



HAL
open science

A cryptic promoter in the LEE1 regulatory region of enterohaemorrhagic *Escherichia coli*: promoter specificity in AT-rich gene regulatory regions

Md. Shahidul Islam, Mark J Pallen, Stephen J.W. Busby

► To cite this version:

Md. Shahidul Islam, Mark J Pallen, Stephen J.W. Busby. A cryptic promoter in the LEE1 regulatory region of enterohaemorrhagic *Escherichia coli*: promoter specificity in AT-rich gene regulatory regions. *Biochemical Journal*, 2011, 436 (3), pp.681-686. 10.1042/BJ20110260 . hal-00596272

HAL Id: hal-00596272

<https://hal.science/hal-00596272>

Submitted on 27 May 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Biochemical Journal (BJ2011/0260)

Revised version

24th March 2011

A cryptic promoter in the *LEE1* regulatory region of enterohaemorrhagic *Escherichia coli*: promoter specificity in AT-rich gene regulatory regions

Md. Shahidul Islam, Mark J. Pallen and Stephen J.W. Busby*

School of Biosciences, University of Birmingham,
Edgbaston, Birmingham B15 2TT, UK

*To whom correspondence should be addressed, s.j.w.busby@bham.ac.uk
+44 (0)121 414 5439, no fax available

Keywords: bacterial promoters, transcription initiation, promoter discrimination, mutational analysis, bacterial multi-subunit RNA polymerase

Running title: Competing promoters in an *E. coli* regulatory region

ABSTRACT

Transcription of the *LEE1* operon in the locus of enterocyte effacement of enterohaemorrhagic *Escherichia coli* is due to the P1 promoter. Mutational and biochemical analyses reveal the existence of an overlapping promoter, designated P1A, which can drive transcript initiation 10 base pairs upstream from the P1 promoter transcript start point. Because of the overlap between P1 and P1A, P1A activity is unmasked only when the P1 promoter is inactivated by mutation. We report here that mutation of the P1 -10 element is less effective in unmasking P1A promoter activity than mutation of the P1 -35 element. This suggests that the P1 promoter -35 element, which corresponds to the consensus, can sequester RNA polymerase even when P1 is inactive and, thereby, prevent RNA polymerase from serving the P1A promoter. We propose that such promoter elements may play a role in enforcing specificity in bacterial regulatory regions that contain alternative possible promoters.

INTRODUCTION

Bacterial gene regulatory regions contain promoters that have evolved to set the level of expression of downstream genes, and the different promoter elements that are recognised by the multi-subunit bacterial RNA polymerase are now well understood. Exhaustive analyses of *Escherichia coli* promoters have shown that the principal promoter elements are the -35 and -10 hexamer elements, and consensus sequences of TTGACA and TATAAT respectively have been derived. Both elements are recognised directly by determinants in the RNA polymerase σ subunit and mutations in either element can lead to big reductions in promoter activity. The A at position 2 of the -10 element is especially important as it is implicated in opening of the transcription 'bubble' as well as promoter recognition. Thus, to a first approximation, the strength of any *E. coli* promoter will depend on the degree to which the base sequences of the functional -10 and -35 elements correspond to the consensus (reviewed in [1]).

Computational analysis of *E. coli* sequences has revealed the existence of potential -10 and -35 elements at higher than expected frequencies in intergenic regulatory regions [2]. Whilst this is probably a by-product of the evolution of gene regulatory regions, it raises the issue of the specificity of promoter recognition by RNA polymerase, and the role played by cryptic promoter elements. This is especially relevant for regulatory regions within horizontally-acquired pathogenicity islands that are generally characterised by a high proportion of AT base pairs [3] and hence contain many potential -10 elements.

In this work, we have focussed on a short promoter fragment, taken from an AT-rich pathogenicity island, the locus of enterocyte effacement (LEE) of enterohaemorrhagic *E. coli* (EHEC) serotype O157:H7. Recall that EHEC serotype O157:H7 causes haemorrhagic colitis and haemolytic uremic syndrome in humans, and the LEE is one of its key pathogenicity determinants [4]. This promoter fragment carries the principal promoter, P1, responsible for the expression of one of the operons encoded by the LEE, *LEE1*. In a recent study, we used mutational analysis to identify the *LEE1* P1 promoter -10 and -35 hexamer elements [5]. Here we report the existence of a second promoter that overlaps *LEE1* P1, whose activity is unmasked when the *LEE1* P1 promoter is inactivated by mutation. Surprisingly, we found that mutations in the *LEE1* P1 promoter -10 and -35 hexamer elements unmask this promoter to different extents. Hence, even when the *LEE1* P1 promoter is inactive due to a -10 element mutation, the activity of the alternative promoter is very low due to sequestration of RNA polymerase by the *LEE1* P1 promoter -35 element.

