Analysis of cis-Acting Elements Required for bfpA Expression in Enteropathogenic Escherichia coli

VICTOR H. BUSTAMANTE, EDMUNDO CALVA, AND JOSE LUIS PUENTE*  
Departamento de Microbiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos 62250, Mexico

Received 3 December 1997/Accepted 25 March 1998

bfpA expression in enteropathogenic Escherichia coli is regulated by growth medium, temperature, and ammonium concentration and requires the BfpT protein (also called PerA), a member of the AraC family of transcriptional activators. Site-directed and PCR random mutagenesis, as well as deletion analysis of the bfpA upstream regulatory region, supported assignment of the promoter elements and demonstrated that the cis-acting elements that mediate BfpT-dependent regulation of bfpA are located between positions −85 and −46. Interestingly, this region shares 73% identity with a 40-bp-long AT-rich tract located upstream of the bfpT gene, which is essential for bfpT autoregulation.

Enteropathogenic Escherichia coli (EPEC) is a common cause of diarrhea, particularly among children under 6 months of age living in developing countries (16). Recently, a three-stage model by which EPEC infections proceed has been proposed (6). The initial stage involves the generalized nonintimate interaction of bacterial microcolonies with the surface of epithelial cells, in a pattern known as the localized adherence phenotype (5). This pattern of attachment requires the 80-kb EPEC adherence factor (EAF) plasmid, which contains a cluster of 14 tandemly arrayed genes; this cluster is sufficient to direct the production of the bundle-forming pilus (BFP), a type IV fimbria associated with microcolony formation and bacterial autoaggregation (8, 21, 23–25).

The expression of bfpA, the gene coding for the structural subunit of BFP (23), occurs during the exponential phase of growth, when it is modulated by the growth medium, ammonium concentration, and temperature (19). Our previous studies revealed that bfpA regulation is under the control of a regulatory region that extends further upstream from the putative −35 and −10 promoter sequences, which seems to determine the coordinate regulation of genes located downstream of bfpA (19, 21). This expression is regulated at the transcriptional level and requires the product of the bfpT gene, which is the first gene of the bfpTIVW locus, localized 6.7 kb downstream of the bfp gene cluster on the EAF plasmid (27). bfpT encodes a 274-amino-acid protein, which belongs to the XylS-AraC family of transcriptional regulators (27). The bfpTIVW locus, previously identified as per, has also been involved in the regulation of the eaeA and esp genes, whose products mediate the second and third stages of EPEC interactions with the host cells (6, 9, 12).

Interactions of BfpT with its target sites have been difficult to study in vitro, since different attempts to overproduce and purify it have been unsuccessful. We previously showed that a DNA fragment containing the sequence between nucleotides −94 and −55 of the bfpA regulatory region was bound by a T7-tagged BfpT fusion protein immobilized on Dynabeads; however, attempts to perform footprinting experiments with this fusion were unsuccessful (27). Thus, an alternative route was to genetically analyze the bfpA regulatory region, as presented here.

Deletion analysis of the bfpA regulatory region. A series of 5′ upstream deletions of the bfpA-cat fusion carried on plasmid pCAT232 (19), containing all of the required elements for expression, were constructed by PCR amplification of the corresponding fragments and cloned into vector pK232-8, which contains a promoterless cat gene (2). The nucleotide sequence of all cloned inserts was determined to confirm the precise positions of the deletions and to ensure that no mutations were introduced by the amplification reaction. The chloramphenicol acetyltransferase (CAT) activity directed by plasmids carrying these bfpA-cat deletions was tested in EPEC B171-8 grown in Dulbecco modified Eagle (DME) medium at 37°C, which are the optimal conditions for bfpA expression, and under conditions that are known to regulate bfpA expression, such as growth in Luria-Bertani (LB) medium at 37°C, DME medium at 25 and 39°C, and DME medium containing 15 mM ammonium sulfate at 37°C, as described before (19).

This analysis (Fig. 1) showed that a bfpA-cat deletion down to position −85 (pCAT85) had similar levels of expression and the same regulatory pattern in response to environmental cues as other fusions containing further upstream sequences. Also, a deletion to position −77 (pCAT77) showed an 84% reduction of the BfpT-dependent expression, although, interestingly, it still responded to regulatory signals to the same extent as the wild type. In contrast, only background activity was detected for deletions to position −54 or −40 (pCAT54 or pCAT40), both of which still contain the promoter (Fig. 1 and data not shown). These results indicated that the sequences required for BfpT-dependent expression of bfpA are located upstream of the −35 region and up to position −85.

