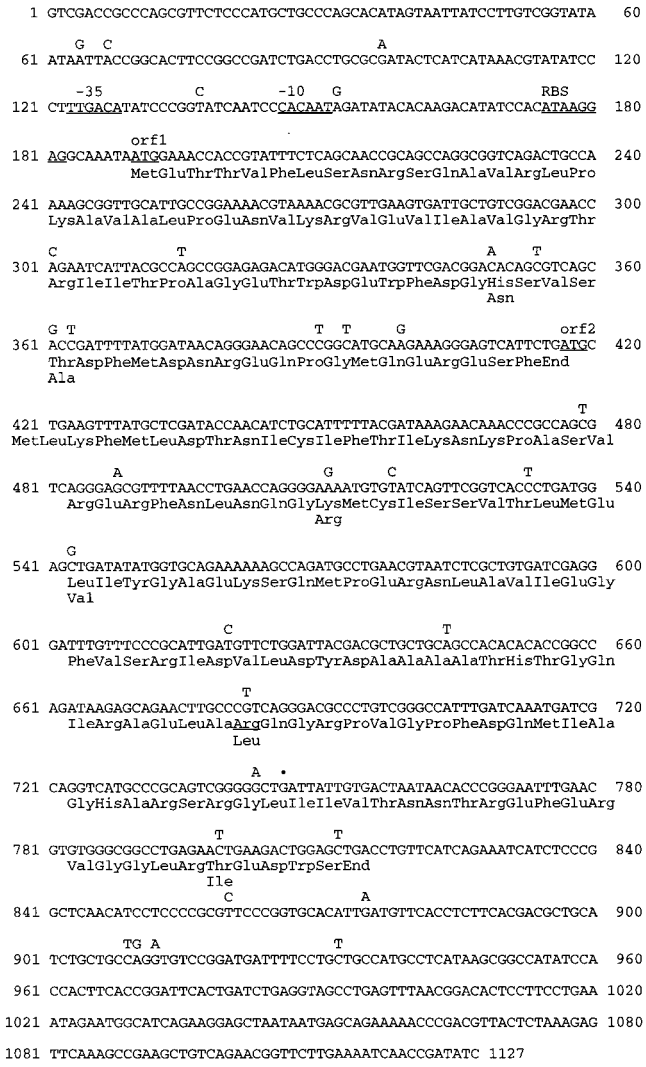

FIG. 3. DNA sequence of the pMYSH6000 *stb* region. The sequence is displayed in the opposite orientation to the maps in Fig. 1 and 2, from the *Sph*I site to the *Eco*RV site. The deduced STBORF1 and STBORF2 peptide sequences are shown below the DNA sequence, with their start points labeled orf1 and orf2. Letters immediately above the DNA sequence show those bases that differ in the homologous region of the F plasmid, and the three-letter codes directly below the peptide sequence show those amino acids that differ in the deduced F TRAORF1 and TRAORF2 peptide sequences. All of the amino acid differences between the pMYSH6000 and F sequences represent changes to similar amino acids except for the arginine (underlined)-to-leucine change at amino acid 89 of STBORF2. The dot above base 746 denotes that the base changed to an A in plasmid pALA1573. This does not alter the amino acid sequence of STBORF2 but creates an amber mutation in pMYSH6000 *trbH* that is transcribed from the opposite strand. The −10 and −35 sequences (underlined) for the putative *stb* promoter are shown, as is the likely ribosome binding site (RBS) for STBORF1.

STBORF1 and STBORF2, and the putative promoter for the *stb* region of the two species (Fig. 3). The inability of the F region to act as a stability determinant is presumably due to one or more of the few sequence differences between the F and the equivalent pMYSH6000 sequences.

Stability of plasmids carrying the 1,127-bp stability region as a function of plasmid copy number. Knowledge of the average plasmid copy number is critical to assessment of the mechanism of a plasmid stabilization function (22). The copy number of a mini-P1 plasmid is dependent on growth rate. For the standard P1 miniplasmid λ-P1:5RCm, there are approximately two copies per average cell in minimal medium and approximately eight per average cell in L broth (2, 3). The copy numbers of the relevant pALA136-derived plasmids in BR825 are similar to those of λ-P1:5RCm (Table 2), and the derivative encoding pMYSH6000 Stb does not appear to have a higher copy number than the pALA136 vector. From this it can be concluded that Stb does not stabilize the plasmid by encoding an additional replicon or by interfering with the copy control mechanism of the P1 replicon to promote higher copy numbers. Table 1 compares the stability of the constructs in cells grown in minimal and complete media. The *Shigella* stabilization locus appears to be as efficient as the P1 active partition locus, even at the very low copy numbers present in minimal medium. The stability achieved (0.1% plasmid loss per generation at a copy number of approximately two per average cell [Tables 1 and 2]) is much better than would be predicted for random distribution of the available copies (22).

