Differential binding of Lrp to two sets of *pap* DNA binding sites mediated by PapI regulates Pap phase variation in *Escherichia coli*

Xiangwu Nou, Bruce Braaten,
Linda Kaltenbach and David A.Low¹

Division of Cell Biology and Immunology, Department of Pathology,
University of Utah, Health Sciences Center, Salt Lake City, UT 84132,
USA

¹Corresponding author

Pyelonephritis-associated pili (Pap) expression in *Escherichia coli* is subject to a phase variation control mechanism that is regulated by the leucine-responsive regulatory protein (Lrp), PapI, and deoxyadenosine methylase (Dam). In previous work, we found that the differential Dam methylation of two target sites in *pap* regulatory DNA, GATC-I and GATC-II, is essential for the transition between active and inactive *pap* transcriptional states. Here, we identify six Lrp binding sites within the *pap* regulatory DNA, each separated by about three helical turns. Lrp binds with highest affinity to three sites (1, 2 and 3) proximal to the *papBAP* promoter. A mutational analysis indicates that the binding of Lrp to sites 2 and 3 inhibits *pap* transcription, which is consistent with the fact that Lrp binding site 3 is located between the –35 and –10 RNA polymerase binding region of *papBAP*. The addition of PapI decreases the affinity of Lrp for sites 1, 2 and 3 and increases its affinity for the distal Lrp binding sites 4 and 5. Mutations within Lrp binding sites 4 and 5 shut off *pap* transcription, indicating that the binding of Lrp to this *pap* region activates *pap* transcription. The *pap* GATC-I and GATC-II methylation sites are located within Lrp binding sites 5 and 2, respectively, providing a mechanism by which Dam controls Lrp binding and Pap phase variation.

Keywords: DNA–protein interaction/Lrp/*pap* operon/phase variation

Introduction

The regulation of pyelonephritis-associated pili (*pap*) expression in *Escherichia coli* is under a complex phase variation control mechanism which involves leucine-responsive regulatory protein (Lrp), PapI, and deoxyadenosine methylase (Dam; Blyn *et al.*, 1990; Braaten *et al.*, 1992, 1994; van der Woude *et al.*, 1992). Lrp is a 19 kDa global regulator which binds to DNA as a dimer (Platko and Calvo, 1993; Wang and Calvo, 1993). Lrp activates some genes and represses others and, in many cases, leucine modulates Lrp activity. However, the *pap* operon is not responsive to leucine. Moreover, Lrp requires PapI as a cofactor to activate *pap* transcription, whereas Lrp appears to act by itself to regulate other members of the Lrp regulon such as *ilvIH* (Nou *et al.*, 1993; Calvo and Matthews, 1994).

Two DNA GATC sites which are targets for Dam methylase are located within the *papI–B* regulatory region. The GATC-II site is 50 bp upstream of the *pBAP* transcription start site, whereas GATC-I is 102 bp upstream of GATC-II (Figure 2A). We have found previously that the GATC-II site of phase OFF cells is protected from methylation, whereas GATC-I is fully methylated (Blyn *et al.*, 1990). Methylation protection of GATC-II requires Lrp but not PapI (Braaten *et al.*, 1991). *In vitro* footprint analyses have shown that in the absence of PapI, Lrp bound close to the GATC-II site (Nou *et al.*, 1993). Together, these data suggest that in phase OFF cells, Lrp binds near GATC-II, thus sterically blocking its methylation. Our recent results indicate that the binding of Lrp near GATC-II blocks basal transcription from the *papBAP* promoter (van der Woude *et al.*, 1995).

Phase ON cells have a methylation pattern that is the converse of phase OFF cells: GATC-I is protected from methylation whereas GATC-II is fully methylated. An *in vitro* footprint analysis showed that the addition of Lrp and PapI to non-methylated *pap* DNA resulted in footprinting of the GATC-I region as well as the GATC-II region (Nou *et al.*, 1993). These results are consistent with the finding that both Lrp and PapI are required for the protection of GATC-I from methylation *in vivo* (Braaten *et al.*, 1991). Recent data suggest that the binding of Lrp, in the presence of PapI, to the GATC-I region activates transcription ~8-fold over the basal transcription level. Lrp binding to the GATC-I region is not sufficient for transcription activation because Lrp mutants which bind to *pap* with the same affinity as wild-type Lrp fail to induce transcription (Platko and Calvo, 1993; van der Woude *et al.*, 1995). These results indicate that Lrp interacts with other parts of the transcriptional apparatus, such as RNA polymerase, when it is bound at the GATC-I region.

Methylation appears to play a direct role in controlling *pap* transcription. First, transcription is shut off in the absence of Dam or in the presence of excess Dam (4-fold above wild-type levels; Blyn *et al.*, 1990). Second, the mutation of the *pap* GATC-I site to the GCTC sequence results in a locked ON phenotype, even in the presence of high Dam levels (Braaten *et al.*, 1994). These results indicate that the methylation of GATC-I shuts off *pap* transcription. A DNA footprint analysis showed that the binding of Lrp to the GATC-I region, facilitated by PapI, was blocked by DNA methylation (Nou *et al.*, 1993). Based on these data, we hypothesized that Lrp and Dam methylase compete for binding to the GATC-I region, thereby controlling the switch between transcriptionally active and inactive states (van der Woude *et al.*, 1992).

Although the methylation of GATC-I inhibits *pap* transcription, the methylation of GATC-II appears to be required for *pap* transcription. Both wild-type and phase

locked ON GCTC-I mutant *E.coli* strains are transcriptionally inactive in a *dam*⁻ background (Braaten *et al.*, 1994). Moreover, the mutation of GATC-II to the GCTC sequence results in a locked OFF phenotype. This phenotype was not caused by the inhibition of binding of Lrp to the GCTC-II region, but appeared to be a direct result of the inability of this site to be methylated by Dam. Preliminary data suggested that the methylation of GATC-II might reduce the affinity of Lrp for this site, aiding in the binding transition of Lrp from GATC-II to GATC-I and the activation of *pap* transcription (Braaten *et al.*, 1994).

Previous analyses using DNase I footprinting indicated that Lrp bound to a 115 bp region around GATC-II in the absence of PapI, and footprinted an additional 60 bp region around GATC-I after PapI addition (Nou *et al.*, 1993). These footprints showed periodic regions of DNase I protection and hypersensitivity, characteristic of bent DNA. Because the cleavage of DNA by DNase I is affected not only by protein binding but also by the DNA structure, it was not possible to determine the regions of *pap* DNA contacted by Lrp. Further, the roles of Lrp binding sites in regulating *pap* phase variation were not explored. Here, we have used a methylation protection analysis to identify six *pap* Lrp binding sites within the *pap* regulatory region, each spaced about three helical turns apart. The GATC-I and GATC-II sites are each located within an Lrp binding site. A mutational analysis indicated that the binding of Lrp to sites 2 and 3 proximal to the *papBap* promoter inhibited transcription, whereas the binding of Lrp to the promoter distal sites 4 and 5 activated transcription. Our results show that PapI facilitates the transition from phase OFF to ON by reducing the affinity of Lrp for the *papBap* promoter proximal binding sites and increasing Lrp affinity for the distal *pap* binding sites.

Results

Identification of Lrp binding sites in the *pap* regulatory region.

