CadC Is the Preferential Target of a Convergent Evolution Driving Enteroinvasive *Escherichia coli* toward a Lysine Decarboxylase-Defective Phenotype

Mariassunta Casalino,1 Maria Carmela Latella,2 Gianni Prosseda,2 and Bianca Colonna2*

*Dipartimento di Biologia, Università Roma Tre, 00146 Rome,† and Dipartimento di Biologia Cellulare e dello Sviluppo, Università “La Sapienza,” 00185 Rome, Italy*

Received 6 February 2003/Returned for modification 8 April 2003/Accepted 10 July 2003

Enteroinvasive *E. coli* (EIEC), like *Shigella*, is the etiological agent of bacillary dysentery, a particularly severe syndrome in children in developing countries. All EIEC strains share with *Shigella* the inability to synthesize lysine decarboxylase (the LDC phenotype). The lack of this function is considered a pathoadaptive mutation whose emergence was necessary to obtain the full expression of invasiveness. Cadaverine, the product of lysine decarboxylation, is a small polyamine which interferes mainly with the inflammatory process induced by dysenteric bacteria. Genes coding for lysine decarboxylase and its transporters constitute a single operon (*cadBA*) and are expressed at low pH under the positive control of CadC. This regulator is an inner membrane protein that is able to sense pH variation and to respond by transcriptionally activating the *cadBA* genes. In this study we show that, unlike in *Shigella*, mutations affecting the *cad* locus in the EIEC strains we have analyzed are not followed by a novel gene arrangement and that the LCD− phenotype is dependent mainly on inactivation of the *cadC* gene. Introduction of a functional CadC restores cadaverine expression in all EIEC strains harboring either an IS2 element or a defective *cadC* promoter. Comparative analysis between the *cad* regions of *S. flexneri* and EIEC suggests that the LCD− phenotype has been attained by different strategies within the *E. coli* species.

In bacteria the evolution toward pathogenic phenotypes revolves around two distinct mechanisms, the acquisition of additional genes encoding virulence determinants and the loss or modification of preexisting genetic material. The acquisition of multiple linked virulence traits by horizontal gene transfer is very important in triggering a virulence phenotype in a commensal organism, which can thus gain access to new host environments (11, 25). The new pathogen then reaches optimal fitness within the novel environment by adaptation of its genome through mutations in preexisting genes. These so-called pathoadaptive mutations improve survival within host tissues, increase the pathogenic potential of the bacteria, and drive the evolution of a microorganism toward a more pathogenic phenotype (26, 40).

In this context, evolution of *Escherichia coli* toward pathogenicity represents an interesting model for the enormous versatility of this microorganism, causing an impressive variety of different diseases (23). Acquisition of virulence genes by horizontal transfer has played an important role in the evolution of pathogenic *E. coli* strains. Virulence determinants have been acquired by *E. coli* cells as parts of plasmids, bacteriophages, transposons, or pathogenicity islands (13). Enteroinvasive *E. coli* (EIEC) strains are facultative intracellular pathogens able to enter epithelial cells of the colon, replicate within them, and move between adjacent cells with a mechanism similar to that of *Shigella*, the causative agent of bacillary dysentery (35, 36). In EIEC and *Shigella*, the critical event in the transition toward a pathogenic life-style probably relies on the acquisition of a large (220- to 260-kb) F-type plasmid (pINV) (35). This plasmid contains all the genes required for invasion and for intra- and intercellular spread, including their positive activators. Plasmid virulence genes are silenced outside the host by one of the major histone-like proteins, H-NS, which acts as a repressor on two positive activators (27). pINVs isolated from EIEC and *Shigella* strains share wide regions of high structural and functional homology (15).

