INFECTION AND IMMUNITY, June 2000, p. 3780–3783 Vol. 68, No. 6
0019-9567/00/$04.00+0

Curli Loci of *Shigella* spp.

HARRY SAKELLARIS,[1] NERISSA K. HANNINK,[1] KUMAR RAJAKUMAR,[1,2]* DIETER BULACH,[1] MEREDITH HUNT,[1] CHIHIRO SASAKAWA,[3] AND BEN ADLER[1]

Department of Microbiology, Monash University, Clayton, Victoria, 3800,[1] and Department of Microbiology and Infectious Diseases, Royal Children's Hospital, Parkville, Victoria, 3052,[2] Australia, and Department of Bacteriology, Institute of Medical Science, University of Tokyo, Minato-ku, Tokyo 108-0071, Japan[3]

Received 15 October 1999/Returned for modification 15 November 1999/Accepted 3 March 2000

An unstable chromosomal element encoding multiple antibiotic resistance in *Shigella flexneri* serotype 2a was found to include sequences homologous to the *csg* genes encoding curli in *Escherichia coli* and *Salmonella enterica* serovar Typhimurium. As curli have been implicated in the virulence of serovar Typhimurium, we investigated the *csg* loci in all four species of *Shigella*. DNA sequencing and PCR analysis showed that the *csg* loci of a wide range of *Shigella* strains, of diverse serotypes and different geographical distributions, were almost universally disrupted by deletions or insertions, indicating the existence of a strong selective pressure against the expression of curli. Strains of enteroinvasive *E. coli* (EIEC), which share virulence traits with *Shigella* spp. and cause similar diseases in humans, also possessed insertions or deletions in the *csg* locus or were otherwise unable to produce curli. Since the production of curli is a widespread trait in environmental isolates of *E. coli*, our results suggest that genetic lesions that abolish curli production in the closely related genus *Shigella* and in EIEC are pathoadaptive mutations.

Bacillary dysentery is a severe diarrheal disease affecting hundreds of millions of people worldwide, leading to more than 500,000 deaths annually (11). The disease is caused by four bacterial species comprising the genus *Shigella*, i.e., *Shigella flexneri*, *S. dysenteriae*, *S. sonnei*, and *S. boydii. Shigella* spp. are transmitted to their hosts via the fecal-oral route and infect the colonic epithelium. Subsequent cell destruction, inflammation, and ulceration of the colon are responsible for the bloody, mucoid diarrhea that is characteristic of the disease. In recent years much has been learned about the sophisticated virulence mechanisms that allow *Shigella* to invade epithelial cells and spread to neighboring cells (5). However, nothing is known about the first step in the infection process, colonization of the host.

Bacterial colonization of the host intestine is generally mediated by fimbrial adhesins (2, 3, 8, 10). However, it is not clear what role fimbriae play in the virulence of *Shigella* (22). Over the last decade, *Escherichia coli* and *Salmonella* spp. have been found to express a surface structure termed thin aggregative fimbriae or curli (14, 20, 21). In *E. coli*, curli mediate the formation of biofilms on inert surfaces (26). However, in *Salmonella enterica* serovar Typhimurium, curli mediate bacterial attachment to mouse intestinal cells in vitro (25), and the expression of curli at 37°C is a phase-variable characteristic that is essential for full virulence in mice (24). These findings demonstrate that curli probably have a role in the colonization of the mouse intestine by serovar Typhimurium. Furthermore, curli are capable of mediating bacterial binding to a wide variety of tissues (15), cell matrix proteins, and plasma proteins (13, 14) and may therefore have additional roles in virulence.

During investigations of a deletable chromosomal element encoding multiple antibiotic resistance in *S. flexneri* serotype 2a (17, 18), members of our group discovered a locus with high sequence similarity to the *csg* gene clusters encoding curli in *E. coli* and serovar Typhimurium. This preliminary finding prompted us to investigate the presence of *csg* loci in a variety of *Shigella* strains.