We showed previously that Lrp protected a 115 bp region around the GATC-II site of *pap* DNA from DNase I cleavage. PapI and Lrp together footprinted an additional 60 bp around the GATC-I site (Nou *et al.*, 1993). These large footprints consisted of periodical protected and hypersensitive DNA regions, which may have resulted from DNA bending as well as the specific binding of Lrp (Hochschild, 1991). To localize further each of the *pap* Lrp binding sites, we used a methylation protection analysis with dimethyl sulfate (DMS) to identify the guanosine and adenosine residues in close contact with Lrp (Wissman and Hillen, 1991). As shown in Figure 1, guanosines in both the top strand (designated G_n) and bottom strand (designated C_n for the complementary guanosine) of the *pap* regulatory region were protected from DMS methylation by Lrp addition. These and other methylation protection data (results not shown), summarized in Figure 2A and B, indicate that the *pap* regulatory region contains six Lrp binding sites. Each of these sites is spaced 30–33 bp apart, center to center, or about three turns of a DNA helix. At lower Lrp levels, the binding of Lrp to sites 1, 2 and 3 occurred (designated

here by [1,2,3]). At higher Lrp levels, sites 4, 5 and 6, designated [4,5,6], were also occupied (Figure 1). The addition of PapI did not result in any new protected bases, although PapI did alter the affinity of Lrp for the six sites identified (see below; Figures 3 and 4). All six Lrp binding sites contain the sequence 'Gnn(n)TTTt' (where 'n' is any nucleotide and 't' indicates that a fourth thymidine was present in all sites except site 4). The 'G' of the consensus binding site was protected from methylation in all six binding sites. Additional guanosines were protected in all sites except site 6. Four adenosines complementary to the consensus thymidines were also protected from methylation in sites [1,2,3]. We were unable to determine if the adenosines in sites [4,5,6] were protected because piperidine cleavage does not occur as efficiently at methylated adenosines compared with methylated guanosines (Wissman and Hillen, 1991). As shown in Figure 2C, the 'G' of the consensus binding site and the 'A' tract are positioned opposite one another across the DNA major groove.

Mutational analysis of the roles of Lrp binding sites in *pap* transcription.

To determine the roles of the Lrp binding sites identified above in controlling *Pap* phase variation, we introduced DNA base-pair replacements containing *Nhe*I restriction sites throughout the *pap* regulatory region (Figure 2B). Our data, summarized in Figure 2A, show that mutations within Lrp binding site 4 or 5 result in locked OFF *pap* transcription phenotypes, whereas mutations within site 2 or 3 lock cells in the phase ON state. All mutations in or close to Lrp binding site 2 or 3 resulted in an increased OFF to ON switch frequency or a locked ON phenotype based on an analysis of strains containing a single-copy *papBap-lac* fusion (Table I). *E.coli* containing mutation #13 (site 3) displayed a locked ON transcription phenotype, as did *E.coli* containing mutation #24 (adjacent to site 2). Both of these mutant strains expressed high levels of β-galactosidase (~5800 Miller units). This level of transcription is ~3-fold higher than that observed for a wild-type population containing 30% phase ON cells (Table I), which is consistent with the locked ON phenotypes of these mutants. Mutants #16 and #17, which contain single base-pair substitutions of G1163 within site 3, also expressed high levels of β-galactosidase and a locked ON phenotype (Figure 2B and Table I). These data suggest that the 'G' in the consensus sequence may be important for Lrp binding (confirmed below). *E.coli* containing mutation #12 (site 2) or #11 (adjacent to site 1) displayed altered switching phenotypes with an ~8-fold higher OFF to ON switch rate compared with wild type. Mutant #11 also had a 2-fold lower ON to OFF switch rate, which resulted in a higher *papBA* transcript level as well as a higher fraction of cells in the phase ON state (Table I). Together, these results show that mutations within or near Lrp binding sites [1,2,3] resulted in either locked ON phenotypes or increased rates of switching from OFF to ON.

In contrast to mutations near or within binding sites [1,2,3], mutations #56 (site 4), #14 (site 5) and #10 (between sites 6 and 1) resulted in locked OFF phenotypes and very low transcript levels (Figure 2A and Table I). Mutation #7, a G1028 to T1028 transversion within the

Fig. 1. The localization of Lrp binding sites in the *pap* regulatory region by methylation protection analysis. The amounts of Lrp shown were added to *pap* regulatory DNA, end-labeled with ^{32}P on the top (A) or bottom (B) DNA strand. A methylation protection analysis using DMS was carried out as described in Materials and methods, and samples were analyzed on polyacrylamide gels. The locations of *pap* top strand guanosines (see *pap* DNA orientation in Figure 2B) protected by Lrp addition are shown in (A) and the positions of *pap* bottom strand guanosines protected by Lrp addition are shown in (B). Bottom strand guanosines are designated as C_n based on the top strand position of the complementary cytosine. The locations of the six Lrp binding sites in the *pap* regulatory region, based on these and additional methylation protection data (results not shown), are indicated to the right.

GATC-I sequence of site 5, also gave a locked OFF phenotype (Figure 2A and B). Mutation #9, located adjacent to site 6, did not significantly affect *pap* switch rates or transcript levels, whereas the complete substitution of site 6 (mutant #57) resulted in only a 2-fold reduction in the phase OFF to ON rate (Table I). Thus, site 6 does

not appear to play a significant role in the regulation of *pap* transcription. Taken together, these results indicate that the binding of Lrp to sites [4,5] has a positive regulatory function. In addition, the *pap* DNA sequence between sites 1 and 6 (defined by mutation #10) also plays a positive role (see Discussion).

Fig. 2. The location of Lrp binding sites in the *pap* regulatory region. (A) The *pap* regulatory region and a summary of the mutation analyses. In the top half, the six Lrp binding sites identified in this study are shown as rectangles. The *pap* GATC-I and GATC-II sequences within sites 5 and 2, respectively, are shown as open circles. The -35 and -10 RNA polymerase binding region for the *papB* promoter is shown as two dark boxes around Lrp binding site 3. The bottom half shows the positions of each of the *pap* mutations and their transcriptional phenotypes: OFF, phase locked off; SW, reversible switching between phase OFF and ON states occurs; ON, phase locked ON; (ON), high OFF to ON rate. (B) The locations of Lrp binding sites and mutations within the *pap* regulatory region. Filled circles show guanines and adenines that are protected from DMS methylation after the addition of a low level ($1.25 \mu\text{g/ml}$) of Lrp (sites [1,2,3]). Filled circles above the sequence show the methylation protection of the top strand, whereas circles below the sequence show the methylation protection of the bottom strand. Open circles show guanines and adenines that are protected from methylation at a higher level of Lrp ($5 \mu\text{g/ml}$). Boxed regions show the six Lrp binding sites, each containing the consensus sequence Gnn(n)TTT(t), where 'n' is any nucleotide. Lrp site 6 is shown as a light box and extends only through the binding site consensus sequence because only a single guanine was protected from methylation in this region. The arrows indicate the relative orientation of each Lrp binding site. The positions of substitution, transversion and transition mutations are shown below the DNA sequence by bars, with the substituted (mutated) base pairs in italics. The -35 and -10 RNA polymerase binding region for the *papB* promoter is shown, as well as the start site for transcription of the *papB* gene. (C) A planar representation of the DNA region around Lrp binding site 2 which contains GATC-II. The positions of bases are projected onto the surface of a cylinder that is then unrolled onto a flat surface. Closed circles show guanines and adenines that are protected from methylation due to Lrp binding.

To determine the relative dominance of sites [1,2,3] compared with sites [4,5] on *pap* transcription, we constructed a mutant (#47) containing mutations #14 (locked OFF) and #13 (locked ON). This double mutant displayed a locked ON phenotype with a β -galactosidase level similar to mutant #13 (Table I). These results indicate that mutation #13 is dominant to mutation #14 (see Discussion).