Although EIEC strains may be developing the full *Shigella* phenotype, they do not have the full set of characters that define *Shigella* strains and are not included in any of the three *Shigella* clusters (16, 28, 29). EIEC strains generally correspond to bioserotypes found in a dozen *E. coli* serogroups, and recent molecular analyses confirm that they are widely distributed among *E. coli* phylogenetic groups (28, 30). Many EIEC strains have *Shigella*-like features: they can be Lac−; nonmotile, low-level indole producing, and/or unable to produce gas during fermentation. Some have the same O antigen that is present in *Shigella*. In contrast to *Shigella*, they have a high metabolic activity since they still retain the ability to catabolize substrates widely utilized by *E. coli* (29, 39). Despite these differences, EIEC and *Shigella* share the inability to catabolize lysine due to the lack of lysine decarboxylase (LDC) activity. Whereas almost 90% of *E. coli* strains are LDC+, all EIEC and *Shigella* strains are LDC− (18, 39). In *E. coli* the ability to use lysine as carbon source is determined by the products of the *cadBA* operon, encoding lysine decarboxylase (*cadA*) and a lysine-cadaverine antiporter (*cadB*), and is coinduced by low pH, anaerobiosis, and lysine. Expression of the *cadBA* operon is dependent on the CadC positive activator, whose gene maps upstream of the operon and is transcribed independently from the same strand (20, 21, 42). Under noninducing conditions,
H-NS represses the cadBA operon (38). While the lack of a functional hns gene also induces high and constitutive expression of the cadBA operon at basic pH, null mutations in cadC completely silence the operon (38, 42). CadC is an inner membrane protein with a carboxy-terminal periplasmic domain, responsible for sensing pH, and an amino-terminal cytoplasmatic domain, providing the DNA binding activity necessary for transcriptional activation (10).

Recent evidence shows that cadaverine, a polyamine resulting from decarboxylation of lysine, negatively interferes with the Shigella pathogenicity process. Cadaverine seems to induce the attenuation of enterotoxicity (18) and also to inhibit the migration of polymorphonuclear leukocytes (PMN) across the intestinal epithelial monolayers (12, 19). The Shigella-induced inflammatory response of the host is a “Trojan’s horse” step in the invasive process, since PMN trafficking across intestinal epithelia opens the tight junctions between epithelial cells and thus facilitates the invasion of the colonic mucosa at basolateral sites (36). It has been proposed that silencing of the cad locus in Shigella represents an important pathoadaptive mutation necessary for increasing the pathogenic potential of bacteria in host tissues (9, 18, 40).

Here we report an analysis of the molecular rearrangements in the cad region of several EIEC strains belonging to different serotypes and isolated in different geographic areas. Our results show that the main strategy adopted for silencing the cad operon relies on mutations in cadC and that the arrangement of genes neighboring the cad locus is colinear to that of E. coli K-12. In most EIEC strains, the introduction of a functional cadC gene restores wild-type (wt) levels of lysine decarboxylase activity. Comparative analysis with respect to the cad region of Shigella flexneri shows that convergent evolution has operated independently to prevent the synthesis of cadaverine and suggests that the rearrangement observed in EIEC could be considered an evolutionary intermediate of a recombination process leading to the loss of the entire region.

MATERIALS AND METHODS

Bacterial strains and general procedures. The bacterial strains used are listed in Table 1. Growth media were Trypticase soy broth and Luria broth medium (34). The solid media contained 1.5% agar. Congo red was used at 0.01% in Trypticase soy agar to check for the Congo red binding (CRb) phenotype. When required, antibiotics were included at the following concentrations: ampicillin, 100 μg/ml; chloramphenicol, 30 μg/ml; kanamycin, 30 μg/ml; spectinomycin, 20 μg/ml; streptomycin, 100 μg/ml; sulfanamide, 600 μg/ml in M9 minimal medium; and tetracycline, 5 μg/ml. To quantify the expression of the cadA gene, strains were grown in modified Falkow lysine decarboxylase medium buffered at pH 5.5 with 100 mM morpholineethanesulfonic acid (MES). Bacterial invasion of HeLa cells was performed as previously described (8, 24) using strains grown at 37°C in Trypticase soy broth to an optical density at 600 nm (OD600) of 0.6. E. coli K-12 MC4100 [F- ΔaraD139 Δ(argF-lac)150cm rpsL150 relA flbB5301 deoC1 pstF25 rbsR] (22) was used as the recipient strain for plasmids carrying gene fusions. β-Galactosidase assay results are provided in Table 2 and were scored based on the zero (negative) line for the E. coli K-12 MC4100 strain.

DNA manipulations. Total and plasmid DNA extraction, restriction digestion, electrophoresis, and purification of DNA fragments were carried out as described previously (3, 8). Southern hybridization was performed using a 32P-labeled DNA probe essentially by the method of Sambrook and Russell (34). Cad-specific probes were obtained by PCR amplification using M1655 total DNA as template and the primer pairs A1 plus A2 (cadA) or B1 plus B3 (cadB) probe. The IS1 and IS2 probes were obtained from λ::IS1 and λ::IS2, respectively (37). PCR was routinely performed using Taq polymerase, or ExTaq (Takara) for fragments larger than 4 kb. The primers used throughout this study are described in Table 2 and were designed mainly based on the sequence of the E. coli K-12 M1655 genome (4).