EXPERIMENTAL

Bacterial strains, plasmids, promoter fragments and primers

E. coli K-12 strain M182, which carries a deletion of the entire lactose operon [6], was used throughout this work, and was grown on MacConkey lactose indicator plates. The LEE20-275 EcoRI-HindIII fragment, which was constructed in our previous work [5], contains 120 base pairs of the EHEC O157:H7 *LEE1* operon regulatory region, including the *LEE1* P1 promoter, and is illustrated in Figure 1A. The vector plasmids used for cloning this fragment and derivatives with different mutations were pRW224, a *lac* expression vector encoding resistance to tetracycline [5,7], and pSR, a *colE1*-based

general cloning vector, encoding resistance to ampicillin [8]. Promoters that are cloned into pSR on EcoRI-HindIII fragments run into the bacteriophage λ *oop* terminator downstream of the HindIII site [9]. Mutations were introduced into the LEE20-275 fragments cloned in pRW224 by using error prone PCR [10] with flanking primers D10520 (5'-CCCTGCGGTGCCCTCAAC-3') and D53463 (5'-GGGGATGTGCTGCAAGGCG -3'), or by using megaprimer PCR [11] as described by Chismon *et al.* [12]. The different bases at the *LEE1* P1 promoter in the LEE20-275 fragment are numbered 1-120, and mutations are denoted by their position and the substituted base, as shown in Figure 1.

Promoter activity assays

E. coli M182 containing pRW224 carrying different mutated LEE20-275 fragments was inoculated from single colonies on MacConkey lactose indicator plates into LB medium supplemented with 35 $\mu\text{g ml}^{-1}$ tetracycline. Cultures were grown aerobically with shaking at 37°C, harvested during exponential growth, and β -galactosidase expression was measured by the Miller method [13]. Recorded activities are the average of at least three independent experiments and are taken as a measure of the promoter activity of the cloned LEE20-275 fragment.

In vitro experiments

The *in vitro* transcription experiments were performed as described by Browning *et al.* [9] using PstI-BamHI DNA fragments purified from caesium chloride preparations of pSR vector plasmid carrying the starting or mutated LEE20-275 fragment. These fragments served as a template for multiple round *in vitro* transcription assays in which 20 ng fragment was incubated in transcription buffer containing 40 mM Tris pH 7.9, 10 mM MgCl_2 , 1 mM DTT, 100 mM KCl, 100 $\mu\text{g ml}^{-1}$ bovine serum albumin, 200 μM GTP, 200 μM ATP, 200 μM CTP, 10 μM UTP and 5 μCi [α - ^{32}P]-UTP. Reactions were started by adding holo *E. coli* RNA polymerase, purchased from Epicenter (Madison). RNA products were analysed on a denaturing 5.5% polyacrylamide gel and visualised using a Fuji phosphor screen and BioRad Molecular Imager FX. Gels were calibrated with the pSR plasmid-encoded 108 base RNA-I transcript and Maxam-Gilbert "G+A" sequencing reactions. Permanganate footprints were performed using the PstI-BamHI DNA fragments labelled at the BamHI end using [γ - ^{32}P]-ATP and polynucleotide kinase. Experiments were performed following the protocol of Savery *et al.* [14] using holo *E. coli* RNA polymerase from Epicenter (Madison). Each reaction contained approximately 3 nM labelled DNA fragment in 20 mM HEPES, pH 8.0, 5 mM MgCl_2 , 50 mM potassium glutamate, 1 mM DTT and 0.5 mg ml^{-1} BSA and 50 nM holo RNA polymerase as required. Products of footprinting reactions were analysed on denaturing 6% polyacrylamide sequencing gels, calibrated with Maxam-Gilbert "G+A" sequencing reactions.

RESULTS AND DISCUSSION

Unexpected promoter activity due to the 64G and 98C substitutions in the LEE20-275 fragment

The LEE20-275 EcoRI-HindIII DNA fragment, illustrated in Figure 1A, covers the *LEE1* P1 promoter. The promoter activity of this fragment can be measured by using the pRW224 *lac* expression vector plasmid, and pRW224 carrying the LEE20-275 fragment confers a Lac⁺ phenotype on the *E. coli* K-12 Δ *lac* strain, M182. In recent work [5], error-prone PCR was exploited to generate a library of random mutations in this fragment cloned in pRW224. We then used MacConkey indicator plates to screen the Lac phenotype of colonies of M182 cells carrying the pRW224 library and selected for single base substitutions that reduced promoter activity and gave Lac⁻ colonies. Strikingly, over 90% of the selected base changes fell in two hexamer elements, TTGACA and TACACA, that were identified as the P1 promoter -35 and -10 hexamers respectively (Figure 1A). Assays with the different mutated LEE20-275 fragments cloned in pRW224 showed that single mutations in the P1 -10 element reduced measured *lac* expression to 3-6% of the level with the starting fragment, whilst mutations in the -35 element reduced expression to 12-27% of the starting level [5].