The binding of Lrp near GATC-I and GATC-II appears to inhibit the methylation of these sites by Dam (Nou et al., 1993). Therefore, the methylation patterns of these sites provide information regarding the binding of Lrp to *pap* in vivo. An analysis of the GATC site methylation patterns of *pap* switch mutants #13 and #24 showed that the GATC-I sites of these locked ON mutants were almost completely protected from methylation, whereas the GATC-II site was fully methylated. These results suggest that Lrp is bound near GATC-I in these mutants, similar to phase ON wild-type cells (Blyn et al., 1990; Nou et al., 1993; Braaten et al., 1994). Mutant #11, which had a higher fraction of phase ON cells because of a lower ON to OFF switch rate, showed a proportional increase in the methylation protection of GATC-I compared with wild-type cells (Table I). In contrast, the locked

OFF mutants #10 and #14 showed the almost complete protection of GATC-II from methylation, whereas the GATC-I site of mutant #10 was fully methylated (mutant #14 lacks a GATC-I site). This methylation pattern suggests that Lrp binds close to GATC-II, similar to phase OFF cells.

The effects of Dam, Lrp and PapI on transcription from mutant *pap* constructs

Recent evidence indicates that Dam methylase plays a dual role in regulating the *pap* transcriptional switch (Braaten et al., 1994). The methylation of GATC-I inhibits the formation of the phase ON state, whereas the methylation of GATC-II is required for the phase ON state. Thus, *pap* transcription is shut down in a *dam*⁻ or a Dam overproducer (~ 4 - to 69 -fold) background (Blyn et al., 1990; Table III). The introduction of mutant *dam* alleles into the *pap* switch mutants showed that mutant #13 remained locked ON in the absence of Dam or under high Dam levels, and thus is totally Dam independent. Mutant #12, which showed a higher rate of switching compared with wild-type cells under normal Dam levels (Table I), became locked ON in the absence of Dam. However, *pap*

transcription in mutant #12 was shut off at high Dam levels, similar to wild-type *pap* (Table III). A methylation analysis showed that at high Dam levels both the GATC-I and GATC-II sites of wild-type and mutant #13 DNAs were fully methylated and the GATC-I site of mutant #12 DNA was also fully methylated (Table III). An analysis of mutant #14 (locked OFF) showed that it remained locked OFF in the absence of Dam or under high Dam levels.

Mutations #12 and #13 are in close proximity to the -35 and -10 RNA polymerase consensus binding sites of the *papBAP* promoter, raising the possibility that the aberrant phase variation phenotypes of *E. coli* containing these mutations might be caused by an alteration in promoter activity rather than the disruption of Lrp binding. If this was so, then it is likely that *pap* transcription would no longer be dependent on Lrp (Braaten *et al.*, 1991). To determine if *pap* switch mutants #12 and #13 are dependent on Lrp, we introduced *lrp* null mutations into both wild-type and regulatory mutant cells. The transcription of *pap* in mutants #12 and #13 was shut off in the absence of Lrp, showing that these regulatory mutants remain dependent upon Lrp for *pap* transcription (Table II).

PapI is required for the methylation protection of the GATC-I site as well as *pap* transcription in *E. coli* containing a wild-type *pap* sequence (Braaten *et al.*, 1991; Nou *et al.*, 1993). Similar to wild-type *pap*, mutant #12 cells were dependent on PapI for transcription and showed little, if any, methylation protection of GATC-I in a *papI* background, consistent with their locked OFF phenotypes (Table II). However, mutant #13 remained locked ON in the absence of PapI, even though a methylation analysis showed that the GATC-I site was not protected from methylation under these conditions (Table II). Together, these results show that *pap* transcription in mutant #12, like wild-type cells, is dependent on both Lrp and PapI, whereas transcription in mutant #13 requires Lrp but is independent of PapI.

Determination of Lrp binding affinities for wild-type and mutant *pap* DNAs

To more fully understand the effects of mutations in the *pap* regulatory region, we analyzed the binding of Lrp, in the presence and absence of PapI, to both wild-type and mutant DNAs *in vitro* using quantitative DNase I footprinting (Figure 3). Lrp bound with highest affinity to sites [1,2,3] of wild-type *pap* DNA and with lower affinity to sites [4,5,6] (Figures 3A and 4A). The addition of PapI altered the affinity of Lrp for its DNA binding sites. With PapI, Lrp had the highest affinity for sites [4,5]. In contrast, the affinity of Lrp for the remaining four sites was reduced. A quantitative analysis showed that Lrp has an ~3-fold higher affinity for the GATC-II DNA region (sites [1,2,3]) than it does for the GATC-I region (sites [4,5]; Figure 4A). The addition of PapI increased the affinity of Lrp for the GATC-I region by ~4-fold and decreased the affinity of Lrp for the GATC-II region by almost 2-fold. Thus, the net effect of PapI is to shift the binding of Lrp from the GATC-II region (sites [1,2,3]) to the GATC-I region (sites [4,5]).

An analysis of mutant #12 (Figures 3B and 4B), which has an altered GATC-II site, showed that the affinity of Lrp for the GATC-II region (sites [2,3]) was reduced ~3-fold compared with wild-type *pap*. In contrast, the affinity

of Lrp for sites [4,5] increased ~1.5-fold over the wild-type level. PapI increased the affinity of Lrp for sites [4,5] and reduced the affinity of Lrp for site 1, similar to wild-type *pap* DNA. These results are consistent with the increased OFF to ON switch rate of this mutant because the binding of Lrp to the GATC-II DNA region, which appears to act negatively on *pap* transcription, was significantly reduced. Moreover, the high affinity binding of Lrp to sites [4,5] around GATC-I, which appears to activate *pap* transcription, still required PapI, consistent with the PapI-dependent phenotype of this mutant.

An analysis of mutant #13, which contains a 6 bp replacement within site 3, showed that the binding of Lrp to sites [2,3] was reduced by ~3-fold, similar to that observed for mutant #12. However, Lrp had a higher affinity for sites [5,6] of mutant #13 DNA compared with either wild-type or mutant #12 DNAs (Figure 3C). Moreover, PapI did not affect the affinity of Lrp for site 5, which contains GATC-I, of mutant #13 DNA. Together, these results are consistent with the PapI-independent phenotype of this mutant (Table II) because the high affinity binding of Lrp to GATC-I (site 5) occurred in the absence of PapI. Similar footprinting results were obtained for mutant #24, which contains a 6 bp replacement within Lrp site 2, and mutants #16 and #17, which contain point mutations within site 3, all displaying locked ON phenotypes (results not shown).

Mutant #14 contains a 6 bp replacement of the GATC-I site, resulting in a locked OFF transcription phenotype (Table I). As shown in Figure 3D, the mutation of the GATC-I site did not reduce the affinity of Lrp by itself for sites 4 or 5, although the affinity of Lrp for sites [6,1,2,3] were reduced to the level that Lrp has for wild-type *pap* DNA in the presence of PapI. Strikingly, PapI did not increase the affinity of Lrp for sites [4,5] around the mutant #14 GATC-I site, as occurred with the wild-type *pap* GATC-I site. Thus, the binding transition of Lrp from the GATC-II to the GATC-I region, mediated by PapI, is blocked in mutant #14 and cells remain in the transcriptionally inactive state. These results indicate that the GATC-I sequence is necessary for the binding of Lrp/PapI to sites [4,5].