Plasmid construction. Plasmids pC280, pCC81, pCC53, pCC90, and pCA681 were generated by cloning Pta ampicillin into pGem-T (Promega) as specified by the manufacturer. For pC280, pCC81, pCC53, and pCC90, the ampicillin was obtained with primer pair B2 plus C2 from genomic DNA of strains HN280, 6.81, 53638, and M90T, respectively. For pCA681, the amplicon was obtained with the oligonucleotide pair A1 plus B2R using genomic DNA of EIEC 681. Plasmid DNA from two independent clones was then used as template for PCR amplification. Amplification products were treated with ExTaq polymerase and labeled DNA probes essentially by the method of Sambrook and Russell (34). The IS1 and IS2 probes were obtained from λ::IS1 and λ::IS2, respectively (37). PCR was routinely performed using Taq polymerase, or ExTaq (Takara) for fragments larger than 4 kb. The primers used throughout this study are described in Table 2 and were designed based on the sequence of the E. coli K-12 M1655 genome (4).
TABLE 2. Nucleotide sequences of primers used for PCR analysis of the cadBA operon

<table>
<thead>
<tr>
<th>Primer</th>
<th>Nucleotide sequence</th>
<th>Target sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>5'-CTT CCC T7G TAC GAG CTA ATT-3'</td>
<td>cadA</td>
</tr>
<tr>
<td>A2</td>
<td>5'-CTGCA GAT AGC ATG AAC-3'</td>
<td>cadA</td>
</tr>
<tr>
<td>B1</td>
<td>5'-CCT ACA CAT GGG ACA AAA TGA-3'</td>
<td>cadB</td>
</tr>
<tr>
<td>B2</td>
<td>5'-GCC ATG TTA CCA ATC AGC-3'</td>
<td>cadB</td>
</tr>
<tr>
<td>B2R</td>
<td>5'-CTG ATG T7G TAA CTT GGC-3'</td>
<td>cadB</td>
</tr>
<tr>
<td>C1</td>
<td>5'-GTC ATG CAC CAT AAA GGT GG-3'</td>
<td>cadC</td>
</tr>
<tr>
<td>C2</td>
<td>5'-GGC T7G T7C ATG CAA AGA C-3'</td>
<td>Intergenic</td>
</tr>
<tr>
<td>I2</td>
<td>5'-GGC GGC G7G AAA GAA AGC-3'</td>
<td>IS2 element</td>
</tr>
<tr>
<td>Y1</td>
<td>5'-GGA AGG TCA T7G ATC AGG-3'</td>
<td>ysl1</td>
</tr>
<tr>
<td>Y2</td>
<td>5'-AAT ATC CGG CAA GGC GAG-3'</td>
<td>yslA</td>
</tr>
<tr>
<td>I1</td>
<td>5'-CAT ACC GGT AAA GAA GCC-3'</td>
<td>IS1 element</td>
</tr>
<tr>
<td>LC1</td>
<td>5'-CGT AGA GAA TTC ATA AAG CAC AAG-3'</td>
<td>Upstream of cadC</td>
</tr>
<tr>
<td>LC2</td>
<td>5'-CAA CGG GAT CAT CAC G7T G7T GCA T-3'</td>
<td>cadC</td>
</tr>
</tbody>
</table>

Table 3. Analysis of regions containing the cad operon in EIEC

<table>
<thead>
<tr>
<th>Strain</th>
<th>Size (kb) of PCR product (primers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>yidA</td>
<td>cadB (Y1-B2R)</td>
</tr>
<tr>
<td>cadA</td>
<td>(A1-A2)</td>
</tr>
<tr>
<td>cadB</td>
<td>(B1-B3)</td>
</tr>
<tr>
<td>IF1</td>
<td>cadC (B2-C2)</td>
</tr>
<tr>
<td>IS2</td>
<td>(C1-12)</td>
</tr>
<tr>
<td>MG1655</td>
<td>7.9 2.2 1.7 2.5 ND</td>
</tr>
<tr>
<td>HN280</td>
<td>7.9 2.2 1.7 3.8 0.8</td>
</tr>
<tr>
<td>HN11</td>
<td>7.9 2.2 1.7 3.8 0.8</td>
</tr>
<tr>
<td>HN13</td>
<td>7.9 2.2 1.7 3.8 0.8</td>
</tr>
<tr>
<td>53638</td>
<td>7.9 2.2 1.7 2.5 ND</td>
</tr>
<tr>
<td>1300</td>
<td>np  np  np np ND</td>
</tr>
<tr>
<td>6-81</td>
<td>7.9 2.2 1.7 1.9 ND</td>
</tr>
</tbody>
</table>

a E. coli K-12, used as control.
b np, no PCR product.
c ND, not done.