Restriction analysis and DNA sample sequencing of *csg* loci in *Shigella*. To test whether the *csg* locus was present in all four species of *Shigella*, oligonucleotide primers were designed for the PCR amplification of an internal portion of the *csg* locus. Primer 4477 was homologous to a 5′-terminal sequence of the *csgE* gene of *E. coli* K-12, while primer 4480 was homologous to a 3′-terminal sequence of *csgA* (Fig. 1). PCR amplification of the *csg* internal fragment from the *E. coli* control strain, DH5α (9), generated a DNA fragment of 2.4 kb, the length predicted from sequence analysis of the *E. coli* K-12 *csg* locus (GenBank accession no. X90754). Similarly, a 2.4-kb fragment was amplified from *S. dysenteriae* serotype 3 strain SBA1304. However, fragments of 3.6, 4.0, and 2.15 kb were amplified from *S. flexneri* serotype 2a (SBA1100), *S. sonnei* (SBA1302), and *S. boydii* serotype 3 (SBA1308), respectively. These results implied that the *csg* locus was present in all four species of *Shigella* but had acquired insertions or undergone internal duplications in *S. flexneri* and *S. sonnei*, while the *S. boydii csg* locus appeared to have undergone a deletion.

To investigate the basis for the variation in length of the *csg* PCR products, each fragment was characterized by restriction mapping (Fig. 1) and sample sequencing. The 3.6-kb PCR product from *S. flexneri* contained sequences with homology to the *csgD*, *csgB*, and *csgA* genes of *E. coli* K-12 and serovar Typhimurium. However, an IS*600* element had inserted downstream of nucleotide 108 of the *csgA* gene (Fig. 1).

Sample sequencing of the 4-kb internal *csg* fragment from *S. sonnei* strain SBA1302 also revealed the presence of *csgD*, *csgB*, and *csgA* homologues. In addition, two distinct IS*1* elements had inserted into two separate locations in the *csg* locus. The first IS*1* element had interrupted the *csgD* gene at a position approximately 0.25 kb downstream of the start of the gene. The second IS*1* element had inserted into the intergenic region between the divergently transcribed *csgD* and *csgB* genes, 255 bp upstream of *csgD*.

The PCR-amplified *csg* fragment from *S. boydii* SBA1308

* Corresponding author. Mailing address: Department of Microbiology, Monash University, Victoria, 3800, Australia. Phone: 61 3 9344 2000. Fax: 61 3 9345 5764. E-mail: kumar.rajakumar@med.monash.edu.au.