Discussion

The data presented here show that the transition between active and inactive *pap* transcriptional states is regulated by the binding of Lrp to two sets of *pap* DNA binding sites. In the absence of PapI, Lrp binds with highest affinity to sites [1,2,3] around GATC-II of the *pap* regulatory DNA. The binding of Lrp at sites [2,3] appears to inhibit *pap* transcription because mutations within or near these sites resulted in locked ON phenotypes or higher OFF to ON switch rates (Table I). In the presence of PapI, the affinity of Lrp for sites [4,5] increased ~4-fold, and the affinity of Lrp for sites [1,2,3] decreased ~2-fold (Figure 4A), resulting in a translocation of Lrp from the GATC-II to the GATC-I region. Recent data have shown that PapI binds specifically to the Lrp moiety of Lrp-*pap* DNA complexes (Kaltenbach *et al.*, 1995). Thus, it is possible that the affinity of Lrp for sites [1,2,3,4,5] is altered by a conformation change induced by the binding of PapI. It seems likely that the binding of Lrp/PapI to the GATC-I

region activates *pap* transcription because the mutation of the GATC-I site inhibits the binding of Lrp/PapI to sites [4,5] and prevents the formation of the phase ON transcription state (Figure 4D, mutant #14).

Previously, we have shown that methylation plays dual roles in *pap* transcriptional regulation (Braaten *et al.*, 1994). The methylation of the GATC-I site inhibits the binding of Lrp/PapI and shuts off transcription. In contrast, the methylation of GATC-II is required for *pap* transcription, possibly by reducing the affinity of Lrp for this DNA region. These data are consistent with recent evidence which indicates that the binding of Lrp at the GATC-II region inhibits the basal transcription level of the *papBAP* promoter, whereas the binding of Lrp near GATC-I activates *pap* transcription ~8-fold above the basal level (van der Woude *et al.*, 1995). Therefore Lrp acts as a

repressor in the absence of PapI and as an activator in the presence of PapI, because PapI is required to shift the binding of Lrp from GATC-II to GATC-I (Figures 3 and 4). Here, our data indicate that Lrp binds to sites [1,2,3] near GATC-II in the phase OFF state (Figures 3A and 4A). Because Lrp binding site 3 is located between the -35 and -10 RNA polymerase consensus binding region (Figure 2A and B), this provides a mechanism (via steric hindrance) by which the binding of Lrp might directly inhibit *pap* transcription. It is less clear how the binding of Lrp near GATC-I activates *pap* transcription. Based on the analysis of Lrp activation mutants, the binding of Lrp is not sufficient for the activation of *pap* transcription, suggesting that Lrp interacts with another component of the regulatory machinery such as RNA polymerase or catabolite gene activator protein (van der Woude *et al.*, 1995).

Fig. 3. A quantitative DNase I footprint analysis of non-methylated wild-type and Lrp binding site mutant *pap* DNAs. The left portion of each figure (section A) shows the addition of increasing amounts of Lrp ($\mu\text{g/ml}$) to *pap* DNAs, whereas the right portion of each figure (section B) shows the addition of increasing amounts of Lrp in the presence of PapI (2 $\mu\text{g/ml}$). To the left of each figure are shown the *pap* nucleotide positions (corresponding to Figure 1), with the GATC-II site at 1130–1133 bp and the GATC-I site at 1028–1031 bp (both GATC sites are shown as closed circles). To the right, the six Lrp binding sites, identified by a methylation protection analysis, are shown. (A) Wild-type *pap* DNA; (B) mutant #12 DNA; (C) mutant #13 DNA; (D) mutant #14 DNA.

The GATC-I site plays an essential role in the modulation of Lrp binding by PapI, as shown by an analysis of mutant #14 which contains a 6 bp substitution of GATC-I. Although the affinity of Lrp for sites [4,5] of mutant #14 DNA was not affected, PapI did not increase the affinity of Lrp for these sites as it did with wild-type *pap* DNA (compare Figure 3A with D). Thus, the stable binding of Lrp/PapI to sites [4,5] requires the GATC-I sequence. The 'G' of GATC-I appears to be in contact with Lrp (Figure 2B); it is essential for transcriptional activation because a G to T transversion (mutation #7) gave a locked OFF phenotype. Although we have not yet analyzed the binding of Lrp to this point mutant, it seems likely that it will reduce the affinity of Lrp/PapI for sites [4,5], as observed with mutant #14. Our previous results indicated that the adenosine of GATC-I must also come into close contact with Lrp because methylation of this 'A' blocks the binding of Lrp/PapI to the GATC-I region (Nou *et al.*, 1993).

The mutation of binding site 3 (mutant #13) resulted in a locked ON transcription phenotype which, in contrast to wild-type cells, was independent of PapI and Dam

methylation (Tables II and III). An *in vitro* footprint analysis showed that Lrp bound to sites [5,6,1] of mutant #13 DNA with highest affinity (Figure 4C). An *in vivo* analysis of methylation patterns showed that the GATC-I site was almost fully protected from methylation, consistent with the *in vitro* binding results, indicating that Lrp binds with high affinity near site 5, which encompasses the GATC-I region. Although the binding of Lrp to site 5 and the transcription activation occurred in the absence of PapI, the high affinity binding of Lrp to site 4 (Figure 4C) and the methylation protection of GATC-I (Tables II and IV) both required PapI. These data are consistent with the observation that *pap* transcription was activated in a Dam overproducer background, conditions under which both *pap* GATC sites are fully methylated (Table III). This raises the possibility that the binding of Lrp to mutant #13 DNA site 5, containing GATC-I, is not blocked by methylation of this site or is not critical for binding to sites [6,1]. This could result from the fact that Lrp binds to three sites ([5,6,1]) of mutant #13 DNA with high affinity, which might stabilize binding compared with

Fig. 4. Summary of the quantitative footprint analyses of Lrp binding to wild-type and mutant *pap* DNAs. In each panel, the relative affinity of Lrp for each *pap* DNA binding site is shown on the left. The affinity of Lrp for sites [1,2,3] of wild-type *pap* DNA has been set at a value of 1.0. Each of the six Lrp binding sites is shown at the bottom of the figure, including the location of the *pap* GATC-I and GATC-II sites. The binding affinity of Lrp is shown by light bars, whereas the affinity of Lrp in the presence of PapI is shown by dark bars. The binding affinities were calculated as the concentration of Lrp at which protection from DNase I cleavage was one-half maximal. Each panel corresponds to the panels shown in Figure 3.

wild-type *pap* in which Lrp/PapI binds with high affinity to only two sites ([4,5]). This hypothesis was supported by our observation that mutation of site 5 does not block transcription when combined with mutation #13 in site 3 (see double mutant #47, Table I). Thus, in contrast to wild-type *pap*, the activation of transcription in mutant #13 does not require site 5 which contains GATC-I.

Our previous results indicated that the methylation of GATC-II by Dam methylase is essential for *pap* transcription (Braaten *et al.*, 1994). We hypothesized that the methylation of this site might inhibit Lrp binding to the GATC-II region, thus preventing the Lrp-mediated repression of transcription (van der Woude *et al.*, 1995). Our finding here, that mutant #13 remained phase locked ON in the absence of Dam methylase, is consistent with this hypothesis because the binding of Lrp to the GATC-II region is already greatly reduced by the mutation of site 3, precluding the need for methylation (Tables II and III, and Figure 4). Similarly, mutant #12 lacking the GATC-II site was phase locked ON in the absence of Dam (Table III), in contrast to our previous finding that a GCTC-II point mutation, which also prevents Dam methylation, was locked in the OFF state (Braaten *et al.*, 1994). We hypothesize that the different phenotypes of these mutations in GATC-II are a result of the 6 bp substitution mutation in GATC-II (mutation #12) disrupting the binding of Lrp to sites [2,3] (Figure 4B), whereas the GCTC-II point mutation does not (Braaten *et al.*, 1994). Thus, in the latter case, the high affinity binding of Lrp to sites [1,2,3] cannot be disrupted by Dam, which prevents the phase ON transition.