TABLE 2. Copy number determinations

Plasmid	Stability locus	Copy no./avg cell[a]	
		L broth	Minimal medium
pALA136	None	12	3
pALA1196	pMYSH6000 1,127-bp minimal stability region	12	2
pALA1557	P1 *par*	7	3

[a] Determined by comparing the amount of plasmid recovered from test cells with that of the standard P1 miniplasmid λ-P1:5RCm in cells grown at 32°C in minimal glucose medium when test and standard cultures were mixed (see Materials and Methods). The values represent the average numbers of copies per plasmid-containing cell in a culture in exponential growth.

Incompatibility properties of the 1,127-bp stability region and its F homolog. Plasmid pCS1367 contains the *Sal*I C and O fragments of pMYSH6000 in a pBR322 vector. This plasmid was maintained stably in the *polA* strain BR825, where replication from the pBR322 origin is blocked (Table 3). We presume that under these conditions, plasmid replication proceeds from the IncFII replicon in *Sal*I-C and is relatively stable due to the presence of the Stb function encoded by *Sal*I-O. The mini-P1 plasmid pALA136 does not exert incompatibility against pCS1367 (Table 3), nor does a stable pALA136 derivative containing the P1 *par* locus (pALA1557 [Table 3]). However, when the derivatives of pALA136 carrying either the *Sal*I O fragment or its 1,127-bp subfragment (pALA1184 or pALA1196) were introduced into the same cell, the resident pCS1367 was lost from many of the cells (Table 3). This result shows that the 1,127-bp *stb* region contains an incompatibility

TABLE 3. Incompatibility properties of the Stb region[a]

Incoming plasmid	Stability region carried	Loss of resident pCS1367 plasmid in 25 generations (%)
pALA136	None	3
pALA1184	pMYSH6000 *Sal*I-O	63
pALA1196	pMYSH6000 1,127-bp minimal stability region	58
pALA1580	F *trbH Eco*RV-*Sph*I	45
pALA1581	F *trbH Eco*RV-*Bam*HI	38
pALA1557	P1 *par*	6

[a] BR825 *polA* cells carrying the resident pCS1367 plasmid were transformed with the plasmids shown, and loss of the resident plasmid was determined as described in Materials and Methods.

determinant, presumably exerting incompatibility by interfering with the function of the same *stb* region in pCS1367. When pALA136 carried the F *trbH* fragment, it also exerted incompatibility against CS1367 (pALA1580 and pALA1581 [Table 3]). Thus, one or more of the few sequence differences between the pMYSH6000 Stb region and its F homolog (Fig. 3) appear to have compromised the overall function of the F locus without destroying the incompatibility determinant.

DISCUSSION

The adjacent *Sal*I O and C fragments of pMYSH6000 contain an IncFII replicon and part of a transfer operon. These sequences are very similar to their NR1/R100 homologs, which are in the same relative positions in the plasmid. The *Shigella* plasmid clearly derived its *tra* segment from the same ancestral source as NR1/R100 and F. It is disrupted by insertion elements and a small deletion. This accounts for the fact that pMYSH6000, like other pINV plasmids, is not proficient for autonomous conjugal transfer (15, 28). The *trbH* sequence may have been present in this common progenitor and was subsequently lost from NR1/R100, or the sequence may have been acquired by some more recent progenitor to pMYSH6000 and F by an insertion event.

We have mapped the plasmid stability function of pMYSH6000 to a region with unusual organization. It begins with a small segment of IS*600* and a small carboxy-terminal fragment of the *traD* open reading frame. These sequences are not sufficient to promote plasmid stability (pALA2503 [Fig. 2]), and as they have known functions not concerning plasmid stability, it is unlikely that they are part of the stability determinant. The bulk of the region consists of the 240-codon F *trbH* homolog which completely overlaps the 75-amino-acid-encoding STBORF1 and 133-amino-acid-encoding STBORF2 open reading frames that are transcribed from the opposite strand. The putative pMYSH6000 *trbH* product was not needed for Stb function. Its role, if any, remains unknown. The deduced pMYSH6000 and F *trbH* products have no homologs among proteins whose sequences are currently available in protein databases (data not shown). TrbH is not required for F conjugal DNA transfer (16), and the transfer region of the conjugal plasmid NR1, which is very similar to that of F (31), does not contain a *trbH* gene (33). Perhaps there is no function for the protein, and the integrity of the *trbH* open reading frame has been preserved in order to maintain translational coupling within the *tra* operon.