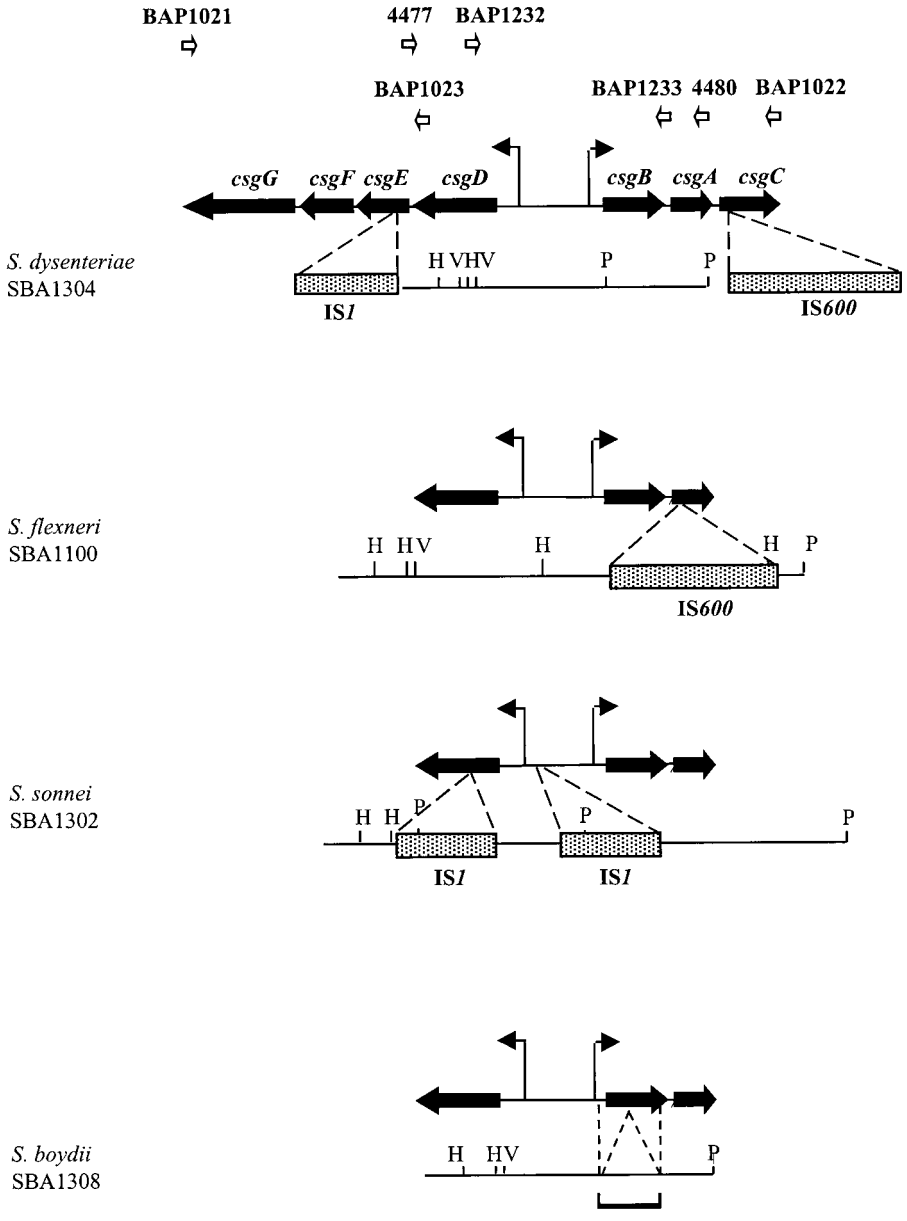

FIG. 1. Deletion and insertion mutations in the *csg* loci of *Shigella* spp. Restriction maps of the internal portion of the *csg* locus, spanning the 5′ terminus of *csgE* and the 3′ terminus of *csgA*, are shown for four strains of *Shigella*. The positions of ISs and deletions in relation to the *csg* genes are indicated by dashed lines. The 250-bp deletion in the *S. boydii* (SBA1308) *csg* gene lies within a 350-bp region indicated by a horizontal bracket. The positions of promoters for the *csgBAC* and *csgDEFG* operons are indicated by bent arrows. The binding sites for primers are indicated by open arrows. Restriction sites include *Hin*dIII (H), *Pst*I (P), and *Eco*RV (V). Primer sequences are as follows: 4477, 5′-GCGGCGAACAGAAATTCTGCC-3′; 4480, 5′-GTTACCAAAGCCAACCTGAGTCACG-3′; BAP1021, 5′-TCAGG ATTCCGGTGGAAC-3′; BAP1022, 5′-TTAAGACTTTTCTGAAGAGG-3′; BAP1023, 5′-CAGCGAAAATACGGTTAAAACGC-3′; BAP1232, 5′-TCCTGCTCAA AGTATCCTGCC-3′); and BAP1233, GATTGCTGCAATTGCTGCTAC-3′.

was shorter than the 2.4-kb fragment amplified from DH5α, suggesting that a deletion had occurred in the *S. boydii csg* locus. In order to locate the site of the predicted 250-bp deletion, the 2.15-kb PCR product was sequenced with primers 4477 and 4480 and the two internal primers BAP1232 and BAP1233 (Fig. 1). Sequencing showed that the deletion had occurred within a 350-bp region which included most of the *csgB* open reading frame (Fig. 1).