The mutation of site 6 (mutant #57) had only a small effect on *pap* transcription, indicating that the binding of Lrp to site 6 does not play a major role in *Pap* phase variation. In contrast, mutation #10, located between sites 6 and 1, resulted in a locked OFF phenotype (Figure 2A and Table I). Preliminary results indicate that the mutation of site 1 also gives a locked OFF phenotype (unpublished data), suggesting that the binding of Lrp at or close to site 1 is essential for *pap* transcription. In contrast, the Lrp binding at sites [2,3] blocks *pap* transcription because the mutation of these sites results in locked ON phenotypes (Figure 2A). Thus, although Lrp appears to bind to sites [1,2,3] in phase OFF cells, the role of site 1 appears to differ from that of sites [2,3]. One possible reason for this difference in Lrp binding site activity is that the translocation of Lrp from sites [1,2,3] to sites [4,5] might require the interaction of PapI with Lrp bound at site 1. If this is so, then the PapI-induced translocation of Lrp from GATC-II to GATC-I should not occur in *pap* site 1 mutants. Experiments are in progress to test this hypothesis.

Previous work by Wang and Calvo (1993) has shown that there are five Lrp binding sites which play roles in the transcription activation of the *ilvIH* operon. However, the arrangement of these sites and the binding by Lrp differ significantly from the *pap* operon. The spacing between *ilvIH* sites varies from 20 to 34 bp, compared with a spacing of 30–33 bp between the Lrp binding sites of *pap* (Figure 2B). Lrp binds with high affinity ($K_d \sim 6$ nM) to sites [1,2] of *ilvIH*. Lrp has a lower affinity for sites [3,4,5] of *ilvIH*, with a K_d of ~ 22 nM. The latter value is similar to our estimate of the binding affinity of

Table I. Phenotypes of *pap* regulatory mutants

Mutation in <i>pap</i> DNA ^a	Switch frequency (events/cell/generation) ^b		β-Galactosidase activity (Miller units)			Methylation protection (% non-methylated) ^d	
	ON→OFF	OFF→ON	OFF	ON	(%ON) ^c	GATC-I	GATC-II
None (wild type)	2.35×10 ⁻²	5.54×10 ⁻⁴	26 ± 4	1440 ± 115	35	31	63
#13 (site 3)	locked ON	NA ^e	NA	5714 ± 899	100	95	0
#12 (site 2, GATC-II)	3.07×10 ⁻²	3.45×10 ⁻³	106 ± 42	1292 ± 190	31	28	NA
#24 (site 2)	locked ON	NA	NA	5282 ± 662	100	94	0
#11 (between sites 1 and 2)	1.23×10 ⁻²	4.24×10 ⁻³	369 ± 15	3272 ± 154	60	61	15
#10 (between sites 1 and 6)	NA	locked OFF	3 ± 1	NA		0	90
#57 (site 6)	2.96×10 ⁻²	1.90×10 ⁻⁴	20 ± 1	781 ± 40	9.8	ND ^f	ND
#9 (between sites 5 and 6)	2.28×10 ⁻²	3.27×10 ⁻⁴	30 ± 9	1703 ± 277	38	35	54
#14 (site 5, GATC-I)	NA	locked OFF	6 ± 1	NA		NA	80
#56 (site 4)	NA	locked OFF	1.9 ± 0.3	NA		ND	ND
#47 (sites 3 and 5) (mutations #13, #14)	locked ON	NA	NA	5053 ± 427	100	NA	0

^aThe mutation number is followed by the Lrp binding site (and GATC site where applicable) affected by the mutation (see Figure 2).

^bThe switch frequency is calculated as the number of switch events per cell per generation (Blyn *et al.*, 1990).

^cWhen β-galactosidase activities were measured, the percentage of cells in the ON phase was also determined, as described previously (Blyn *et al.*, 1990).

^dMethylation protection was calculated as the percentage of GATC sites that were digested by *Mbo*I (Braaten *et al.*, 1994).

^eNot applicable.

^fNot done.

Table II. Effects of Lrp and PapI on the phenotypes of *pap* regulatory mutants

Mutation in <i>pap</i> DNA	Switch frequency (events/cell/generation) ^a		β-Galactosidase activity (Miller units)		Methylation protection (% non-methylated) ^b	
	ON→OFF	OFF→ON	OFF	ON	GATC-I	GATC-II
In <i>Lrp</i> ⁻ background						
Wild type	NA ^c	locked OFF	8 ± 9	NA	0	0
#13 (site 3)	NA	locked OFF	6 ± 1	NA	0	0
#12 (site 2, GATC-II)	NA	locked OFF	9 ± 1	NA	0	0
In <i>PapI</i> ⁻ background						
Wild type	NA	locked OFF	3 ± 1	NA	0	70
#13 (site 3)	locked ON	NA	NA	4320 ± 420	0	0
#12 (site 2, GATC-II)	NA	locked OFF	41 ± 1	NA	0	NA

^aSwitch frequency is calculated as the number of switch events per cell per generation (Blyn *et al.*, 1990).

^bMethylation protection was calculated as the percentage of GATC sites that were digested by *Mbo*I (Braaten *et al.*, 1994).

^cNot applicable.

Table III. Effects of Dam methylase on the phenotypes of *pap* regulatory mutants

Mutation in <i>pap</i> DNA	Switch frequency (events/cell/generation) ^a		β-Galactosidase activity (Miller units)		Methylation protection (% non-methylated) ^b	
	ON→OFF	OFF→ON	OFF	ON	GATC-I	GATC-II
In <i>Dam</i> ⁻ strain						
Wild type	NA	locked OFF	84 ± 6	NA ^c	NT ^d	NT
#13 (site 3)	locked ON	NA	NA	4488 ± 411	NT	NT
#12 (site 2, GATC-II)	locked ON	NA	NA	2965 ± 435	NT	NT
#14 (site 5, GATC-I)	NA	locked OFF	7 ± 1	NA	NT	NT
In <i>Dam</i> overproducer strain (4-fold up)						
Wild type	NA	locked OFF	70 ± 34	NA	0	0
#13 (site 3)	locked ON	NA	NA	5779 ± 415	0	0
#12 (site 2, GATC-II)	NA	locked OFF	19 ± 1	NA	0	NA
#14 (site 5, GATC-I)	NA	locked OFF	69 ± 2	NA	NA	20

^aSwitch frequency is calculated as the number of switch events per cell per generation (Blyn *et al.*, 1990).

^bMethylation protection was calculated as the percentage of GATC sites that were digested by *Mbo*I (Braaten *et al.*, 1994).

^cNot applicable.

^dNot tested (previous analysis showed that the two *pap* GATC sites are 100% non-methylated in this *dam* background; unpublished data).

Fig. 5. Pap phase variation model. At the upper left, the six Lrp binding sites identified in this study are shown. Lrp binding sites [1,3,4,6] are shown as boxes, whereas Lrp sites [2,5], which contain the *pap* GATC-II and GATC-I sites, respectively, are shown as circles. Fully methylated GATC sites are shown as closed circles; hemi-methylated GATC sites are shown as half-closed circles. The -35 and -10 RNA polymerase binding sites of the *pap* BA promoter are shown as small shaded boxes flanking site 3. The circled numbers refer to steps in the *pap* transcription phase switching referred to in the text. Step 1: binding of Lrp to sites [1,2,3] blocks basal transcription and generates the phase OFF state. Step 2: PapI binds to Lrp. After DNA replication, a hemi-methylated GATC-II site is generated by DNA replication. Step 3: Lrp-PapI complexes bind to sites [4,5] and the GATC-II site is methylated by Dam, generating the phase ON state.