In complementary experiments, site-directed mutagenesis was used to construct specific mutations in the LEE20-275 fragment, including the 98C mutation at position 2 of the P1 -10 hexamer. Whilst screening Lac⁻ M182 colonies that carried pRW224 with the LEE20-275 98C mutant fragment, we unexpectedly found a single Lac⁺ colony. Sequence analysis showed that, in this colony, the LEE20-275

fragment contained a second accidentally generated substitution, 64G, located 34 base pairs upstream of the 98C base change (Figure 1B). Table 1 lists measurements of the promoter activity of this fragment, together with controls. As expected, the 98C substitution reduces expression to 4% of the starting level, whilst the measured promoter activity of the LEE20-275 fragment with both the 98C and 64G substitutions is 20-fold higher, which is nearly 80% of the starting level. Data in Table 1 also show that, in isolation, the 64G mutation causes but a modest 14% increase in promoter activity.

Characterisation of the *LEE1* P1A promoter

Since it is known that many *E. coli* promoters are completely inactivated by base changes at position 2 of the -10 hexamer [15-17], the simplest explanation for the effect of the 64G mutation is that it unmasks a cryptic promoter in the LEE20-275 fragment. To investigate this, transcripts formed, after labelled nucleoside triphosphates were added to binary complexes of purified RNA polymerase and the LEE20-275 fragment, were analysed. In this experiment, RNA polymerase runs to a downstream transcription terminator, and the RNA transcripts are labelled by using [α^{32} P]-labelled UTP and analysed by gel electrophoresis (Figure 2A). With the starting LEE20-275 fragment, the major transcript is ~102 bases, which corresponds to the *LEE1* P1 transcript that starts at position 107A (Figure 1A). As expected, this transcript is completely absent when the DNA fragment carries the 98C substitution. However, with the LEE20-275 fragment carrying both the 98C and 64G substitutions, a new ~112 base transcript is observed. This corresponds to a transcript starting at position 97T, suggesting that the 64G substitution has unmasked a promoter (Figure 1B). To confirm this, we used potassium permanganate footprinting to compare regions of DNA duplex unwinding in binary complexes of purified RNA polymerase and the LEE20-275 fragment either without or with the 98C and 64G substitutions. Recall that potassium permanganate modifies T residues in the single stranded 'bubble' produced after local unwinding of promoter DNA around the transcription start at promoters [14]. Results in Figure 2B show that, without the substitutions, there is extensive unwinding that starts just downstream of the P1 promoter -10 hexamer, and that this unwinding is suppressed by the 98C substitution. In contrast, with the fragment carrying both the 98C and 64G substitutions, clear duplex opening is seen at positions 86, 88, 89 and 93, just upstream from the transcript start at 97T. This argues that 98C and 64G substitutions cause RNA polymerase to recognise a new promoter (P1A), with a transcript start that is 10 base pairs upstream from the P1 promoter start (Figure 1B).

In order to identify sequence elements essential for the P1A promoter, error prone PCR was used to generate 3 independent preparations of the LEE20-275 fragment with the 98C and 64G substitutions. They were cloned into pRW224, the mixture of resulting recombinant plasmids was transformed into *E. coli* strain M182, and transformants were grown on MacConkey indicator plates. As expected, the majority of colonies scored as Lac⁺, but, after screening over 15,000 transformants, we identified 4 Lac⁻ colonies that each carried a supplementary base change. These were an A for C change at position 88 (88C), and A for G substitution at position 89 (89G), a T for C substitution at position 90 (90C), and the insertion of a single T between positions 71 and 72 (Figure 1B). Table 2 shows measurements of the promoter activity of each of the four mutant fragments. The data show that the 88C substitution reduces promoter activity by over 20-fold, whilst the other three substitutions have smaller effects. These data, together with the *in vitro* results in Figure 2, identify the hexamer, TAATGT, from position 87 to position 92 as the P1A -10 hexamer, upstream of the transcript start at position 97. Hence, the 88C mutation at position 2 of the hexamer has the biggest effects on P1A activity and is within the region of duplex unwinding following RNA. We tentatively assign TGGTTT from position 62 to position 67 as the P1A -35 hexamer (Figure 1B). Although this corresponds to the consensus TTGACA at only 2 out of 6 positions, the concordant G at position 64 is the G that was created by the 64G substitution that unmasked the P1A promoter. Thus, the high promoter activity of the LEE20-275 fragment carrying the 98C and 64G substitutions is likely due to improvement of the -35 element from TGATTT to TGGTTT.

Activity of *LEE1* P1A without the 64G substitution

The chance isolation of the 64G substitution in the LEE20-275 fragment, in combination with the 98C substitution, led us to identify the P1A promoter. We wanted to understand whether P1A had been generated by the 64G mutation, or whether it is a cryptic promoter, present in the starting LEE20-275 fragment, but silenced because of competition by the P1 promoter. To investigate this, we exploited the

88C substitution at position 2 of the P1A promoter -10 hexamer element that inactivates P1A. Table 3 illustrates an experiment designed to quantify the activity of P1A in the LEE20-275 fragment and in different mutated derivatives, by measuring the effect of the 88C substitution on β -galactosidase expression after the fragments were cloned into pRW224.