STBORF1 and STBORF2 are preceded by putative promoter and ribosome binding sites. Deletions into the open reading frames from either side destroy Stb function. Thus, it is likely that proteins are produced from these open reading frames and that they are involved in the activity of Stb. This hypothesis is strengthened by the observation that pairs of very similar open reading frames are found in several other systems. One of these, the *vagC* and *vagD* genes from the *Salmonella dublin* virulence plasmid, has been indirectly implicated in plasmid maintenance (24). The *vapB* and *vapC* genes of *Dichelobacter nodosus*, although first found as chromosomal genes (13), have also been found on a plasmid (4). The existence of such highly related open reading frame pairs in such different bacterial species suggests either an ancient origin and strong genetic conservation for these genes or that they have been widely distributed among different species by some lateral transfer mechanism. The sequences described in Table 4 form a tight family group, both in the primary sequence of the putative proteins and in the organization of the open reading frames. In every case, the likely start codon for the STBORF2 homolog overlaps the stop codon of the STBORF1 homolog by one or four bases (Table 4). Unlike the case for pMYSH6000 and F, none of the other open reading frame pairs described in Table 4 lie within a reverse open reading frame which is equivalent to the *trbH* gene.

TABLE 4. *stb* homologs form a family of open reading frame pairs[a]

Species; ORF pair	% Similarity (% identity) to *S. flexneri* sequence		Overlap
	1st ORF	2nd ORF	
Shigella flexneri; STBORF1, STBORF2	100.0 (100.0)	100.0 (100.0)	−1
Escherichia coli (F); TRAORF1, TRAORF2	98.7 (97.4)	98.5 (97.0)	−1
Haemophilus influenzae			
Hi0321, Hi0322	56.2 (28.8)	58.3 (37.1)	−4
Hi0948, Hi0947	66.7 (44.0)	78.0 (59.9)	−1
Dichelobacter nodosus; VapB, VapC	51.4 (32.4)	62.9 (40.9)	−4
Salmonella dublin; VagC, VagD	48.7 (27.0)	58.7 (40.6)	−4

[a] The deduced amino acid sequences from the *S. flexneri* pMYSH6000 STBORF1 and STBORF2 open reading frames (ORFs) were aligned with their homologs from the F plasmid (TRAORF1, TRAORF2), from two loci from the chromosome of *Haemophilus influenzae* (Hi0321, Hi0322 and Hi0948, Hi0947 [7]), from the *vapB*/*vapC* locus of *Dichelobacter nodosus* (VapB, VapC [13]), and from the *vagC*/*vagD* locus of the virulence plasmid of *Salmonella dublin* (VagC, VagD [24]). In the case of the VagD open reading frame, we have inserted a T-A at bp 1141 in the sequence of Pullinger and Lax (24), which is necessary to maintain the reading frame. We speculate that this position, or some base at or near it, may have been omitted previously due to a sequencing error. The similarities were found by using the BLAST program of the Genetics Computer Group package. In each case, the STBORF1 homolog and the equivalent STBORF2 homolog are found together, with the stop and start codons overlapping by 1 or 4 bp (last column). The STBORF2 homolog is always in the −1 reading frame register with respect to the STBORF1 homolog. The similarities between Hi0321 and VagC, Hi0322 and VapC, Hi0947 and VapC, and Hi0948 and the F TRAORF2 were previously noted (7), as was the similarity between the *vap* and F open reading frames (13).

What is the nature of the Stb stability function present on the 1,127-bp fragment of pMYSH6000? Evidence has been presented that the system promotes a better-than-random distribution of the available copies to daughter cells. Of several different classes of plasmid stability functions, only active partition systems such as F *sop* and P1 *par*, and postsegregational killing systems such as F *ccd* and P1 *doc*/*phd*, are known to be able to do this (22, 32). The Stb system also exerts incompatibility against a coresident plasmid containing Stb. This is also typical of both partition and postsegregational killing systems. In the former case, the incompatibility is determined by the *cis*-acting partition site; in the latter, it is determined by the antidote protein that blocks postsegregational killing (22, 32). The F Stb homolog has retained the incompatibility determinant but has lost some other sequence required for the proper function of Stb by mutation. We are investigating the possibility that the single nonconservative change at amino acid 89 of STBORF2 (Fig. 3) is responsible for this loss of Stb function in F.