Lrp for sites [1,2,3] of *pap* ($K_d \sim 27$ nM; taken from the data in Figures 3 and 4). It appears that the activation of *ilvIH* transcription involves the occupancy of all five Lrp binding sites, a condition in which Lrp bound to site 5, 70 bp upstream of the transcription start site, might allow a direct contact with RNA polymerase. In contrast, our data for *pap* show that Lrp alternates between binding sites [1,2,3] in the absence of PapI, and sites [4,5] in the presence of PapI. Assuming that this *in vitro* analysis applies *in vivo*, in phase ON cells Lrp bound to *pap* sites [4,5] would be at least 140 bp upstream of the transcript start site, necessitating DNA bending to allow the contact between Lrp and RNA polymerase. It does not seem likely that site 6 is occupied in ON phase cells *in vivo* because PapI, which is required for *pap* transcription, reduces the affinity of Lrp for this site (Figure 4A).

The Lrp consensus binding sequence deduced from this study for *pap* [Gnn(n)TTTt] (Figure 2) differs somewhat from that reported previously for *ilvIH* (AgaATTT-TATtT), although both contain a 'T' tract. Our mutational analyses indicate that the consensus 'G' in sites [1,3,5] is important for Lrp binding (Figure 2B and Table I). However, sites outside this consensus also appear to be important for Lrp binding (see mutants #10 and #24). These DNA regions could function directly as contact sites for Lrp or indirectly through effects on the DNA structure.

Previous results have shown that the binding of Lrp to *ilvIH* DNA is highly cooperative (Wang and Calvo, 1993). Our data suggest that Lrp binds cooperatively to *pap* DNA sites [2,3]. Footprint analyses showed that Lrp bound to sites [1,2,3] with equal affinity (Figures 3 and 4). The disruption of site 3 inhibited the binding of Lrp to sites

[3,2] (mutant #13), as did the disruption of site 2 (mutant #12) (Figure 4). These results indicate that the binding of Lrp to sites [2,3] is cooperative. Based on preliminary data it does not seem likely that site 1 is required for Lrp cooperative interactions because the disruption of this site has only a small effect on Lrp affinity for sites [2,3] (L.Kaltenbach, unpublished data).

A model for Pap phase variation which includes the data presented here is shown in Figure 5. Lrp binds with highest affinity to sites [1,2,3], blocking basal transcription (step 1) (van der Woude *et al.*, 1995). The GATC-I sequence at site 5 is not bound by Lrp, and is methylated by Dam. Under conditions in which PapI levels are induced (Blyn *et al.*, 1990), PapI binds specifically to the Lrp moiety of Lrp-*pap* DNA complexes (step 2) (Kaltenbach *et al.*, 1995). However, Lrp-PapI complexes are locked in position because the methylation of GATC-I prevents the binding of Lrp/PapI to sites [4,5] (Nou *et al.*, 1993; Braaten *et al.*, 1994). We hypothesize that the Lrp/PapI binding transition shown in step 3 occurs following DNA replication because Lrp/PapI binds to the GATC-I region if this GATC site is hemi-methylated (Nou *et al.*, 1993). Following DNA replication, PapI decreases the affinity of Lrp for sites [1,2,3] and increases Lrp affinity for sites [4,5], resulting in the translocation of Lrp from the GATC-II to the GATC-I region (step 3). The methylation of GATC-II is required for *pap* transcription, possibly by inhibiting the binding of Lrp to the GATC-II region (Braaten *et al.*, 1994). The binding of Lrp to sites [4,5] activates *pap* transcription, possibly via the direct contact of Lrp with RNA polymerase (van der Woude *et al.*, 1995). The transition from the phase ON to phase OFF state can occur after DNA replication if Lrp dissociates

Table IV. *Escherichia coli* strains, plasmids and bacteriophage used in this study

<i>E.coli</i> strain, plasmid or bacteriophage	Description	Reference or source
<i>E.coli</i>		
MC4100	<i>F⁻araD139 Δ(lacIPOZYA-argF)U169 rpsL thi-1</i>	Casadaban (1976)
DL379	MC4100 λ246 lysogen	Blyn <i>et al.</i> (1989)
DL520	MC4100 λ246-16 lysogen	this study
DL524	MC4100 λ246-17 lysogen	this study
DL1504	MC4100 λ354 lysogen	Braaten <i>et al.</i> (1994)
DL2207	MC4100 λ354-7 lysogen	this study
DL2078	MC4100 λ354-9 lysogen	this study
DL2079	MC4100 λ354-10 lysogen	this study
DL2080	MC4100 λ354-11 lysogen	this study
DL2120	MC4100 λ354-12 lysogen	this study
DL2121	MC4100 λ354-13 lysogen	this study
DL2122	MC4100 λ354-14 lysogen	this study
DL2124	MC4100 λ354-24 lysogen	this study
DL2193	MC4100 λ354-47 lysogen	this study
DL2496	MC4100 λ354-56 lysogen	this study
DL2497	MC4100 λ354-57 lysogen	this study
DL1611	DL1504 <i>dam-13::Tn9</i> (cam ^r)	Braaten <i>et al.</i> (1994)
DL2150	DL2120 <i>dam-13::Tn9</i> (cam ^r)	this study
DL2151	DL2121 <i>dam-13::Tn9</i> (cam ^r)	this study
DL2152	DL2122 <i>dam-13::Tn9</i> (cam ^r)	this study
DL1650	DL1504(pPY1025, pTP166)	Braaten <i>et al.</i> (1994)
DL2156	DL2120(pPY1025, pTP166)	this study
DL2157	DL2121(pPY1025, pTP166)	this study
DL2158	DL2122(pPY1025, pTP166)	this study
DL2046	DL1504 <i>mbf-20</i> (<i>lrp</i> null mutation)	this study
DL2161	DL2120 <i>mbf-20</i>	this study
DL2162	DL2121 <i>mbf-20</i>	this study
DL1741	MC4100 λ354 Δ <i>papI</i> lysogen	this study
DL2187	MC4100 λ354-12 Δ <i>papI</i> lysogen	this study
DL2188	MC4100 λ354-13 Δ <i>papI</i> lysogen	this study
Plasmids		
pTZ18R	<i>amp</i> pMB1 replicon	US Biochemical
pDAL262	pTZ18R containing a 271 bp <i>TaqI</i> – <i>SphI</i> <i>papI</i> DNA fragment	Nou <i>et al.</i> (1993)
pRS550	<i>amp</i> – <i>kan</i> – <i>lacZYA</i> pMB1 replicon	Simons <i>et al.</i> (1987)
pRS551	<i>amp</i> – <i>kan</i> – <i>lacZYA</i> pMB1 replicon	Simons <i>et al.</i> (1987)
pDAL337	pTZ19U containing a 1.76 kb <i>papIB</i> regulatory sequence	Braaten <i>et al.</i> (1994)
pDAL354	pRS550 containing a 1.76 kb <i>papIB</i> regulatory sequence	Braaten <i>et al.</i> (1994)
pDAL354-7	pDAL354 with <i>pap</i> bp 1027 G to T transversion	this study
pDAL354-9	pDAL354 with <i>pap</i> bp 1048–1053 replaced by 'GCTAGC'	this study
pDAL354-10	pDAL354 with <i>pap</i> bp 1070–1075 replaced by 'GCTAGC'	this study
pDAL354-11	pDAL354 with <i>pap</i> bp 1101–1106 replaced by 'GCTAGC'	this study
pDAL354-12	pDAL354 with <i>pap</i> bp 1048–1053 replaced by 'GCTAGC'	this study
pDAL354-13	pDAL354 with <i>pap</i> bp 1161–1166 replaced by 'GCTAGC'	this study
pDAL354-24	pDAL354 with <i>pap</i> bp 1117–1121 replaced by 'CATAT'	this study
pDAL354-47	pDAL354 with mutations #13 and #14 in combination	this study
pDAL354-56	pDAL354 with <i>pap</i> bp 997–1009 replaced by 'TTGCTAGCTACTA'	this study
pDAL354-57	pDAL354 with <i>pap</i> bp 1059–1068 replaced by 'AGCTAGCAAT'	this study
pDAL246	pRS551 containing a 1.6 kb <i>papIB</i> DNA sequence	Blyn <i>et al.</i> (1990)
pDAL246-16	pDAL246 with <i>pap</i> bp 1163 C to A transversion	this study
pDAL246-17	pDAL246 with <i>pap</i> bp 1163 C to T transition	this study
pPY1025	pGB2 containing <i>lacI^q</i>	P.Youderian
pTP166	pBR322 containing <i>ptac</i> – <i>dam</i>	Marinus <i>et al.</i> (1984)
Bacteriophage		
λRS45	<i>amp'</i> – <i>lacZYA imm</i> ²¹	Simons <i>et al.</i> (1987)
λ354	λRS45–pDAL354 recombinant phage	Braaten <i>et al.</i> (1994)
λ354-7	λRS45–pDAL354-7 recombinant phage	this study
λ354-9	λRS45–pDAL354-9 recombinant phage	this study
λ354-10	λRS45–pDAL354-10 recombinant phage	this study
λ354-11	λRS45–pDAL354-11 recombinant phage	this study
λ354-12	λRS45–pDAL354-12 recombinant phage	this study
λ354-13	λRS45–pDAL354-13 recombinant phage	this study
λ354-14	λRS45–pDAL354-14 recombinant phage	this study
λ354-24	λRS45–pDAL354-24 recombinant phage	this study
λ354-47	λRS45–pDAL354-47 recombinant phage	this study
λ354-56	λRS45–pDAL354-56 recombinant phage	this study
λ354-57	λRS45–pDAL354-57 recombinant phage	this study
λ246-16	λRS45–pDAL246-16 recombinant phage	this study
λ246-17	λRS45–pDAL246-17 recombinant phage	this study