This AT-rich region contains two 8-bp-long direct-repeat elements, as well as two 10-bp-long inverted-repeat elements, which were designated IRS1 and IRS2 (Fig. 2). Although the precise role of these elements in BfpT binding remains unclear, since they are not sufficient for full activation (Fig. 2), it should be noticed that binding to tandem elements has been reported for other members of the AraC family, such as AraC, MelR, and VirF (3, 14, 28). Moreover, AT-rich sequences are necessary for the regulatory activity of other, closer homologs of BfpT, such as Rns and CfaD (regulation of the CS1 and CFA/I fimbrial operons in enterotoxigenic E. coli, respectively) (11, 18) and VirF (regulation of plasmid-encoded invasion...
proteins in *Shigella* species) (26). Interestingly, another common feature of Rns, CfaD, and VirF is that they seem to overcome the negative regulation by H-NS at their respective promoters, a mechanism that might also account for *bfpA* repression at temperatures below 37°C (11, 18, 26).

**Mutational analysis of the *bfpA* regulatory region.** To pinpoint the position of cis-acting regulatory elements required for *bfpA* expression, random mutations were generated in the *bfpA* regulatory region contained in pCAT232, which was amplified under PCR conditions that enhance error-prone copying, as described previously (15). The PCR products were subcloned back into vector pKK232-8 (Ampr) and transformed into *Escherichia coli* HB101 carrying plasmid pBTA-BH1 (Km r), which contains the *bfpT* regulatory locus (27). Muta-

tions that reduced *bfpA-cat* expression were identified by selecting colonies that did not grow in concentrations of chloramphenicol noninhibitory for strains carrying the wild-type fusion (pCAT232), while transformants carrying mutations that improved *bfpA-cat* expression were screened for their ability to grow in a chloramphenicol concentration that inhibits the growth of a strain carrying the wild-type fusion. Candidates were assayed for CAT activity, as described before (19). Plasmid DNA from these clones was purified and the nucleotide sequence of the *bfpA-cat* regulatory region was determined, allowing the identification of two groups of mutations.

**Promoter mutations.** The sequence of the −35 promoter region of *bfpA* (TTGGGT) contains the most conserved residues of the consensus hexamer (TTGACA) at the first three positions. A T-to-C transition (T-35C) or a G-to-A transition (G-33A), at the first and third positions, respectively (Fig. 2), decreased expression of *bfpA* to the background level, showing that this sequence is critical for *bfpA* expression, in contrast to what is observed for the majority of the positively controlled promoters (20). Furthermore, a G-to-T transversion at position −29 (G-29T; 1 base downstream from the −35 hexamer) produced nearly a twofold increase in *bfpA-cat* expression under all growth conditions tested (Fig. 2 and 3B). This mutation did not modify the transcrip
tional start point (see below; Fig. 4B, lane 2) and caused a ninefold increase in the basal BfpT-independent expression levels (Fig. 3B), suggesting that it generated a stronger promoter. This is consistent with the fact that, in *E. coli* promoters, a T is the most frequently found residue 1 base downstream from the −35 hexamer (10).

Moreover, to better characterize the *bfpA* promoter, two site-directed mutations at the −10 hexamer were independently generated by PCR (1). A T-to-G transversion at position −7 (T-7G) that reduced the identity of the putative −10 region with the consensus sequence abolished *bfpA-cat* expression (Fig. 2), whereas a G-to-T transversion at position −12 (G-12T) (pSNE10-232), which brought the similarity of the putative −10 region closer to the consensus, caused more than a twofold increase in CAT activity, although its regulation in response to environmental cues was similar to that of the wild-type fusion (Fig. 2 and 3B). Interestingly, in the absence of BfpT, the G-12T mutation produced a 30-fold increase in the *bfpA-cat* basal level of expression (Fig. 3B). Primer extension analysis of this promoter mutant showed that transcription initiates at the wild-type position either in the wild-type EPEC strain or in its *bfpT* mutant derivative (Fig. 4A and B, lanes 1, 3, and 4), ruling out the possibility of having generated an alternative promoter. In summary, these results further support the assignment of the *bfpA* promoter (Fig. 2).

**Mutations upstream of the promoter.** Further analysis of mutants with mutations randomly generated by PCR revealed that single deletions or a single insertion at different positions upstream from the promoter but downstream from the −85

---

**FIG. 2.** Nucleotide sequence alignment of the *bfpA* (upper line) and *bfpT* (lower line) −5′ regulatory regions. The upstream regulatory sequence up to position −85 contains all of the cis-acting elements required for the BfpT-dependent activation of *bfpA* (this work). The *bfpA* sequence between positions −85 and −46 shares a 73% identity over a 40-bp-long region (shaded bar) with the sequence between positions −65 and −26 of the *bfpT* regulatory region, which has been shown to be required for *bfpT* autoactivation (17). Brackets enclose the region bound by a T7-tag-BfpT fusion protein (27). Thin and thick broken arrows indicate the positions of each deletion and mutation, respectively, that affected CAT activity. The activities listed in parentheses are expressed as percentages of pCAT232 activity, which was assigned a value of 100% (see Fig. 1). Horizontal arrows above or below the sequences denote the inverted (IRS) or direct-repeat elements. Promoter elements −35 and −10 and the transcription start sites are also indicated.
position (e.g., an insertion or a deletion of one A at the 10-A tract between positions 265 and 274, a deletion of one T at the 8-T tract between positions 244 and 251, or a deletion of one T between positions 241 and 242 [plasmids pMG60, pMG51, pMG58, and pMG65, respectively]) decreased bfpA expression to less than 2% (Fig. 2 and 3A). Interestingly, this reduced level of expression still required BfpT and was regulated by the growth medium, temperature, and ammonium concentration (Fig. 3A).