Results in Table 3 show that the 88C substitution has very little effect on the measured promoter activity of the starting LEE20-275 fragment. However, when the LEE20-275 fragment carries mutations in the P1 -10 element (98C), or in the P1 -35 element (76G), or in both elements of P1 (76G 98G), the 88C substitution reduces expression by over 85%. The simplest explanation for these observations is that P1A is a cryptic promoter that is silenced by competition from P1, but becomes active as P1 is inactivated.

Surprisingly, the introduction of the 76G substitution in the P1 -35 element substantially increased the activity of the LEE20-275 fragment carrying the 98G substitution in the P1 -10 element. This increased expression is completely suppressed by the 88C substitution, and, hence, must be due to the P1A promoter. From this, we conclude that, even though the P1 promoter is inactivated by the 98G substitution, the activity of P1A remains low because RNA polymerase still makes abortive contacts with the P1 consensus TTGACA -35 hexamer element. Hence, P1A activity increases as these interactions are weakened, due to substitutions such as 76G. In a related experiment, one of the 7 consecutive T residues immediately upstream of the P1 promoter -35 hexamer element (Δ 71T) was deleted, in order to change the length of the spacer between the P1A promoter -10 and -35 hexamer elements closer to the optimal 17 base pairs [15]. Results in Table 3 show that the promoter activity of the resulting LEE20-275 fragment carrying the Δ 71T change is reduced ~3-fold by the 88C substitution. Hence strengthening P1A in the LEE20-275 fragment increases its contribution to measured promoter activity.

Conclusions

Many bacterial gene regulatory regions, especially those that are AT-rich, contain multiple promoters, and it is generally assumed that they are a by-product of evolution. Here we have identified a cryptic promoter, P1A, which overlaps the *LEE1* P1 promoter of EHEC serotype O157:H7. Our results suggest that the P1A -10 hexamer element lies between the P1 -10 and -35 hexamer elements, and the P1A -35 hexamer element lies upstream of the P1 -35 element (Figure 1). The P1A promoter was discovered following the chance isolation of the 64G substitution that improves its -35 element. Our data argue that the P1A promoter can function even without the 64G substitution and thus it is a true cryptic promoter. The P1A promoter overlaps with P1 and hence the two promoters compete and are mutually exclusive. In our conditions, with the starting LEE20-275 fragment, the P1 promoter clearly wins the competition. However, as P1 is weakened, or P1A is strengthened, the scale tips towards P1A. Interestingly, when P1 is inactivated by a substitution at position 2 of its -10 hexamer, measured P1A activity remains low due to the P1 -35 element, which corresponds exactly to the consensus. Thus the P1 promoter, even when inactive, can retain the ability to sequester RNA polymerase, thereby blocking access to the P1A promoter.

Potentially, clusters of potential promoter elements in regulatory regions can have many different consequences. For example, they may, quite simply, give rise to multiple transcription starts and recent high resolution analysis of the *E. coli* transcriptome has identified many cases of this [18]. This can lead to complex patterns of regulation. Alternatively, these elements could be the remains of now defunct regulatory systems, or be awaiting future adoption. Our finding that the P1 -35 element retains function even when P1 is inactive, together with similar observations with the *E. coli gal* operon regulatory region [19] underscores that residual promoter elements at non-functional promoters can affect the distribution of RNA polymerase. At present, the role of the *LEE1* P1A is unclear. However, we note that, although consensus -35 hexamer elements are rare at *E. coli* promoters [20], the *LEE1* P1 -35 element is conserved in all of the pathogenic *E. coli* listed in the xBASE database that contain a locus of enterocyte effacement [21], as well as in *Citrobacter rodentium* [22]. We speculate that the rationale for retaining TTGACA as the *LEE1* P1 -35 element is to focus RNA polymerase at P1 and to reduce the use of neighbouring promoters such as P1A.

ACKNOWLEDGEMENTS

This work was supported by a BBSRC project grant and by a Commonwealth PhD studentship to M.S.I. We are grateful to Kerry Hollands for help with the footprinting and transcription experiments.