Although no deletions or insertions were evident from PCR analysis of the *S. dysenteriae* serotype 3 *csg* locus, the primers used did not encompass the entire *csg* locus. This left open the possibility that mutations existed in other parts of the locus. To address this question, primers BAP1021 and BAP1022 were used to PCR amplify the entire *csg* locus of *S. dysenteriae* (Fig. 1). PCR products of 4.4 kb, the expected length of the intact *E. coli csg* locus, were amplified from DH5α and the curliated *E. coli* strains YMel (19) and χ7122 (16). However, a product of approximately 6 kb was amplified from *S. dysenteriae* serotype 3, suggesting that one or more IS elements may have inserted into the locus, outside of the region previously investigated. To determine the sites of the proposed insertion, the 6-kb PCR product was sequenced with primers BAP1021, BAP1022, and BAP1023 (Fig. 1). Sequence analysis showed that an IS*1* element had inserted downstream of nucleotide 98 in the *csgE*

TABLE 1. Presence and size of the *csg* locus in *Shigella* strains

Strain	Serotype	Country of isolation	Size of *csg* PCR product (kb)
S. flexneri			
SBA1316	2b	Australia	NP[a]
SBA1317	2a	Australia	6.4
SBA1318	2a	Australia	6.4
SBA1319	2a	Australia	6.4
SBA1320	2a	Australia	5
SBA1321	2a	Australia	5
SBA1322	2a	Australia	6.4
SBA1323	2a	Australia	6.4
SBA1173	1b	Australia	NP
SBA1387	5a	Japan	5
SBA1388	2a	Japan	5
SBA1389	3a	Japan	NP
SBA1390	2b	Japan	5
SBA1391	4a	Japan	5
SBA1392	6	Japan	NP
SBA1401	2b	Australia	6.4
SBA1402	2a	Australia	6.4
SBA1403	4	Australia	NP
SBA1404	2a	Australia	5
SBA1405	6	Australia	NP
SBA1406	6	Australia	NP
SBA1407	3a	Australia	7.4
SBA1308	4a	Australia	6
S. dysenteriae			
SBA1393	1	Japan	NP
SBA1394	3	Japan	6
SBA1395	9	Japan	7
SBA1396	6	Japan	6
SBA1397	4	Japan	NP
SBA1398	5	Japan	2.2
S. boydii			
SBA1381	1	Japan	6
SBA1382	2	Japan	5.5
SBA1383	3	Japan	3
SBA1384	4	Japan	5.5
SBA1385	7	Japan	4.4
SBA1386	8	Japan	3
S. sonnei			
SBA1375		Japan	6.4
SBA1376		Japan	6
SBA1377		Japan	6.4
SBA1378		Japan	6
SBA1379		Japan	6
SBA1380		Japan	6
SBA1399		Australia	7
SBA1400		Australia	6.4

[a] NP, no product.

gene, while an IS*600* element had inserted downstream of nucleotide 71 of the *csgC* gene.

Survey of insertion and deletion mutations in the *csg* locus of *Shigella* strains. Initial analysis of the four *Shigella* strains representing each species suggested that mutations within the *csg* locus are probably widespread phenomena. To test this hypothesis, 43 *Shigella* strains, representing a wide range of serotypes isolated over several years from patients in Australia and Japan, were surveyed by colony PCR with the primers BAP1021 and BAP1022, which flank the complete *csg* locus. The results (Table 1) demonstrated that insertions into the *csg* locus are widespread in *Shigella* spp. The size variation in the *csg* loci suggests that different types of insertion events have occurred. The smaller PCR products are consistent with the insertion of single IS elements, while the larger products (6.4 to 7.4 kb) are consistent with the insertion of multiple IS elements similar to those in *S. dysenteriae* serotype 3 strain SBA1304 and *S. sonnei* strain SBA1302 (Fig. 1). However, in many strains the *csg* locus was either partially deleted or not detected at all.