RESULTS

Genetic variability of the cadBA operon in EIEC. EIEC is not a homogenous group. EIEC strains differ with respect to serotype, plasmid content, and biochemical features. Moreover, recent studies of the genetic relationships between pathogenic and commensal E. coli strains by multilocus enzyme electrophoresis or ribosomal DNA restriction fragment length polymorphism confirm the presence of EIEC among different clusters of E. coli species (28, 30). Despite this diversity, these strains are all defective in the synthesis of lysine decarboxylase (18, 39). Since no data are available to explain the origin of this phenotype, we studied the genetic organization of the cad region in several EIEC strains and its relationship to that of S. flexneri.

Some EIEC and S. flexneri strains used in this study were isolated previously during a 2-year epidemiological survey of diarrheal diseases in children from Somalia (5, 6, 24), while others were obtained from the Reference Centers of the Pasteur Institute and of the Walter Reed Army Institute of Research (35). The EIEC strains were isolated in different geographic areas, belong to different serotypes, and display different plasmid contents, but they are all positive in invasivity assays. Moreover, no EIEC strains produce lysine decarboxylase (LDC− phenotype) or cadaverine, even under inducing conditions (Table 1).

To detect molecular rearrangements in the cad operon leading to the LDC− phenotype, we first analyzed EIEC strains for the presence of the cadA gene by PCR amplification using primers (A1 and A2) (Table 2) flanking the complete gene. As reported in Table 3, EIEC strains generate a DNA fragment with the same size (2.2 kb) as that of the E. coli K-12 control strain MG1655 (4). Only EIEC 13.80 lacks the cadA product. The presence of a cad region is also confirmed by the appearance of a cadA-specific 2.4-kb PstI fragment in Southern blots of genomic digests from all EIEC strains except 13.80 (data not shown). Taken together, these results suggest that despite their inability to produce cadaverine, most of the EIEC strains asayed contain the cadA gene.

We then analyzed the same EIEC strains for the presence of the cadB gene. First, we amplified the cadB region using primers (B1 and B3) flanking the complete cadB gene. We obtained a PCR product of the same size (1.7 kb [Table 3]) as the one from the control strain. Also in this case, no product was observed using EIEC 13.80 genomic DNA as template. Unlike the Southern analysis with the cadA probe, hybridization of PstI-digested genomic DNA with a cadB probe revealed a certain variability among our EIEC strains (Fig. 1). Only in EIEC 53638 did the restriction fragments containing the cadB gene have the same size as in E. coli K-12: a 2.4-kb PstI fragment (containing almost the entire cadA gene and part of the cadB gene) and a 7.3-kb PstI fragment (containing most of the cadB gene). Surprisingly, in four strains (HN280, HN11, HN13, and 4608-58), the 7.3-kb PstI fragment was replaced by an 8.6-kb fragment, suggesting the acquisition of insertion sequences or the presence of internal duplications. In agreement with the results obtained by PCR analysis with cadB primers, it is logical to expect that the rearrangement is located outside the cadB region and therefore within the cadC gene or its upstream region.

Detection of IS elements in the cadC gene. To find which kinds of molecular alterations have caused the loss of lysine

Downloaded from http://iai.asm.org on October 15, 2017 by guest
specific probe (data not shown). The location and orientation of the IS2 element were then confirmed by PCR analysis with a primer pair (C1 and I2) internal to the cadC gene and to the IS2 sequence, respectively (Table 3; Fig. 2).

Amplification of EIEC 6.81 genomic DNA with primers B2 and C2 gave a shorter amplicon (1.9 kb) than that of the control strain, MG1655 (2.5 kb [Table 3]), suggesting that this strain has suffered a deletion upstream of the cadB locus. The 1.9-kb amplicon from EIEC 6.81 was then cloned into pGem-T, yielding pCC681. Sequence analysis reveals that in this case an IS2 element is inserted into the cadC gene (Fig. 2). Thus, unlike the IS2 insertion in the cadC gene of the EIEC strains described above, the IS1 insertion has promoted a severe deletion (1,385 bp) at the 5′ end of the cadC gene.