This negative effect could have resulted from slight but significant local distortions in the DNA spatial structure, which would bring out of phase the BfpT-binding sites with respect to the promoter, probably altering its interactions with other molecules, i.e., RNA polymerase. In this regard, it has been observed for other regulatory proteins, such as CRP and FNR, that the exact spacing of their binding sites with respect to the promoter is crucial for activation (7, 29). Further analysis of site-directed mutants with full or half-turn insertions will be required to test this hypothesis. In contrast, two mutants with an A-to-G transition in the same region (A-65G or A-66G) rendered only a moderate positive effect on bfpA expression (Fig. 2).

The PCR random mutagenesis strategy did not render a wider variety of mutations, as illustrated by those that were generated by site-directed mutagenesis. In this respect, it is also possible that the effect of other mutations is not large enough to be detected by our screening procedure, contrasting with the larger effect caused by several single-base deletion or insertion mutants, which allowed their easy and recurrent isolation. In summary, new mutagenesis and screening schemes should be explored to exhaust all the possibilities.

The BfpT-independent expression of a bfpA promoter mutant is still repressed in LB medium. The BfpT-independent expression showed by pSNE10-232 (Fig. 5) was still repressed also possible that the effect of other mutations is not large enough to be detected by our screening procedure, contrasting with the larger effect caused by several single-base deletion or insertion mutants, which allowed their easy and recurrent isolation. In summary, new mutagenesis and screening schemes should be explored to exhaust all the possibilities.

The BfpT-independent expression of a bfpA promoter mutant is still repressed in LB medium. The BfpT-independent expression showed by pSNE10-232 (Fig. 5) was still repressed...
in LB medium but not by ammonium. Furthermore, a derivative of this mutant with a deletion to position −40 (pSNE10-40) behaved in the same manner (Fig. 5). These observations suggested that the different levels of bfpA expression in LB and DME medium could be mediated by a mechanism that acts directly on its promoter, while ammonium repression occurs through a different mechanism that requires BfpT and sequences upstream of the promoter. Moreover, since bfpA expression is selectively repressed upon entrance to stationary phase, we cannot exclude the possibility that the different levels of expression in DME and LB medium depend, at least partially, on how long the exponential phase of growth is sustained and that this phenomenon might be directly or indirectly mediated by RpoS (19, 30).

The bfpA and bfpT regulatory regions share a common motif. Recently, we have observed that bfpT expression is autoregulated and also modulated by the growth medium, temperature, and ammonium concentration (17). Considering these observations, we expected that common elements could be present in the regulatory regions of bfpA and bfpT. The nucleotide sequence alignment of these regions revealed the presence, as part of the minimal regulatory region of bfpT, of a sequence that shares 73% identity with the region between residues −85 and −46, which was shown to mediate regulation and BfpT-dependent expression of bfpA (Fig. 2). In contrast, no significant sequence similarities could be found with the bfpA region upstream from position −84 or downstream from position −46 (Fig. 2). Interestingly, the sequence comprised between positions −84 and −65, which proved to be critical in bfpA activation and is part of the bfpA-bfpT homologous motif, is located two full turns further upstream in bfpA with respect to bfpT, suggesting that BfpT can activate transcription from different locations with respect to the promoter, as long as the correct phase is maintained (Fig. 2), as has been described for many regulatory proteins in E. coli (4, 7, 29).

Concluding remarks. This study led us to determine that the sequence between positions −85 and −55 is essential for the BfpT-dependent activation and ammonium regulation of bfpA, probably constituting the BfpT-binding motif. The region between positions −55 and −35, which resembles an UP element (13, 22), probably accounts for the stronger promoter activity shown by bfpA in comparison with that of bfpT, which lacks this element (Fig. 2) (17), although this hypothesis remains untested. All of this provides the basis toward further understanding the molecular mechanisms that control the expression of BFP and possibly other virulence factors in EPEC.

We particularly thank Francisco Santana for excellent technical assistance. We thank Enrique Morett for critical reading of the manuscript. J.L.P. thanks Dave Bieber for his collaborative effort and many helpful discussions at the early stage of this work.

V.H.B. was supported by a Ph.D. fellowship from the Consejo Nacional de Ciencia y Tecnología (CONACyT 1027P-N).

REFERENCES


