Targets for the MalI repressor at the divergent Escherichia coli K-12 malX-malI promoters.

Lloyd, Georgina; Godfrey, Rita; Busby, Stephen

DOI:
10.1111/j.1574-6968.2010.01907.x

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Lloyd, G, Godfrey, R & Busby, S 2010, 'Targets for the MalI repressor at the divergent Escherichia coli K-12 malX-malI promoters.', *FEMS Microbiology Letters*. https://doi.org/10.1111/j.1574-6968.2010.01907.x

RESEARCH LETTER

Targets for the MalI repressor at the divergent *Escherichia coli* K-12 *malX-malI* promoters

Georgina S. Lloyd, Rita E. Godfrey & Stephen J.W. Busby

School of Biosciences, University of Birmingham, Birmingham, UK

Correspondence: Stephen J.W. Busby, School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. Tel.: +44 121 414 5439; fax: +44 121 414 5925; e-mail: s.j.w.busby@bham.ac.uk

Received 16 November 2009; revised 16 December 2009; accepted 4 January 2010. Final version published online 5 February 2010.

DOI:10.1111/j.1574-6968.2010.01907.x

Editor: Robert Gunsalus

Keywords
Escherichia coli; *malX-malI*; divergent promoters; MalI repressor; operator targets.

Abstract

Random mutagenesis has been used to identify the target DNA sites for the MalI repressor at the divergent *Escherichia coli* K-12 *malX-malI* promoters. The *malX* promoter is repressed by MalI binding to a DNA site located from position −24 to position −9, upstream of the *malX* promoter transcript start. The *malI* promoter is repressed by MalI binding from position +3 to position +18, downstream of the *malI* transcript start. MalI binding at the *malI* promoter target is not required for repression of the *malX* promoter. Similarly, MalI binding at the *malX* promoter target is not required for repression of the *malI*. Although the *malX* and *malI* promoters are regulated by a single DNA site for cyclic AMP receptor protein, they function independently and each is repressed by MalI binding to a different independent operator site.

Introduction

The *Escherichia coli malX* and *malY* genes encode proteins for the transport and metabolism of an as yet unidentified substrate (Zdych *et al.*, 1995; Clausen *et al.*, 2000). They are cotranscribed from a single promoter (the *malX* promoter) whose activity is completely dependent on binding of the cyclic AMP receptor protein (CRP) to a single target centred at position −41.5, i.e. between base pairs −41 and −42, upstream from the *malXY* transcript start (Reidl & Boos, 1991; Lloyd *et al.*, 2008). Upstream of *malX*, the divergent *malI* gene encodes a transcription repressor that represses *malXY* expression (Reidl *et al.*, 1989). Expression of the *malI* gene is dependent on a single promoter that controls divergent transcription initiation from a location that is 85 base pairs upstream from the *malX* promoter transcription startpoint (Lloyd *et al.*, 2008). The *malI* promoter is factor-independent, but can be activated ~1.6-fold by CRP binding to its target at the *malX* promoter, which is centred at position −43.5 with respect to the *malI* promoter transcription startpoint (Fig. 1).

Sequence analysis shows that MalI is a typical member of the LacI family of transcription repressors (Reidl *et al.*, 1989; Weickert & Adhya, 1992). Most members of this family function as dimers that bind to inverted repeats, and Reidl *et al.* (1989) identified the sequence 5′-GATAAAACGTTT TATC-3′ as a likely target for MalI-dependent repression of the *malX* promoter. In this work, we describe a genetic screen to prove that this sequence, located from position −24 to position −9 at the *malX* promoter, and overlapping the −10 hexamer element, is indeed the binding target for MalI. The *malX-malI* regulatory region contains a closely related sequence, 5′-GGTAAAACGTTTTATC-3′, from position +3 to position +18, downstream of the transcription start of the *malI* promoter. We describe a similar genetic screen to prove that this is the target for MalI-dependent autoregulation of the *malI* promoter.