These results suggest that in 5 of 7 EIEC strains examined, the lack of lysine decarboxylase activity depends on IS1 or IS2 insertional inactivation of the cadC gene. EIEC strain 13.80 seems to harbour a wide deletion encompassing the entire cad region, since we never observed homology to cad probes (Fig. 1; Table 3). In contrast, in EIEC 53638 (Fig. 1 and 2; Table 3), the organization of the cad region is similar to that of E. coli K-12. The most likely explanation for the absence of lysine decarboxylase activity in this strain includes the possibility of point mutations or small deletions within cad genes. Alternatively, the absence of lysine decarboxylase activity may be due to mutations in other (albeit yet unknown) regulators mapping outside the cad region.

To verify whether, in EIEC strains conserving the cadBA operon, the LCD− phenotype is determined only by the absence of CadC, we transformed these strains with a plasmid (pCC55) harboring a functional cadC gene. Introduction of pCC55 into the four EIEC cadC::IS2 strains (HN280, HN11, HN13, and 4608-58 [Fig. 2]) restores the expression of the cadBA operon, as confirmed by the excretion of a large amount of cadaverine into the medium (Fig. 3). Also in the EIEC strain with an apparently wt cad region (EIEC 53638), we observed complementation by pCC55 (Fig. 3). Interestingly, sequence analysis of the cadC locus of strain 53638 revealed the presence in the promoter region of a transversion (A to C) positioned in the −10 box (5′-AATTCT-3′). To verify the promoter activity of cadC, we cloned the cad promoter region (−341 to +40) from MG1655 and EIEC 53638 in the promotor probe vector pRS414, obtaining pLacC55 and pLacC53, respectively. Compared to the wt strain, β-galactosidase expression of the cadC-lacZ hybrid gene was almost silenced under the control of the cadC promoter of 53638 (5 MU with pFlacC53 versus 400 MU with pLacC55 in the MC4100 background), suggesting that in this strain the LCD− phenotype was obtained through the inactivation of the cadC promoter.

Since the presence of a functional cadC regulator is not able to restore cadaverine secretion in EIEC 6.81 (which carries a severe cadC rearrangement due to an IS1 insertion and an IS1-mediated deletion), we suppose that in this strain the LCD− phenotype also stems from a mutation in the cadBA operon. In fact, sequence analysis of this operon confirms that in this strain the cadA gene is also inactivated by several missense mutations in the coding sequence.

Analysis of the region flanking the cadA gene. The results obtained by PCR, Southern blotting, and sequence analysis indicate that in all EIEC strains (with the single exception of
strain 13.80), the cad region is indeed present and shows colinearity with the E. coli K-12 genome extending beyond the pheU gene, as indicated by the PstI digestion patterns (Fig. 1). Also in EIEC 1380, despite the complete loss of the cad genes, PCR-based analysis with oligonucleotides internal to the neighboring genes indicates that colinearity with the E. coli K-12 chromosome is reestablished upstream of yjdF and downstream of dsbD (data not shown). To verify whether the region flanking cadA is also conserved in EIEC strains, we performed PCR analysis with a primer pair (B2R and Y1) corresponding to sequences internal to the cadB and yjdI genes (Fig. 2). Since the amplicon obtained has the same size (7.9 kb [Table 3]) and the same PstI restriction pattern (data not shown), we conclude that in our EIEC strains (except EIEC 13.80), the region adjacent to cadA is also colinear with the E. coli K-12 chromosome. This is schematically illustrated in Fig. 2. These results, together with the results of the Southern blot analysis of the region adjacent the cadC gene, strongly suggest that in the majority of EIEC strains the inactivation of the cad region is caused only by localized insertions or deletions (mainly in the cadC gene).