amp, ampicillin; kan, kanamycin.

from sites [4,5] and rebinds to sites [1,2,3]. This should be favored under low PapI levels (Blyn *et al.*, 1990).

Materials and methods

Bacterial strains, plasmids, bacteriophage and media

The *E. coli* strains, plasmids and bacteriophage used in this study are shown in Table IV. The construction of *E. coli* λ lysogens containing single-copy wild-type and mutant *pap* DNA sequences was carried out as described previously (Braaten *et al.*, 1994). Congenic *dam*⁻ and *lrp*⁻ strains were constructed by phage P1 transduction, and *Dam* overproducing strains were made as described previously (Braaten *et al.*, 1994). Luria-Bertani (LB) broth, LB agar, M9 minimal broth (M9) and M9 agar were prepared according to Miller (1972). The carbon source for M9 medium was glycerol at a final concentration of 0.2%. The chromogenic substrate 5-bromo-4-chloro-3-indolyl- β -D-galactoside (X-gal; Gold Biotechnology) was used at 40 μ g/ml of medium. Antibiotics, when used, were at the following final concentrations: ampicillin, 100 μ g/ml; kanamycin, 25 μ g/ml; spectinomycin, 100 μ g/ml; and tetracycline, 15 μ g/ml.

Site-specific mutagenesis and construction of double-site and *papI* deletion mutants

The site-directed mutagenesis protocol of Kunkel (1985) was used to mutate specific *pap* regulatory sequences as described previously, except that T7 DNA polymerase was used for second-strand synthesis (Braaten *et al.*, 1994). Mutagenic primers were made using an Applied Biosystems 380B DNA synthesizer as follows: mutant #7, 5'-TTCATTGACTATCTTTTATGC-3' OH; mutant #9, 5'-TTTATGCTGTAAATTGCTAGGCCATGATGTTTTTA-3' OH; mutant #10, 5'-CCATGATGTTTTTATGCTAGCAGCTCTTGCTATTA-3' OH; mutant #11, 5'-CTATTAGTGTTTTGTGCTAGCTAATTTGTTTTGTGG-3' OH; mutant #12, 5'-GTTTTGTGGGTTAAAGCTTTAAATCAATATT-3' OH; mutant #13, 5'-ATTTACAACATAAAGCTAGCATTTAACTTATTGCG-3' OH; mutant #14, 5'-AAATTCATTAGACGGCTAGCTATGCTGTAAATCA-3' OH; mutant #56, 5'-GCGTTTTATTTTCTGCTTGCTAGCTACGTAAAAATTCATTAGACG-3' OH; and mutant #57, 5'-GCTGTAAATTCATTTGCCATAGCTAGCAATTCTGAGTACCCTCTTGC-3' OH.

Mutations #16 and #17 were constructed by induction with hydroxylamine as follows. Phage λ 246 (Blyn *et al.*, 1990) was treated with hydroxylamine as described previously (Silhavy *et al.*, 1984) and used to lysogenize *E. coli* strain MC4100. Lysogens were plated on M9 minimal medium containing kanamycin and X-gal indicator at 23 and 37°C. Two isolates, DL520 and DL524, that showed locked ON phenotypes at both temperatures, were picked and analyzed. A DNA sequence analysis showed that DL520 contained a C to A transversion at 1163 bp (mutation #16), whereas strain DL524 contained a C to T transition at 1163 bp (mutation #17).

Mutant #24 was constructed by an overlapping PCR using two internal primers (primer A, 5'-TTTGTGTTTGTGCATATGGTTAAAGAT-3'; primer B, 5'-TCTTTTAACCATATGCACAAAACAAA-3') and two external primers (primer C, 5'-CGCATTTCTGACCGACT-3'; primer D, 5'-CCTGTTTCAGAAATCCAG-3'). Plasmid pDAL337, containing the entire *pap* regulatory region, was used as a template (Braaten *et al.*, 1994). An overlapping PCR was carried out as described previously (Ausubel *et al.*, 1994).

The double mutant #47, containing mutations #13 and #14, was constructed as follows. Both plasmids pDAL354-13 and pDAL354-14 were digested with *NspV* and *ApaI*. The *NspV*-*ApaI* DNA fragment was then digested with *DdeI* and the two DNA fragments containing mutations #13 and #14 were religated to the large DNA fragment of vector pDAL354, digested previously with *NspV* and *ApaI* (Braaten *et al.*, 1994).

The *papI* deletion derivatives λ 354 Δ *papI*, λ 354-12 Δ *papI* and λ 354-13 Δ *papI* were constructed by replacing the *SphI*-*NspV* DNA fragment containing the entire *papI* coding sequence with a 21mer double-stranded oligonucleotide containing the *SphI* and *NspV* end sites. This double-stranded oligonucleotide was made by annealing the following single-strand DNAs together: 5'-CTCACTGTAACAAAGTTTCTT-3' and 5'-CGAAGAACTTTGTTACAGTGAGCATG-3'.

Southern blot, switch frequency and β -galactosidase analyses

A Southern blot analysis was performed as described previously (Blyn *et al.*, 1990), except that chromosomal DNA was isolated using a

QIAGEN QIAamp tissue kit following the manufacturer's instructions. A quantitation of radioactivity was carried out using a Bio-Rad model GS-250 imager system. Switch frequencies (number of switch events per cell per generation) were determined as described previously (Blyn *et al.*, 1989), and β -galactosidase activities were determined as described previously (Braaten *et al.*, 1994).

DNA methylation protection, gel retardation and DNase I quantitative footprint analyses

A gel retardation analysis was carried out as described previously with the following modifications. Samples (20 μ l) contained 20 000 c.p.m. [³²P]dATP-labeled DNA fragment, 3 μ g herring sperm DNA and 2 μ g acetylated bovine serum albumin in binding buffer (60 mM Tris-HCl, pH 7.5, 40 mM KCl, 100 mM NaCl, 0.1 mM EDTA and 1 mM dithiothreitol). The reaction mixes were incubated 15 min at room temperature and analyzed on high ionic strength polyacrylamide gels, as described previously (Nou *et al.*, 1993). Similar results were obtained on low ionic strength gels, but with a reduced resolution.