Materials and methods

The starting materials for this work were the EcoRI–HindIII *malX100* and *malI100* fragments described by Lloyd *et al.* (2008). These fragments were inserted into the polylinker of the low copy number *lac* expression vector plasmid, pRW50, encoding resistance to tetracycline (Lodge *et al.*, 1992). Recombinant pRW50 derivatives were propagated in the Δ*lac E. coli* K-12 strain, M182, or its Δ*crp* derivative, as in Hollands *et al.* (2007). Inserts in pRW50 were manipulated

FEMS MICROBIOLOGY LETTERS

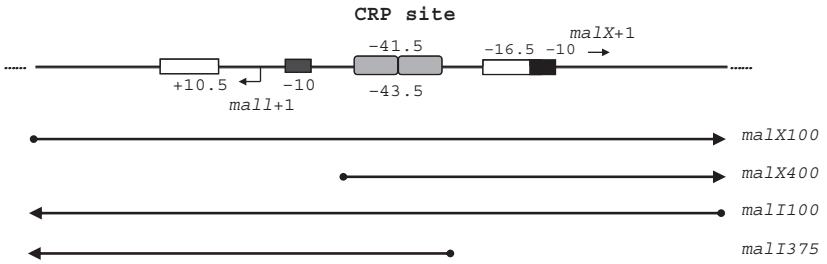

Fig. 1. Organization of the *malX-malI* intergenic region. The top line of the figure illustrates the divergent *malX* and *malI* transcription startpoints, the corresponding −10 hexamer elements (rectangles shaded black), the shared DNA site for CRP (rectangles shaded grey), and the two 16 base pair elements, suggested to be targets for the MalI repressor (unshaded white rectangles). Coordinates above and below the diagram are numbered from the *malX* and *malI* transcription startpoints, respectively. The lower four lines of the figure illustrate the extent of the *malX100*, *malX400*, *malI100*, and *malI375* promoter fragments that are flanked by EcoRI and HindIII sites. The horizontal arrow heads indicate the HindIII sites and the direction of the *malX* promoter (in the case of the *malX100* and *malX400* fragments) and the *malI* promoter (in the case of the *malI100* and *malI375* fragments).

after PCR using the flanking primers D10520 (5′-CCCT GCGGTGCCCCTCAAG-3′) and D10527 (5′-GCAGGTC GTTGAACTGAGCCTGAAATTCAGG-3′) described in Lloyd *et al.* (2008). The shorter *malX400* fragment was generated from *malX100* by PCR using primer D10527 together with D62262 (5′-GACGAATTCCGTTGCGTA ATGTG-3′). Likewise, the shorter *malI375* fragment was generated from *malI100* by PCR using primer D10527 together with D65378 (5′-GGAATTCCAAATTTTAGTGA GGCATAAATCAC-3′). DNA sequences are numbered with the respective transcription start sites labelled as +1 and upstream and downstream sequences are assigned negative and positive coordinates, respectively.

Plasmid pACYC184 was used as a vector for cloning of the *malI* gene, together with the control empty derivative pACYC-Δ*HN* (Mitchell *et al.*, 2007). The *malI* gene, together with its promoter and flanking sequences, was amplified by PCR using genomic DNA from *E. coli K-12* strain MG1655 as a template and primers D63433 (5′-CGA TAAGCTTCAAAACGTTTTATCAAATTTTAGTG-3′) and D63434 (5′-TGGTGCATGCGCAGATAAAGAGAGGATTAT TTCGC-3′). The product was restricted with HindIII and SphI and cloned into plasmid pACYC184 to generate plasmid pACYC-*malI*, which encodes *malI* and resistance to chloramphenicol.

Error-prone PCR, using the flanking D10520 and D10527 primers and *Taq* DNA polymerase, was used to generate libraries of random mutations in the *malX400* or *malI375* promoter fragments, with the respective fragments cloned in pRW50 as the starting templates, using the conditions described by Barne *et al.* (1997). For each promoter, the products of four PCR reactions were restricted with EcoRI and HindIII, purified separately, and cloned into pRW50. After transformation into *E. coli* strain M182 carrying pACYC-*malI*, colonies carrying recombinants were screened on MacConkey lactose indicator plates containing 35 μg mL^{-1} tetracycline and 25 μg mL^{-1} chloramphenicol. Lac^{+} candidates were selected and purified, and for each candidate, the entire EcoRI–HindIII insert was sequenced. Mutations are denoted by their location with respect to the corresponding transcript start and the substituted base on the coding nontemplate strand. Activities of different *malX* and *malI* promoters cloned in pRW50 were deduced from measurements of β-galactosidase expression in M182 or its Δ*crp* derivative, carrying plasmid pACYC-*malI* or the control empty pACYC-Δ*HN* plasmid.