Comparative analysis of the cad region in EIEC and S. flexneri. Data reported above indicate that in EIEC the lack of lysine decarboxylase activity has been attained by different mutational strategies, including point mutations, insertions, and large deletions. Although S. flexneri, the most prevalent Shigella species causing bacillary dysentery, has high homology to EIEC in terms of plasmid and chromosomal sequences, as well as of pathogenicity processes (15, 29, 36), only a single mutational strategy has been reported (9) to account for the LDC− phenotype. To ascertain the possible contribution of convergent evolution to this phenotype, we analyzed seven S. flexneri strains (at least one for each of the serotypes 1 to 6 and representative of the two S. flexneri clusters I and III [Table 1]) isolated from different geographic areas (the United States, Japan, and Africa). While we did not obtain amplicons corresponding to cadA and cadB sequences by PCR analysis (Table 4), we observed cadB hybridization signals in all strains except

---

**FIG. 2.** Comparative analysis of the cad operon of EIEC and E. coli K-12. Arrows indicate the orientation of the IS sequences and of the cad open reading frames. The locations of the primers used are shown below the appropriate region. The positions of the PstI restriction sites were derived from PCR amplicons and confirmed by sequence analysis. Sequences of E. coli K-12 MG1655 were as reported previously (4).

**FIG. 3.** Complementation assay with a functional cadC gene. Cells with (+) or without (−) pCC55, a recombinant plasmid carrying a wt cadC gene, were grown anaerobically in modified Falkow lysine decarboxylase medium at pH 5.5. The amount of cadaverine excreted into the medium was determined by measuring the OD340 of TNP-cadaverine and is expressed as micromoles of cadaverine (OD 340 corresponds to 250 μmol). Values are the average of three independent experiments. For this assay, we used HN13-1, a tetracycline-sensitive mutant of HN13 previously isolated in our laboratory (24).
SFZM49 and SFZM43. As shown in Fig. 1, the presence of sequences homologous to the cadB gene only in a single PstI fragment (6.1 kb), as opposed to the 2.4- and 7.3-kb E. coli K-12 PstI control fragments, indicates that genetic rearrangements causing deletions have altered the S. flexneri cad region. Following PCR amplification with oligonucleotide pairs I1 and B2R, we were able to detect an additional IS1 fragment (6.1 kb), as opposed to the 2.4- and 7.3-kb E. coli cadB flanking remnants of the S. flexneri genes were completely deleted. All strains conserving remnants of the cad locus displayed the same 1.2-kb product (instead of a 2.5-kb fragment [Table 4]).

The 1.2-kb fragment generated using S. flexneri M90T DNA as template was cloned into pGem-T to obtain pCC90. Sequence analysis of pCC90 revealed that an IS1 insertion sequence was present upstream of the cadB gene and that cadC was completely deleted. All S. flexneri strains containing remnants of the cad genes were similar to M90T (Fig. 4), as confirmed by PCR analysis with oligonucleotide pairs I1 and B2 (Table 4). The recently published genome sequences of two S. flexneri serotype 2a strains, SFZM49 and SFZM43, are shown schematically. The same organization is displayed by S. flexneri serotype 2a SFZM46, SFZM50, YSH6000, SFZM53, and M90T DNA. The amplicon obtained has the same size (4.7 kb) as the same 1.2-kb product (instead of a 2.5-kb fragment [Table 4]).

To test for structural changes in the region flanking the cadB remnant of our S. flexneri strains, we performed a PCR analysis with primer Y2, internal to the yfaA gene, and primer B2R (Table 4). The amplicon obtained has the same size (4.7 kb) and the same PstI restriction pattern as that of S. flexneri 2457T and SFZM49. By PCR analysis with the oligonucleotide pair I1 plus B2R, we were able to detect an additional IS1 element (Fig. 4), located as in the previously described serotype 2a strains (9, 14, 43). Nevertheless, as illustrated in Fig. 1, the complete absence of hybridization signals from any of the cad probes in SFZM49 (serotype 3, cluster I) and in SFZM43 (serotype 6, cluster III) indicates that besides the rearrangement reported in Fig. 4, the LCD− phenotype in S. flexneri can also be obtained by deletion of the entire cad region. A PCR-based analysis was performed to verify whether the linear arrangement of the genes flanking the cad locus was similar to that described in S. flexneri 2a (9, 14, 43). The results obtained indicate that in these strains the colinearity with S. flexneri 2a is maintained, since the yfaA and dsbD genes are conserved in the same chromosome location upstream and downstream, respectively, of the cad disrupted locus (data not shown).