REFERENCES

1. Busby, S., Kolb, A. and Buc, H. (2009) Where it all begins: an overview of promoter recognition and open complex formation. In *RNA Polymerases as Molecular Motors* (Buc, H. & Strick, T., eds) chapter 1, pp. 13-17, RSC Publishing, Cambridge
2. Huerta, A.M. and Collado-Vides, J. (2003) Sigma70 promoters in *Escherichia coli*: specific transcription in dense regions of overlapping promoter-like signals. *J. Mol. Biol.* **333**, 261-278
3. Hacker, J., Blum-Oehler, G., Mühldorfer, I. and Tschäpe, H. (1997) Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Mol. Microbiol.* **23**, 1089-1097
4. Elliott, S.J., Wainwright, L.A., McDaniel, T.K., Jarvis, K.G., Deng, Y.K., Lai, L.C., McNamara, B.P., Donnenberg, M.S. and Kaper, J.B. (1998) The complete sequence of the locus of enterocyte effacement (LEE) from enteropathogenic *Escherichia coli* E2348/69. *Mol. Microbiol.* **28**, 1-4
5. Islam, M.S., Bingle, L.E.H., Pallen, M.J. and Busby, S.J.W. (2010) Organization of the *LEE1* operon regulatory region of enterohaemorrhagic *Escherichia coli* O157:H7 and activation by GrlA. *Mol. Microbiol.* **79**, 468-483
6. Casadaban, M.J. (1976) Transposition and fusion of the *lac* genes to selected promoters in *Escherichia coli* using bacteriophage lambda and Mu. *J. Mol. Biol.* **5**, 541-555
7. Lodge, J., Fear, J., Busby, S., Gunasekaran, P. and Kamini, N-R. (1992) Broad host range plasmids carrying the *Escherichia coli* lactose and galactose operons. *FEMS Microbiol. Lett.* **95**, 271-276
8. Kolb, A., Kotlarz, D., Kusano, S. and Ishihama, A. (1995) Selectivity of the *Escherichia coli* RNA polymerase E sigma 38 for overlapping promoters and ability to support CRP activation. *Nucleic Acids Res.* **23**, 819-826
9. Browning, D., Savery, N., Kolb, A. and Busby, S. (2009) Assays for transcription factor activity. In *Methods in Molecular Biology: DNA-Protein Interactions* (Moss, T., and Leblanc, B. eds). 3rd Edition, volume 543, pp. 369-387, Springer Science, Heidelberg
10. Leung, D.W., Chen, E. and Goeddel, D.V. (1989) A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. *Technique* **1**, 11-15
11. Sarkar, G. and Sommer S.S. (1990) The "megaprimer" method of site directed mutagenesis. *Biotechniques* **8**, 404-407
12. Chismon, D.L., Browning, D.F., Farrant, G.K. and Busby, S.J. (2010) Unusual organisation, complexity and redundancy at the *Escherichia coli hcp-hcr* operon promoter. *Biochem. J.* **430**, 61-68
13. Miller, J. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Plainview, NY
14. Savery, N.J., Belyaeva, T. and Busby, S. (1996) Introduction to protein: DNA interactions, DNase I footprinting, hydroxyl radical footprinting, permanganate footprinting and supplementary protocols. In *Essential Techniques: Gene Transcription*. (Docherty, K., ed.) pp. 1-5 and 21-33, BIOS Scientific Publishers, Oxford
15. Rosenberg, M. and Court, D. (1979) Regulatory sequences involved in the promotion and termination of RNA transcription. *Annu. Rev. Genet.* **13**, 319-353
16. Miroslavova, N.S. and Busby, S.J.W. (2006) Investigations of the modular structure of bacterial promoters. *Biochem. Soc. Symp.* **73**, 1-10

17. Jayaraman, P-S., Cole, J. A. and Busby, S.J. (1989) Mutational analysis of nucleotide sequence at the FNR-dependent *nirB* promoter in *Escherichia coli*. *Nucleic Acids Res.* **17**, 135-145
18. Mendoza-Vargas, A., Olvera, L., Olvera, M., Grande, R., Vega-Alvarado, L., Taboada, B., Jimenez-Jacinto, V., Salgado, H., Juarez, K., Contreras-Moreira, B., Huerta, A.M., Collado-Vides, J. and Morett, E. (2009) Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in *E. coli*. *PLoS ONE* **4**, e7526
19. Johnston, F., Ponnambalam, S. and Busby, S. (1987) Binding of *Escherichia coli* RNA polymerase to a promoter carrying mutations that stop transcription initiation. *J. Mol. Biol.* **195**, 745-748
20. Mitchell, J.E., Zheng, D., Busby, S.J., and Minchin, S.D. (2003) Identification and analysis of 'extended -10' promoters in *Escherichia coli*. *Nucleic Acids Res.* **31**, 4689-4695
21. Chaudhuri, R.R., Loman, N.J., Snyder, L.A.S., Bailey, C.M., Stekel, D.J. and Pallen, M.J. (2008) xBASE2: a comprehensive resource for comparative bacterial genomics. *Nucleic Acids Res.* **36**, D543-546
22. Schauer, D.B. and Falkow, S. (1993) Attaching and effacing locus of *Citrobacter freundii* biotype that causes transmissible murine colonic hyperplasia. *Infect. Immun.* **61**, 2486-2492

