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, U.K.

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 recognized 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 recognized directly by determinants in the RNA polymerase σ subunit bacterial RNA polymerase are now well understood. In a recent study, we used mutational analysis to identify the LEE1 P1 promoter −10 and −35 hexamer elements [5]. In the present paper, 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 the present study, and was grown on MacConkey lactose indicator plates. The LEE20–275 EcoRI/HindIII fragment, which was constructed in our previous work [5], contains 120 bp of the EHEC O157:H7 LEE1 operon regulatory region, including the LEE1 P1 promoter, and is illustrated in Figure 1(A). 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′-CCCTGCGGTGCCCCTCAAC-3′) and D5346 (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.

Abbreviations used: DTT, dithiothreitol; EHEC, enterohaemorrhagic Escherichia coli; LB, Luria–Bertani; LEE, locus of enterocyte effacement. 1 To whom correspondence should be addressed (email s.j.w.busby@bham.ac.uk).
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 (Luria–Bertani) medium supplemented with 35 μg·ml⁻¹ tetracycline. Cultures were grown aerobically with shaking at 37°C, harvested during exponential growth, and β-galactosidase expression was measured using the Miller method [13]. Activities recorded 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 of fragment was incubated in transcription buffer containing 40 mM Tris/HCl (pH 7.9), 10 mM MgCl₂, 1 mM DTT (dithiothreitol), 100 mM KCl, 100 μg·ml⁻¹ BSA, 200 μM GTP, 200 μM ATP, 200 μM CTP, 10 μM UTP and 5 μCi of [α-³²P]UTP. Reactions were started by adding holo E. coli RNA polymerase, purchased from Epicenter. RNA products were analysed on a denaturing 5.5% polyacrylamide gel and visualized using a Fuji phosphor screen as i n g l e L a c expression to 3–6% of the starting level [5]. Assays with the different mutated LEE20–275 fragments cloned into pRW224 showed that single mutations in the P1 promoter reduced measured lac expression 3–6% of the level with the starting fragment, whereas 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 promoter. While 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 bp upstream of the 98C base change (Figure 1B). Figure 2 lists measurements of the promoter activity of this fragment, together with controls. As expected, the 98C substitution reduces expression to 4% of the

Figure 1  Nucleotide sequence and map of the LEE20–275 LEE1 regulatory region fragment

(A) 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 fragment. The P1 promoter −35 and −10 hexamer elements are boxed, and the transcription 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 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 transcription 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 Δ71T ‘up’ mutations.

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 1(A), 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 Δlac strain M182. In recent work [5], error-prone PCR was used 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 −35 and −10 hexamer elements, respectively (Figure 1A).

Assays with the different mutated LEE20–275 fragments cloned into pRW224 showed that single mutations in the P1 −35 element reduced lac expression to 3–6% of the level with the starting fragment, whereas mutations in the −35 element reduced expression to 12–27% of the starting level [5].
starting level, whereas 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. The results in Figure 2 also show that, in isolation, the 64G mutation causes only a modest (14%) increase in promoter activity.

**Characterization 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 NTPs 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 using [α-32P]UTP and analysed by gel electrophoresis (Figure 3A). 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 greatly reduced 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. 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 3(B) 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 of the transcript start at 97T. This argues that 98C and 64G substitutions cause RNA polymerase to recognize a new P1A promoter, with a transcript start that is 10 bp upstream of the P1 promoter start (Figure 1B).

To identify sequence elements essential for the P1A promoter, error-prone PCR was used to generate three independent preparations of the LEE20–275 fragment with the 98C and 64G substitutions. They were cloned into pRW24, 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 15000 transformants, we identified four Lac− colonies that each carried a supplementary base change. These were an A for C substitution at position 88 (88C), an 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). Figure 4 shows measurements of the promoter activity of each of the four mutant fragments. The results show that the 88C substitution reduces promoter activity over 20-fold, whereas the other three substitutions have smaller effects. These results, together with the in vitro results in Figure 3, identify the hexamer TAATGT, from position 87 to position 92, as the LEE1 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 TAATGT, from position 87 to position 92, as the LEE1 P1A promoter. Thus the high promoter activity of the LEE20–275 fragment carrying the 98C and 64G substitutions is probably due to improvement of the −35 element from TGATTTT to TGTTTT.