Only a serotype 7 strain of *S. boydii*, SBA1385, produced a 4.4-kb fragment, suggesting that the *csg* locus may be intact. To test if this strain produced curli, SBA1385 and the positive control *E. coli* strains, YMel and χ7122, were grown on CFA agar (4) for 48 h at 25°C and were negatively stained with ammonium phosphotungstate for examination by electron microscopy as previously described (7). Although curli were clearly visible on the positive control *E. coli* strains, they were not produced by SBA1385 (data not shown). The most likely explanations for the absence of curli in SBA1385 include the possibility of point mutations or small deletions that were undetectable by agarose gel electrophoresis. Alternatively, the absence of curli may have been due to extragenic mutations such as those in *rpoS*, which are known to affect curli expression in *E. coli* (13).

Significance of mutations in the *csg* loci of *Shigella* strains. The insertion and precise excision of IS elements into genes have been described as a possible mechanism for the control of gene expression. However, this is only likely to be significant when insertion and excision occur at high frequencies. For example, the expression of exopolysaccharide synthesis in *Pseudomonas atlantica* is mediated by an IS element that excises from the *eps* locus at frequencies as high as 0.5 (1). This generates genetically distinct subpopulations that are preadapted to environmental change (1). In contrast, precise excision of IS*1* in *E. coli* occurs at frequencies of less than 10^{-5} (12). Furthermore, since it appears that curli loci in *Shigella* are often interrupted by multiple IS elements, the restoration of curli expression by the simultaneous excision of more than one IS element seems unlikely. Rather than being involved in the control of curli expression, we propose that insertions into the *csg* locus are common because of a strong selection against the expression of curli in *Shigella*. This is supported by the finding that up to a quarter of *Shigella* strains may have partial or complete deletions of the *csg* locus.

Since curli are expressed in a wide variety of *E. coli* strains (13), we propose that the widespread loss of curli in the closely related genus *Shigella* may represent a pathoadaptive mutation. Such loss-of-function mutations overcome selections against an organism that is in transition from a free-living or commensal niche to a virulence niche (23). In this way, genes that hinder the colonization of or survival within the new virulence niche are lost in the process of evolution towards a pathogenic mode of life. It is possible that during the divergence of *Shigella* and *E. coli*, the expression of curli in the new virulence niche became a selective disadvantage in *Shigella*. Supporting this hypothesis is the observation that although the ability to express curli at 37°C increases the infectivity of serovar Typhimurium, the pathogen is cleared much more rapidly than are natural variants unable to express curli at body temperature (24), suggesting that curli are good targets for the host immune-clearance systems. In serovar Typhimurium the controlled expression of curli provides the pathogen with a net selective advantage, i.e., colonization of the host intestine (25), which outweighs the disadvantage of increased exposure to the host immune system. The ability to express curli is therefore maintained. We propose, however, that as *Shigella* and *Salmonella* have followed different evolutionary paths, the disadvantage of increased exposure to the host immune system has outweighed selective pressures to maintain curli in *Shigella*.

If host factors have selected against the expression of curli in *Shigella*, a similar phenomenon would be expected in enteroinvasive strains of *E. coli* (EIEC), which share common virulence determinants with *Shigella* and produce similar diseases in humans (6). To test this hypothesis, 11 EIEC strains isolated from patients in Australia, South Africa, and Japan were examined by PCR with primers BAP1021 and BAP1022 for insertions or deletions in the *csg* locus. Six strains had insertions ranging in size from 0.5 to 2.5 kb. The five remaining strains, however, appeared to have intact *csg* loci, as judged by agarose electrophoresis of the PCR products. These strains were tested for their ability to produce curli by electron microscopy of negatively stained cells. When grown on CFA agar at 25°C, the control *E. coli* strains, YMel and χ7122, clearly produced curli. In contrast, curli were not produced by any of the EIEC strains (data not shown). These results are consistent with the previous finding that EIEC strains do not bind fibronectin (13), a characteristic associated with curli in *E. coli*.