Table 1. Mutational analysis of the LEE20-275 promoter fragment

Promoter fragments	Promoter sequences from positions 62 to 102	β -galactosidase activity (Miller units \pm SD)
	62 70 80 90 102 ● ● ● ● ●	
LEE20-275 (WT)	5'-TGATTTTTTTGTTGACATTTAATGATAATGTATTTTACACA-3'	5513 \pm 261
LEE20-275 98C	5'-TGATTTTTTTGTTGACATTTAATGATAATGTATTTT CC CACA-3'	233 \pm 4 (4.2)
LEE20-275 64G 98C	5'-TG G TTTTTTTTGTTGACATTTAATGATAATGTATTTT CC CACA-3'	4375 \pm 133 (79.4)
LEE20-275 64G	5'-TG G TTTTTTTTGTTGACATTTAATGATAATGTATTTTACACA-3'	6276 \pm 20 (114)

The table shows measured β -galactosidase activities in cultures of *E. coli* strain M182 carrying pRW224 containing the LEE20-275 (WT) fragment and different derivatives. Cultures were grown aerobically at 37°C in LB medium to an optical density of \sim 0.5 at 650 nm. Activities were measured in triplicate, giving a mean and standard deviation (SD). Activities expressed as a percentage of activity with the starting LEE20-275 fragment are shown in parentheses. The central part of the table shows the fragment base sequence from position 62 to position 102, with the P1 promoter -10 and -35 hexamer elements grey shaded. Base substitutions in the different fragments are underlined and highlighted in boldface type.

Table 2. Identification of the functional elements of the P1A promoter

Promoter fragments	Promoter sequences from positions 62 to 102	β -galactosidase activity (Miller units \pm SD)
	62 70 80 90 102 ● ● ● ● ●	
LEE20-275 64G 98C	5'- <u>TGGTTTTTTT</u> GTTGACATTTAATGATAATGTATTTCCACA-3'	4375 \pm 133
LEE20-275 64G 98C 88C	5'- <u>TGGTTTTTTT</u> GTTGACATTTAATGAT CAT GTATTTCCACA-3'	190 \pm 45 (4.3)
LEE20-275 64G 98C 89G	5'- <u>TGGTTTTTTT</u> GTTGACATTTAATGATAG T GTATTTCCACA-3'	517 \pm 28 (11.8)
LEE20-275 64G 98C 90C	5'- <u>TGGTTTTTTT</u> GTTGACATTTAATGATA ACG TATTTCCACA-3'	263 \pm 15 (6.0)
LEE20-275 64G98C	5'- <u>TGGTTTTTTT</u> GTTGACATTTAATGATA ACG TATTTCCACA-3'	939 \pm 8 (21.5)
InsT (71-72)		

The table shows measured β -galactosidase activities in cultures of *E. coli* strain M182 carrying pRW224 containing the LEE20-275 64G 98C fragment and different mutations. Cultures were grown aerobically at 37°C in LB medium to an optical density of \sim 0.5 at 650 nm. Activities were measured in triplicate, giving a mean and standard deviation (SD). Activities expressed as a percentage of activity with the starting LEE20-275 64G 98C fragment are shown in parentheses. The central part of the table shows the fragment base sequence from position 62 to position 102, with the location of P1 promoter -10 and -35 hexamer elements grey shaded and predicted cryptic promoter -10 and -35 elements are underlined. Base substitutions and insertion in the different fragments are highlighted in boldface type.

Table 3. P1A promoter activity quantified by effects of the 88C substitution

Promoter fragments	Promoter sequences from positions 62 to 102	β -galactosidase activity (Miller units \pm SD)
	62 70 80 90 102 ● ● ● ● ●	
LEE20-275 (WT)	5'- <u>TGATTTTTTT</u> GTTGACATTTAATGATAATGTATTTTACACA-3'	5513 \pm 261
LEE20-275 88C	5'- <u>TGATTTTTTT</u> GTTGACATTTAATGAT CAT GTATTTTACACA-3'	4596 \pm 30
LEE20-275 98C	5'- <u>TGATTTTTTT</u> GTTGACATTTAATGATAATGTATTTT CC CACA-3'	233 \pm 4
LEE20-275 98C 88C	5'- <u>TGATTTTTTT</u> GTTGACATTTAATGAT CAT GTATTTT CC CACA-3'	10 \pm 1
LEE20-275 76G	5'- <u>TGATTTTTTT</u> GTT GGC ATTTAATGATAATGTATTTTACACA-3'	1568 \pm 24
LEE20-275 76G 88C	5'- <u>TGATTTTTTT</u> GTT GGC ATTTAATGAT CAT GTATTTTACACA-3'	220 \pm 13
LEE20-275 98G	5'- <u>TGATTTTTTT</u> GTT GGC ATTTAATGATAATGTATTTT GC CACA-3'	202 \pm 15
LEE20-275 76G 98G	5'- <u>TGATTTTTTT</u> GTT GGC ATTTAATGATAATGTATTTT GC CACA-3'	1289 \pm 3
LEE20-275 76G 98G 88C	5'- <u>TGATTTTTTT</u> GTT GGC ATTTAATGAT CAT GTATTTT GC CACA-3'	11 \pm 1
LEE20-275 Δ 71	5'- <u>TGATTTTTTT</u> -GTTGGCATTTAATGATAATGTATTTTCCACA-3'	3107 \pm 115
LEE20-275 Δ 71 88C	5'- <u>TGATTTTTTT</u> -GTTGGCATTTAATGAT CAT GTATTTTCCACA-3'	1103 \pm 24