**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. Figure 5 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 Figure 5 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
Figure 3  \textit{In vitro} run-off transcription and potassium permanganate footprinting analyses

(A) 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; and 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) 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, whereas 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 potassium permanganate. Partial sequences of the WT and 64G 98C fragments are shown on the left-hand and right-hand sites respectively, with reactive sites marked by asterisks.

<table>
<thead>
<tr>
<th>Promoter fragments</th>
<th>Promoter sequences from positions 62 to 102</th>
<th>$\beta$-galactosidase activity (Miller units ± SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEE20-275 64G 98C</td>
<td>5'-TGTTTTTTTGTACATTTATGATATTTCCACA-3'</td>
<td>4375 ± 133</td>
</tr>
<tr>
<td>LEE20-275 64G 98C 88C</td>
<td>5'-TGTTTTTTTGTACATTTATGATATTTCCACA-3'</td>
<td>190 ± 45 (4.3)</td>
</tr>
<tr>
<td>LEE20-275 64G 98C 89G</td>
<td>5'-TGTTTTTTTGTACATTTATGATATTTCCACA-3'</td>
<td>517 ± 28 (11.8)</td>
</tr>
<tr>
<td>LEE20-275 64G 98C 90C</td>
<td>5'-TGTTTTTTTGTACATTTATGATATTTCCACA-3'</td>
<td>263 ± 15 (6.0)</td>
</tr>
<tr>
<td>InstT (71-72)</td>
<td>5'-TGTTTTTTTGTACATTTATGATATTTCCACA-3'</td>
<td>939 ± 8 (21.5)</td>
</tr>
</tbody>
</table>

Figure 4  Identification of the functional elements of the P1A promoter

Measured $\beta$-galactosidase activities in cultures of \textit{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 absorbance of ~0.5 at 650 nm. Activities were measured in triplicate, giving a mean ± S.D. Activities expressed as a percentage of activity with the starting LEE20–275 64G 98C fragment are shown in parentheses. The central part of the Figure shows the fragment base sequence from position 62 to position 102, with the location of P1 promoter −10 and −35 hexamer elements shaded in grey and predicted cryptic promoter −10 and −35 elements underlined. Base substitutions and insertion in the different fragments are highlighted in bold.

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 in 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 seven consecutive T residues immediately upstream of the P1 promoter −35 hexamer element (Δ71T) 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 bp [15]. Results in Figure 5 show that the promoter activity of the resulting LEE20–275 fragment carrying the Δ71T change is reduced ~3-fold by the 88C substitution. Hence strengthening P1A in the LEE20–275 fragment increases its contribution to the promoter activity measured.
Competing promoters in an E. coli regulatory region

Figure 5  P1A promoter activity quantified by effects of the 88C substitution

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 attenuance of ∼0.5 at 650 nm. Activities were measured in triplicate, giving a mean ± S.D. The central part of the Figure shows the fragment base sequence from position 62 to position 102, with the P1 promoter −10 and −35 hexamer elements shaded in grey and P1A promoter −10 and −35 elements underlined. Base substitutions in the different fragments are highlighted in bold, whereas the Δ71 deletion is shown by a dash. For each set of promoter fragments, the effect of the 88C mutation gives an estimate of P1A 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. In the present study, 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. The results of the present study 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. Under 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.

AUTHOR CONTRIBUTION

This paper arises from a joint project involving all three authors, who made an equal contribution to its conception, and to the design and interpretation of the experiments. The experimental work was performed by Md. Shahidul Islam.

ACKNOWLEDGEMENTS

We thank Kerry Hollands for help with the footprinting and transcription experiments.

FUNDING

This work was supported by the BBSRC (Biotechnology and Biological Sciences Research Council) [project grant number BBE-020860] and by a Commonwealth Ph.D. studentship to Md.S.I.

REFERENCES


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

Received 10 February 2011/24 March 2011; accepted 8 April 2011
Published as BJ Immediate Publication 8 April 2011, doi:10.1042/BJ20110260

© The Authors Journal compilation © 2011 Biochemical Society