Our work demonstrates that there is a strong selective pressure against the maintenance of curli in EIEC, as in *Shigella* spp. The observation that curli are expressed in 60% of environmental isolates of *E. coli* (13) but are absent from all strains of the closely related genus *Shigella* and EIEC, two bacterial groups with very similar mechanisms of pathogenesis, supports the hypothesis that mutations abolishing curli expression in these strains are pathoadaptive.

Nucleotide sequence accession numbers. Nucleotide sequences of the sites of IS element insertion into the *csg* locus have been deposited into the GenBank database under the accession numbers AF237724, AF237725, AF237726, and AF237727.

We are very grateful to D. Lightfoot for supplying *Shigella* strains, Roy Robbins-Browne for supplying EIEC strains, and Arne Olsen and Roy Curtiss for supplying curliated *E. coli* strains.

This work was supported by a grant from the National Health and Medical Research Council, Canberra, Australia.

REFERENCES

1. **Bartlett, D. H., M. Wright, and M. Silverman.** 1988. Variable expression of extracellular polysaccharide in the marine bacterium *Pseudomonas atlantica* is controlled by genome rearrangement. Proc. Natl. Acad. Sci. USA **85:**3923–3927.
2. **Bieber, D., S. W. Ramer, C. Y. Wu, W. J. Murray, T. Tobe, R. Fernandez, and G. K. Schoolnik.** 1998. Type IV pili, transient bacterial aggregates, and virulence of enteropathogenic *Escherichia coli*. Science **280:**2114–2118.
3. **Darwin, K. H., and V. L. Miller.** 1999. Molecular basis of the interaction of *Salmonella* with the intestinal mucosa. Clin. Microbiol. Rev. **12:**405–428.
4. **Evans, D. G., D. J. Evans, Jr., and W. Tjoa.** 1977. Hemagglutination of human group A erythrocytes by enterotoxigenic *Escherichia coli* isolated from adults with diarrhea: correlation with colonization factor. Infect. Immun. **18:**330–337.
5. **Finlay, B. B., and S. Falkow.** 1997. Common themes in microbial pathogenicity revisited. Microbiol. Mol. Biol. Rev. **61:**136–169.
6. **Floderus, E., T. Pal, K. Karlsson, and A. A. Lindberg.** 1995. Identification of *Shigella* and enteroinvasive *Escherichia coli* strains by a virulence-specific, monoclonal antibody-based enzyme immunoassay. Eur. J. Clin. Microbiol. Infect. Dis. **14:**111–117.
7. **Froehlich, B. J., A. Karakashian, H. Sakellaris, and J. R. Scott.** 1995. Genes for CS2 pili of enterotoxigenic *Escherichia coli* and their interchangeability with those for CS1 pili. Infect. Immun. **63:**4849–4856.
8. **Gaastra, W., and A. M. Svennerholm.** 1996. Colonization factors of human enterotoxigenic *Escherichia coli* (ETEC). Trends Microbiol. **4:**444–452.
9. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166:**557–580.
10. **Herrington, D. A., R. H. Hall, G. Losonsky, J. J. Mekalanos, R. K. Taylor, and M. M. Levine.** 1988. Toxin, toxin-coregulated pili, and the *toxR* regulon are essential for *Vibrio cholerae* pathogenesis in humans. J. Exp. Med. **168:**1487–1492.
11. **Katz, S. L.** 1986. Diseases of importance in developing countries, vol. 2, p. 165. National Academy Press, Washington, D.C.
12. **Lu, S. D., D. Lu, and M. Gottesman.** 1989. Stimulation of IS*1* excision by bacteriophage P1 *ref* function. J. Bacteriol. **171:**3427–3432.