The table shows measured β -galactosidase activities in cultures of *E. coli* strain M182 carrying pRW224 containing the LEE20-275 fragment and different mutations. Cultures were grown aerobically at 37°C in LB medium to an optical density of \sim 0.5 at 650 nm. Activities were measured in triplicate, giving a mean and standard deviation (SD). The central part of the table shows the fragment base sequence from position 62 to position 102, with the P1 promoter -10 and -35 hexamer elements grey shaded and P1A promoter -10 and -35 elements are underlined. Base substitutions in the different fragments are highlighted in boldface type whereas the Δ 71 deletion is shown by a dash. For each set of promoter fragments, the effect of the 88C mutation gives an estimate of P1A activity.

FIGURE LEGENDS

Figure 1. Nucleotide sequence and map of the LEE20-275 *LEE1* regulatory region fragment

A. The figure shows the base sequence of the LEE20-275 fragment, carrying the *LEE1* P1 promoter, flanked by upstream EcoRI and downstream HindIII restriction sites. The sequence is numbered 1-120, starting with the first base of the cloned *LEE1* regulatory region. The P1 promoter -35 and -10 hexamer elements are boxed and the transcript start site is indicated by a bent arrow. Mutations in the -35 and -10 promoter elements that cause reduction of P1 activity are indicated by vertical arrows.

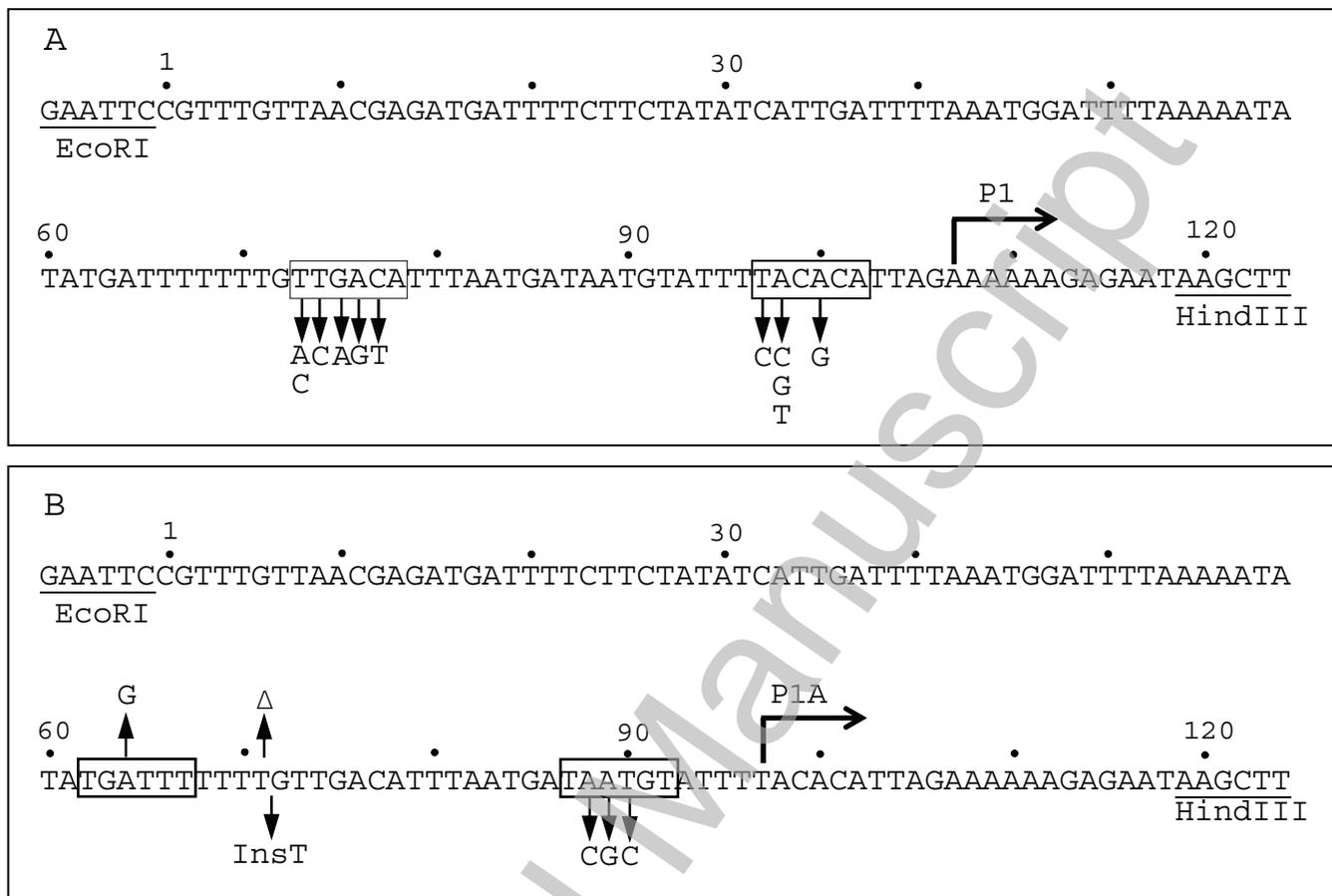
B. The figure shows the base sequence of the LEE20-275 fragment highlighting the sequence determinants of the P1A promoter. The P1A promoter -35 and -10 elements are boxed and the transcript start site is indicated by a bent arrow. The locations of the 88C, 89G, 90C and InsT (71-72) mutations that reduce P1A activity are indicated, together with the 64G and Δ 71T 'up' mutations.