DNA footprinting using DMS was performed by treating *pap* DNA-Lrp complexes, formed as described previously (Nou *et al.*, 1993), with 0.2 volumes of 3 M DMS in reaction buffer (0.3 M sodium cacodylate, pH 7.3, 1 mM EDTA) for 3 min at room temperature. The reaction was stopped by the addition of an equal volume (0.2 volumes of DNA-protein mix) of 2.5 M β -mercaptoethanol. DNA was precipitated by ethanol addition and washed twice with 70% ethanol before resuspension in 7 μ l 20 mM sodium acetate. Samples were then heated to 90°C for 30 min, and piperidine was added to 10% (v/v) in a final volume of 70 μ l. Samples were heated to 90°C for 30 min and vacuum evaporated. DNA samples were then resuspended in 5 μ l formamide loading buffer and analyzed on a 6% polyacrylamide sequencing gel.

Quantitative DNase I footprinting was carried out as described previously for DNase I footprinting (Braaten *et al.*, 1994) with the following modifications. Highly purified Lrp (Braaten *et al.*, 1994) was a gift from Felix Vajdos and Chris Hill (University of Utah, UT). PapI was isolated from a glutathione-S-transferase-PapI fusion protein, purified by affinity chromatography on glutathione agarose. After cleavage with thrombin, PapI containing two additional amino acids at the N-terminus (gly-ser) was isolated. A constant level of PapI (sufficient to supershift 100% of Lrp-DNA complexes) and increasing amounts of Lrp were added to 100 μ l of a mixture containing 5 ng ³²P end-labeled *pap* DNA probe (1 $\times$ 10⁵ c.p.m./ μ l), 2 μ g sonicated herring sperm DNA and 5 μ g acetylated bovine serum albumin in DNA binding buffer (40 mM Tris-HCl, pH 7.5, 60 mM KCl, 1 mM dithiothreitol, 0.1 mM EDTA). The binding was allowed to reach equilibrium by incubating for 20 min at room temperature. DNA-protein complexes were treated with 10 μ l 1 μ g/ml DNase I in 22 mM CaCl₂ and 22 mM MgCl₂ for 1 min. The reaction was stopped by the addition of 50 μ l stop buffer containing 50% glycerol and 60 mM EDTA. 10 μ l were then loaded on a native high ionic strength polyacrylamide gel to monitor the binding (Ausubel *et al.*, 1994). The remaining sample was precipitated by ethanol addition, resuspended in 10 μ l formamide loading buffer and analyzed on a 6% polyacrylamide sequencing gel. The quantitation of radioactivity was carried out using a Bio-Rad model GS-250 imager system.

Acknowledgements

We thank Marjan van der Woude and Barbara Graves for helpful discussions, and the Protein-DNA Core Facility, Cancer Center, University of Utah, UT, for the production of the oligonucleotides used in this study. The facility is supported by National Cancer Institute grant no. 5P30-CA42014-07. We are very grateful to Felix Vajdos and Chris Hill at the University of Utah, UT, for supplying highly purified Lrp, and to Brian Ernsting and Rowena Matthews for supplying the Lrp-overproducing strain used for purification. This work was supported by National Institutes of Health grant No. 2R01 AI23348 to D.A.L.

References

- Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1994) *Current Protocols in Molecular Biology*. John Wiley and Sons, New York, NY.
- Blyn, L.B., Braaten, B.A., White-Ziegler, C.A., Rolfson, D.H. and Low, D.A. (1989) Phase-variation of pylonophthrit-associated pili in *Escherichia coli*: evidence for transcriptional regulation. *EMBO J.*, **8**, 613-620.
- Blyn, L.B., Braaten, B.A. and Low, D.A. (1990) Regulation of *pap* pilin

- phase variation by a mechanism involving differential *dam* methylation states. *EMBO J.*, **9**, 4045–4054.
- Braaten, B.A., Blyn, L.B., Skinner, B.S. and Low, D.A. (1991) Evidence for a methylation-blocking factor (*mbf*) locus involved in *pap* pilus expression and phase variation in *Escherichia coli*. *J. Bacteriol.*, **173**, 1789–1800.
- Braaten, B.A., Platko, J.V., Woude, M.W.v.d., Simons, B.H., Graaf, F.K.d., Calvo, J.M. and Low, D.A. (1992) Leucine-responsive regulatory protein controls the expression of both the *pap* and *fan* pili operons in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **89**, 4250–4254.
- Braaten, B.A., Nou, X., Kaltenbach, L.S. and Low, D.A. (1994) Methylation patterns in *pap* regulatory DNA control the pyelonephritis-associated pili phase variation in *E. coli*. *Cell*, **76**, 577–588.
- Calvo, J.M. and Matthews, R.G. (1994) The leucine-responsive regulatory protein (Lrp), a global regulator of metabolism in *Escherichia coli*. *Microbiol. Rev.*, **58**, 466–490.
- Casadaban, M. (1976) Transposition and fusion of the *lac* genes to selected promoters in *E. coli* using bacteriophage lambda and Mu. *J. Mol. Biol.*, **104**, 541–555.
- Hochschild, A. (1991) Detecting cooperative protein–DNA interactions and DNA loop formation by footprinting. *Methods Enzymol.*, **208**, 343–364.
- Kaltenbach, L.S., Braaten, B.A. and Low, D.A. (1995) Specific binding of PapI to Lrp-*pap* DNA complexes. *J. Bacteriol.*, in press.
- Kunkel, T.A. (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl Acad. Sci. USA*, **82**, 488–492.
- Marinus, M.G., Poteete, A. and Arraj, J.A. (1984) Correlation of DNA adenine methylase activity with spontaneous mutability in *Escherichia coli* K-12. *Gene*, **28**, 123–125.
- Miller, J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Nou, X., Skinner, B., Braaten, B., Blyn, L., Hirsh, D. and Low, D. (1993) Regulation of pyelonephritis-associated pili phase variation in *Escherichia coli*: binding of the PapI and Lrp regulatory proteins is controlled by DNA methylation. *Mol. Microbiol.*, **7**, 545–553.
- Platko, J.V. and Calvo, J.M. (1993) Mutations affecting the ability of *Escherichia coli* Lrp to bind DNA, activate transcription, or respond to leucine. *J. Bacteriol.*, **175**, 1110–1117.
- Silhavy, T.J., Berman, M.L. and Enquist, L.W. (1984) *Experiments with Gene Fusions*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Simons, R.W., Houman, F. and Kleckner, N. (1987) Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. *Gene*, **53**, 85–96.
- van der Woude, M.W., Braaten, B.A. and Low, D.A. (1992) Evidence for global regulatory control of pilus expression in *Escherichia coli* by Lrp and DNA methylation: model building based on analysis of *pap*. *Mol. Microbiol.*, **6**, 2429–2435.
- van der Woude, M.W., Kaltenbach, L.S. and Low, D.A. (1995) Leucine-responsive regulatory protein plays dual roles as both an activator and a repressor of the *E. coli pap* operon. *Mol. Microbiol.*, **17**, 303–312.
- Wang, Q. and Calvo, J.M. (1993) Lrp, a global regulatory protein of *Escherichia coli*, binds co-operatively to multiple sites and activates transcription of *ilvIH*. *J. Mol. Biol.*, **229**, 306–318.
- Wissman, A. and Hillen, W. (1991) DNA contacts probed by modification protection and interference studies. *Methods Enzymol.*, **208**, 365–379.

Received on June 2, 1995; revised on September 7, 1995