**DISCUSSION**

Enteric bacteria like E. coli, Salmonella, and Shigella experience dramatic pH fluctuations in nature, particularly during pathogenesis. To face the extremely low pH of the stomach, as well as the passage of volatile fatty acids in the intestine, these bacteria have evolved complex, acid-inducible survival strategies. The low pH induces a programmed molecular response which, among other things, activates the transcription of several amino acid decarboxylases acting as pH-homoostatic systems (2). In Salmonella and E. coli, the three major amino acid decarboxylase systems operate on glutamate, arginine, and lysine and are able to increase the pH by consuming a proton during decarboxylation and then exchanging the decarboxylation end product for a new substrate via a membrane-bound antporter (2). For the cad operon, it has been shown that the acid-induced synthesis of cadaverine from lysine by CadA and its subsequent excretion through the lysine-decadaverine antporter CadB lead to neutralization of the external pH, thus protecting the cell from acidification (20, 42). Moreover, in vivo experiments have revealed that in E. coli and in S. enterica serovar Typhimurium the induction of the cadBA operon by low pH raises the pH and provides a survival advantage over cad mutants (2).

Although the cad system may be important during bacterial transit through the intestinal tract, in Shigella and EIEC such a system has been completely lost (9, 18, 39). Recently, a detailed analysis of the inflammatory process induced by Shigella has revealed that cadaverine, the end product of pH-induced lysine decarboxylation, attenuates the ability of the bacteria to induce secretion of PMN across model intestinal epithelia (19). The migration of PMN is critical step in Shigella...
pathogenicity process since it facilitates bacterial invasion of the basolateral membrane domain (36). Cadaverine seems to reduce the ability of S. flexneri to lyse the phagocytic vacuole, thus preventing the bacterium from interacting with the cytoskeleton. On the basis of this observation, it has been postulated that loss of the ability to disseminate intra- and intercellularly perturbs signaling pathways that are essential to the promotion of PMN transepithelial migration (12).

It has been suggested that evolution from E. coli ancestors to pathogenic lineages expressing the Shigella phenotype has occurred several times and is the result of acquisition of the main virulence traits through horizontal gene transfer followed by the occurrence of so-called pathoadaptive mutations, i.e., mutations affecting resident genes, which, if expressed, would lower the pathogenic potential (11, 16, 25, 29, 40). Although EIEC strains share essential steps of the pathogenicity pathway with Shigella (23, 35), their metabolic activity is nevertheless more closely related to that of the commensal E. coli strains (39). It is still an open question whether EIEC strains are potential precursors of Shigella or a distinct form, albeit still better adapted to life outside the host (16, 29).

Like Shigella, EIEC strains lose the ability to use lysine as a carbon source. In this work, we have analyzed the molecular rearrangements in the cad operon which determine the silencing of the lysine decarboxylase activity in a group of EIEC strains and have compared them to those present in several S. flexneri strains. Our results show that in EIEC, the major event triggering the loss of the LDC⁺ phenotype is the transposition of IS1 or IS2 elements into the cadC locus (Fig. 2). This gene encodes a regulator essential for the activation of the cadBA operon, and its insertional inactivation results in the complete silencing of cadBA expression (10, 21, 42). All the EIEC strains we have examined from Eastern Africa harbor an IS2 sequence exactly in the same position in cadC. The lack of lysine decarboxylase activity depends only on this insertion, since expression in trans of a functional CadC regulator restores cadaverine secretion to the E. coli K-12 standard level (Fig. 3). Also, EIEC 4608, a strain which differs in serotype, geographic origin, plasmid profile, and lactose utilization from the other EIEC cadC::IS2 strains, contains an IS2 element at exactly the same location (Fig. 2). Inactivation of cadC seems to be a common strategy for silencing cadA, since another strain (53638 [Table 1]) carrying a mutation in the cadC promoter (Fig. 1) is complemented by the introduction of a wt cadC gene (Fig. 3). Furthermore, the fact that cadC is the preferential target of a convergent evolution driving the bacterium toward the LCD⁻ phenotype is confirmed by the presence in EIEC strain 6.81 of an IS1 element which, unlike the IS2 insertion, has provoked a partial deletion of cadC (Fig. 2). A possible explanation of the inactivation of the cadC gene derives from evidence indicating that the CadC regulator may play a direct or indirect role in the control of other genes or operons involved in the adaptation of the bacterium to the host environment. In this respect, it has been recently reported (32, 33) that the mere expression of the CadC protein, even in the absence of cadaverine, leads to a reduction in the amount of both OmpF and OmpC porins, thus altering the permeability of the outer membrane.