Figure 2. *In vitro* run-off transcription and potassium permanganate footprinting analyses

A. The figure shows an autoradiogram of an analysis by gel electrophoresis of 32 P-labelled RNA transcripts made by RNA polymerase holoenzyme from DNA fragments carrying the starting LEE20-275 sequence (lanes 1-3), a derivative with the 98C mutation (lanes 4-6), or a derivative with the 64G and 98C mutations (lanes 7-9). The RNA polymerase concentration was: lanes 1, 4 and 7, no enzyme; lanes 2, 5 and 8, 200 nM; lanes 3, 6 and 9, 400 nM. The gel was calibrated with the pSR plasmid-encoded 108-base RNA-I transcript (lane 10) and Maxam-Gilbert sequence reactions (GA). The P1 and P1A transcripts are indicated by asterisks, and the proposed corresponding initiation sites are shown by bent arrows on the starting (WT) and mutant (64G 98C) sequences.

B. The figure shows an autoradiogram that identifies the potassium permanganate sensitive sites in complexes of holo RNA polymerase with a DNA fragment carrying the starting LEE20-275 sequence (WT), a derivative with the 98C mutation, or a derivative with the 64G and 98C mutations. Lanes 1, 3 and 5 show the results after control incubations without RNA polymerase, while lanes 2, 4 and 6 show the analysis of samples with 50 nM RNA polymerase. The gel was calibrated using Maxam-Gilbert sequence reactions (GA) and relevant positions are indicated. Asterisks identify residues that display RNA polymerase-dependent reactivity to permanganate. Partial sequences of the WT and 64G 98C fragments are shown on the left and right sites respectively, with reactive sites marked by asterisks.

Islam *et al.* Figure 1

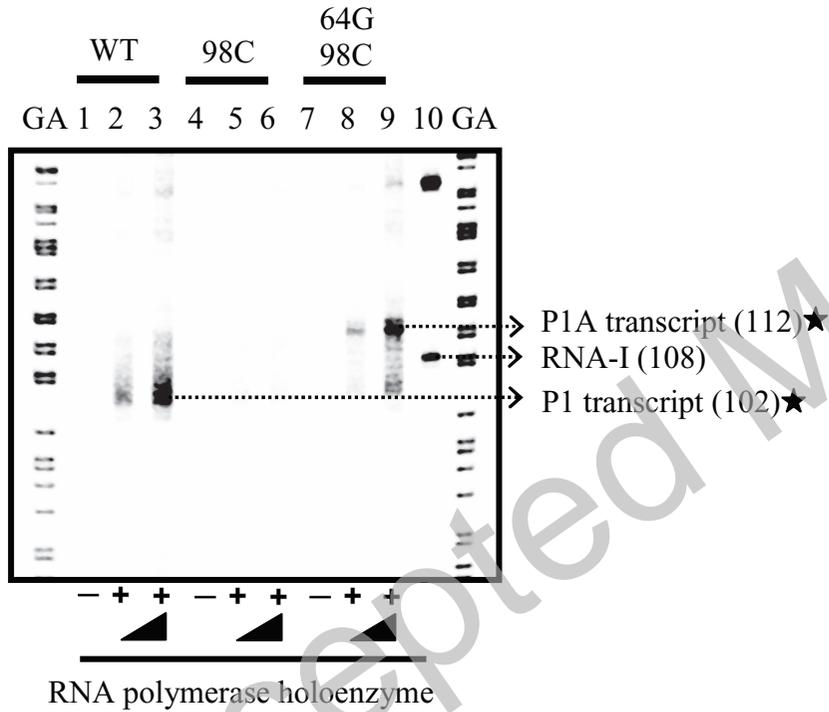


THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BJ20110260

Accepted Manuscript

Islam et al. Figure 2

A: *In vitro* transcription assay



B: *In vitro* permanganate footprinting