13. **Olsen, A., A. Arnqvist, M. Hammar, S. Sukupolvi, and S. Normark.** 1993. The RpoS sigma factor relieves H-NS-mediated transcriptional repression of *csgA*, the subunit gene of fibronectin-binding curli in *Escherichia coli*. Mol. Microbiol. **7:**523–536.
14. **Olsen, A., A. Jonsson, and S. Normark.** 1989. Fibronectin binding mediated by a novel class of surface organelles on *Escherichia coli*. Nature **338:**652–655.
15. **Olsén, A., M. J. Wick, M. Mörgelin, and L. Björck.** 1998. Curli, fibrous surface proteins of *Escherichia coli*, interact with major histocompatibility complex class I molecules. Infect. Immun. **66:**944–949.
16. **Provence, D. L., and R. Curtiss III.** 1992. Role of *crl* in avian pathogenic *Escherichia coli*: a knockout mutation of *crl* does not affect hemagglutination activity, fibronectin binding, or curli production. Infect. Immun. **60:**4460–4467.
17. **Rajakumar, K., D. Bulach, J. Davies, L. Ambrose, C. Sasakawa, and B. Adler.** 1997. Identification of a chromosomal *Shigella flexneri* multi-antibiotic resistance locus which shares sequence and organizational similarity with the resistance region of the plasmid NR1. Plasmid **37:**159–168.
18. **Rajakumar, K., C. Sasakawa, and B. Adler.** 1996. A spontaneous 99-kb chromosomal deletion results in multi-antibiotic susceptibility and an attenuation of contact hemolysis in *Shigella flexneri* 2a. J. Med. Microbiol. **45:**64–75.
19. **Richenberg, H. V., and G. Lester.** 1955. The preferential synthesis of beta-galactosidase in *Escherichia coli*. J. Gen. Microbiol. **13:**279–284.
20. **Römling, U., Z. Bian, M. Hammar, W. D. Sierralta, and S. Normark.** 1998. Curli fibers are highly conserved between *Salmonella typhimurium* and *Escherichia coli* with respect to operon structure and regulation. J. Bacteriol. **180:**722–731.
21. **Sjobring, U., G. Pohl, and A. Olsen.** 1994. Plasminogen, absorbed by *Escherichia coli* expressing curli or by *Salmonella enteritidis* expressing thin aggregative fimbriae, can be activated by simultaneously captured tissue-type plasminogen activator (t-PA). Mol. Microbiol. **14:**443–452.
22. **Snellings, N. J., B. D. Tall, and M. M. Venkatesan.** 1997. Characterization of *Shigella* type 1 fimbriae: expression, FimA sequence, and phase variation. Infect. Immun. **65:**2462–2467.
23. **Sokurenko, E. V., D. L. Hasty, and D. E. Dykhuizen.** 1999. Pathoadaptive mutations: gene loss and variation in bacterial pathogens. Trends Microbiol. **7:**191–195.
24. **Sukupolvi, S., A. Edelstein, M. Rhen, S. J. Normark, and J. D. Pfeifer.** 1997. Development of a murine model of chronic *Salmonella* infection. Infect. Immun. **65:**838–842.
25. **Sukupolvi, S., R. G. Lorenz, J. I. Gordon, Z. Bian, J. D. Pfeifer, S. J. Normark, and M. Rhen.** 1997. Expression of thin aggregative fimbriae promotes interaction of *Salmonella typhimurium* SR-11 with mouse small intestinal epithelial cells. Infect. Immun. **65:**5320–5325.
26. **Vidal, O., R. Longin, C. Prigent-Combaret, C. Dorel, M. Hooreman, and P. Lejeune.** 1998. Isolation of an *Escherichia coli* K-12 mutant strain able to form biofilms on inert surfaces: involvement of a new *ompR* allele that increases curli expression. J. Bacteriol. **180:**2442–2449.

Editor: A. D. O'Brien