An IS1 sequence is always inserted in S. flexneri between the cadB remnant and pheU, and the rearrangements induced by this element probably account for the complete deletion of cadC observed in S. flexneri (Fig. 4). In this context, EIEC 6.81 may represent the “missing link” in the evolutionary transition toward the genetic organization of most S. flexneri strains, where rearrangements are more severe. In fact, in EIEC 6.81 the cadBA operon is still present although it is inactivated by a frameshift mutation in the cadB gene (Fig. 3), while in S. flexneri it is almost completely deleted and has been replaced by IS1 and IS600 insertions. Moreover, it is worth stressing that unlike EIEC, S. flexneri contains only remnants of cadB and has completely lost colinlinearity with the E. coli K-12 genome in the region downstream of the cadB remnant (Fig. 4). Comparative analysis of different serotypes of S. flexneri (Fig. 1 and 4) suggests that in Shigella the cad operon has undergone a multistep rearrangement, involving not only the insertion of IS elements but also independent deletion events resulting in the lack of the entire cadA and cadC genes. Furthermore, the absence of cad hybridization signals in one EIEC strain (13.80) and in two S. flexneri strains (SFZM43 and SFZM49) (Fig. 1) may represent the final result of several progressive steps leading to complete lack of the cad locus.

IS sequences play a crucial role in bacterial evolution, either directly, by modifying gene expression through deletions or insertions, or indirectly, by promoting homologous recombination (17). In Shigella, the lack of motility has been correlated with insertion of IS1 primarily into the master operon flhDC (1). Also in this case, the insertion of IS1 elements, differing in number and insertion position among the different Shigella species, has determined the appearance of progressive genetic defects, resulting in the complete loss of flagella. Moreover, in Shigella, multiple IS elements often interrupt genes encoding thin aggregative fimbriae or curli, triggering genetic rearrangements resulting in deletions with different extensions (31). The recent availability of the S. flexneri 2a SF301 and 2457T genome sequences further reveals how critical the role of IS elements might be in chromosome dynamics, since these elements are found in astonishingly large amounts compared to those in the commensal E. coli strains (14, 43).

The phylogenetic relationships of Shigella strains to each other and to E. coli strains have been analyzed by sequencing eight housekeeping genes in four separate regions around the chromosome (16, 28, 29). These data show that the majority of Shigella strains fall into three main clusters within E. coli and that EIEC strains, as well as S. sonnei strains and some S. dysenteriae strains (D1, D8 and D10), are outside these clusters. Our results show that the majority of S. flexneri strains we have analyzed have undergone the same modification in the cad operon, in excellent agreement with that described for S. flexneri 2a 2457T (9, 43) and SF301 (14). Nevertheless, the complete absence of any remnants of the cad genes in two S. flexneri strains (Fig. 1) indicates that it is difficult to find a correlation between a specific kind of rearrangement and a specific serotype and/or cluster. The disparity of the rearrangements in the EIEC cad region leading to loss of lysine decarboxylase activity provides good evidence of the heterogeneity present in this group of pathogenic E. coli strains (28, 30) and supports the idea that the lack of lysine decarboxylase activity is accomplished through diverse strategies within the E. coli species.

Our results suggest that each time a commensal E. coli strain acquires the invasive plasmid, it encounters a strong selective pressure for silencing the lysine decarboxylase gene. In the
transition toward pathogenicity, the bacterium generally loses some traits that are important for survival in the environment but redundant for the life inside the host. The loss of several catabolite pathways widely present in E. coli, as well as the loss of flagella, exemplifies how Shigella sheds functions that are not crucial for survival inside eucaryotic cells (11, 16, 26, 29, 40). In this case, the same product, cadaverine, which could be useful for buffering low pH in the stomach, is also detrimental for the pathogenicity process (2, 12, 18, 19, 33). In this context, the pathoadaptive mutations responsible for the silencing of cadA could be the result of choosing the lesser of two evils.

ACKNOWLEDGMENTS

We thank P. J. Sansonetti and C. Sasakawa for supplying EIEC and S. flexneri reference strains, G. Micheli for critical reading of the manuscript, and E. Tappi for providing technical assistance.

This work was supported by grants from MIUR (COFIN, Progetti Ateneo), from CNR, and, in part, from the Foundation “Istituto Pasteur-Fondazione Cenci Bolognetti.”

REFERENCES