The *Shigella* Stb sequence shows little or no homology with active partition systems or postsegregational killing systems. However, the general organization of the STBORF1-STBORF2 region resembles that of postsegregational killing systems, which have similar-size products and similar overlapping or very closely apposed start and stop codons. We are currently testing the pMYSH6000 Stb system for a postsegregational killing activity.

ACKNOWLEDGMENTS

We are grateful for the expert assistance of Marilyn Powers and Colleen Rumbenhouer in synthesizing oligonucleotides and operation of the automated sequencing machine. We thank S. Makino and M. Yoshikawa for plasmids containing the pMYSH6000 sequences and T. Maurelli for introducing us to pINV plasmids and for encouragement and advice. We acknowledge the National Cancer Institute for allocation of computing time and staff support at the Frederick Biomedical Supercomputing Center of the NCI-Frederick Cancer Research and Development Center.

This research was sponsored by the National Cancer Institute, DHHS, under contract with ABL.

REFERENCES

1. **Abeles, A.** 1986. P1 plasmid replication: purification and DNA-binding activity of the replication protein RepA. J. Biol. Chem. **261:**3548–3555.
2. **Abeles, A. L., L. D. Reaves, B. Youngren-Grimes, and S. J. Austin.** 1995. Control of P1 plasmid replication by iterons. Mol. Microbiol. **185:**903–912.
3. **Austin, S. J., and B. G. Eichorn.** 1992. Random diffusion can account for *topA*-dependent suppression of partition defects in low-copy-number plasmids. J. Bacteriol. **174:**5190–5195.
4. **Billington, S. J., M. Sinistraj, M. Cheetham, B. F. Ayres, E. K. Moses, M. E. Katz, and J. I. Rood.** 1996. Identification of a native *Dichelobacter nodosus* plasmid and implications for the evolution of the *vap* regions. Gene **172:** 111–116.
5. **Bradshaw, H. D., Jr., B. A. Traxler, E. G. Minkley, Jr., E. W. Nester, and M. P. Gordon.** 1990. Nucleotide sequence of the *traI* (helicase I) gene from the sex factor F. J. Bacteriol. **172:**4127–4131.
6. **Das, A., D. Court, and S. Adhya.** 1976. Isolation and characterization of conditional lethal mutants of *Escherichia coli* defective in transcription termination factor Rho. Proc. Natl. Acad. Sci. USA **73:**1959–1963.
7. **Fleischmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J. F. Tomb, B. A. Dougherty, J. M. Merrick, K. McKenny, G. Sutton, W. FitzHugh, C. A. Fields, J. D. Gocayne, J. D. Scott, R. Shirley, L.-I. Liu, A. Glodek, J. M. Kelley, J. F. Weidman, C. A. Phillips, T. Spriggs, E. Hedblom, M. D. Cotton, T. R. Utterback, M. C. Hanna, D. T. Nguyen, D. M. Saudek, R. C. Brandon, L. D. Fine, J. L. Fritchman, J. L. Fuhrmann, N. S. M. Geoghagen, C. L. Gnehm, L. A. McDonald, K. V. Small, C. M. Frazer, H. O. Smith, and J. C. Venter.** 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science **269:**468–470.
8. **Friedman, S. A., and S. J. Austin.** 1988. The P1 plasmid-partition system synthesizes two essential proteins from an autoregulated operon. Plasmid **19:**103–112.
9. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166:**557–580.
10. **Jaffé, A., T. Ogura, and S. Hiraga.** 1985. Effects of the *ccd* function of the F plasmid on bacterial growth. J. Bacteriol. **163:**841–849.
11. **Jalakumari, M. B., and P. A. Manning.** 1989. Nucleotide sequence of the *traD* region in the *Escherichia coli* sex factor. Gene **81:**195–202.
12. **Karow, M., and C. Georgopoulos.** 1992. Isolation and characterization of the *Escherichia coli msbB* gene, a multicopy suppressor of null mutations in the high-temperature requirement gene *htrB*. J. Bacteriol. **174:**702–710.
13. **Katz, M. E., R. A. Strugnell, and J. I. Rood.** 1992. Molecular characterization of a genomic region associated with virulence in *Dichelobacter nodosus*. Infect. Immun. **60:**4586–4592.
14. **Lehnherr, H., E. Maguin, S. Jafri, and M. B. Yarmolinsky.** 1993. Plasmid addition genes of bacteriophage P1: *doc*, which causes cell death on curing of prophage, and *phd*, which prevents host death when the prophage is retained. J. Mol. Biol. **233:**414–428.
15. **Makino, S.-I., C. Sasakawa, and M. Yoshikawa.** 1988. Genetic relatedness of the basic replicon of the virulence plasmid in shigellae and enteroinvasive *Escherichia coli*. Microb. Pathog. **5:**267–274.
16. **Maneewannakul, K., and K. Ippen-Ihler.** 1993. Construction and analysis of F plasmid *traR*, *trbJ*, and *trbH* mutants. J. Bacteriol. **175:**1528–1531.
17. **Martin, K. A., M. A. Davis, and S. J. Austin.** 1991. Fine-structure analysis of the P1 plasmid partition site. J. Bacteriol. **173:**3630–3634.
18. **Martin, K. A., S. A. Friedman, and S. J. Austin.** 1987. Partition site of the P1 plasmid. Proc. Natl. Acad. Sci. USA **84:**8544–8547.
19. **Matsutani, S., H. Ohtsubo, Y. Maeda, and E. Ohtsubo.** 1987. Isolation and characterization of IS elements repeated in the bacterial chromosome. J. Mol. Biol. **196:**445–455.
20. **Miki, T., K. Yoshioka, and T. Horiuchi.** 1984. Control of cell division by sex factor F in *Escherichia coli*. I. The 42.84-43.6 F segment couples cell division of the host bacteria with replication of plasmid DNA. J. Mol. Biol. **174:**605–625.
21. **Miller, J. H.** 1992. A short course in bacterial genetics: a laboratory manual and handbook for *E. coli* and related bacteria. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
22. **Nordström, K., and S. J. Austin.** 1989. Mechanisms that contribute to the stable segregation of plasmids. Annu. Rev. Genet. **23:**37–69.
23. **Ohtsubo, H., T. B. Ryder, Y. Maeda, K. A. Armstrong, and E. Ohtsubo.** 1986. DNA replication of resistance plasmid R100 and its control. Adv. Biophys. **21:**115–133.
24. **Pullinger, G. D., and A. J. Lax.** 1992. A *Salmonella dublin* virulence plasmid locus that affects bacterial growth under nutrient-limited conditions. Mol. Microbiol. **6:**1631–1643.
25. **Radnedge, L., M. A. Davis, and S. J. Austin.** 1996. P1 and P7 plasmid partition: ParB protein bound to its partition site makes a separate discriminator contact with the DNA that determines species specificity. EMBO J. **15:**1155–1162.
26. **Rosen, J. I., T. B. Ryder, H. Inokuchi, H. Ohtsubo, and E. Ohtsubo.** 1980. Genes and sites involved in replication and incompatibility of an R100 plasmid derivative based on nucleotide sequence analysis. Mol. Gen. Genet. **179:**527–537.
27. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
28. **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.
29. **Sasakawa, C., K. Kamata, T. Sakai, S. Y. Murayama, S. Makino, and M. Yoshikawa.** 1986. Molecular alteration of the 140-megadalton plasmid associated with loss of virulence and Congo red binding activity in *Shigella flexneri*. Infect. Immun. **51:**470–475.
30. **Silva, R. M., S. Saadi, and W. K. Maas.** 1988. A basic replicon of virulence-associated plasmids of *Shigella* spp. and enteroinvasive *Escherichia coli* is homologous with a basic replicon in plasmids of IncF groups. Infect. Immun. **56:**836–842.
31. **Womble, D. D., and R. H. Rownd.** 1988. Genetic and physical map of plasmid NR1: comparison with other IncFII antibiotic resistance plasmids. Microbiol. Rev. **52:**433–451.
32. **Yarmolinsky, M. B.** 1995. Programmed cell death in bacterial populations. Science **267:**836–837.
33. **Yoshioka, Y., Y. Fujita, and E. Ohtsubo.** 1990. Nucleotide sequence of the promoter-distal region of the *tra* operon of plasmid R100, including *traI* (DNA helicase I) and *traD* genes. J. Mol. Biol. **214:**39–53.