Results and discussion

Identification of the functional MalI-binding target at the *malX* promoter

Figure 1 shows a diagram illustrating the *malX-malI* intergenic region with the transcription start sites for the *malX* and *malI* promoters, the corresponding −10 elements, and the DNA site for CRP that is located at position −41.5 with respect to the *malX* transcription start and position −43.5 with respect to the *malI* transcription start. Figure 1 also shows the locations of two 16 base pair elements, suggested to be the operator targets for the MalI repressor. The aim of the work described here was to investigate this suggestion and to determine the functional operator(s) for each promoter.

In a previous work, Lloyd *et al.* (2008) described how the *malX* promoter could be assayed by cloning the *malX100* fragment into the *lac* expression vector plasmid, pRW50. Measurements of β-galactosidase expression in M182 or its Δ*crp* derivative showed the *malX* promoter to be a typical Class II CRP-dependent promoter, which is consistent with the location of the DNA site for CRP (West *et al.*, 1993). Lloyd *et al.* (2008) also reported that expression of the *malX* promoter::*lac* fusion carried by pRW50 is unaffected by the introduction of a multicopy plasmid carrying the *malX-malI* intergenic region, suggesting that the level of chromosomally encoded MalI is insufficient to repress the *malX* promoter significantly. Thus, to set up a system to measure

MalI-dependent repression of the *malX* promoter, we cloned the *malI* gene into plasmid pACYC184 to generate pACYC-*malI*. Measurements of β-galactosidase expression in M182 cells carrying pRW50 with the *malX100* promoter show that the presence of pACYC-*malI* causes an ∼30-fold reduction in expression, compared with the control with the empty pACYC-Δ*HN* plasmid (Table 1, upper panel). The experiment was then repeated with M182 cells carrying pRW50 with the *malX400* promoter fragment, in which the *malX* promoter sequence upstream of the DNA site for CRP had been removed (illustrated in Fig. 1). The data in Table 1 (upper panel) show that neither *malX* promoter activity nor repression by MalI is substantially affected by the deletion, and thus sequences upstream of the DNA site for CRP must play little or no role.

On MacConkey lactose indicator plates, colonies of M182 carrying pRW50 with either the *malX100* or *malX400* promoter fragments, together with pACYC-*malI*, appear as white Lac^- colonies. In contrast, if pACYC-*malI* is replaced with pACYC-Δ*HN*, colonies have a bright red, clear Lac^+ appearance. Thus, to pinpoint the operator sequences essential for repression of the *malX* promoter by MalI, we used error-prone PCR to generate a library of random mutations in the *malX400* promoter fragment and screened for mutations that resulted in pink or red colonies of cells containing pACYC-*malI*. We reasoned that such colonies

Table 1. Measurement of *malX* promoter activities

Promoter fragment cloned in pRW50	Activity in M182 pACYC-Δ*HN*	Activity in M182 pACYC-*malI*	Repression ratio due to MalI
malX100	1622 ± 170	51 ± 4	31.8
malX400	1735 ± 49	57 ± 3	30.4
malX400 – 24C	3657 ± 130	940 ± 50	3.9
malX400 – 22C	3452 ± 123	881 ± 126	3.9
malX400 – 18G	1131 ± 48	372 ± 12	3.0
malX400 – 17T	8332 ± 37	4925 ± 71	1.7
malX400 – 16A	2676 ± 7	1256 ± 10	2.1
malX400 – 15C	2312 ± 59	1063 ± 11	2.2
malX400 – 14A	6475 ± 52	2101 ± 82	3.1
malX400 – 14C	1895 ± 32	1097 ± 22	1.7
Promoter fragment cloned in pRW50	Activity in M182 pACYC-Δ*HN*	Activity in M182 pACYC-*malI*	Repression ratio due to MalI
malI100	2118 ± 63	138 ± 4	15.3
malI375	1575 ± 28	89 ± 6	17.6
malI375+5C	1728 ± 78	472 ± 14	3.7
malI375+8G	1990 ± 92	1137 ± 35	1.8
malI375+9G	1913 ± 141	744 ± 16	2.6
malI375+11A	2649 ± 191	1415 ± 77	1.9
malI375+12C	2277 ± 149	1196 ± 85	1.9
malI375+13C	2340 ± 54	1407 ± 18	1.7
malI375+16T	2923 ± 71	345 ± 17	8.5
malI375 – 49T	6023 ± 406	956 ± 25	6.3
Promoter fragment cloned in pRW50	Activity in M182Δ*crp* pACYC-Δ*HN*	Activity in M182Δ*crp* pACYC-*malI*	Repression ratio due to MalI
malI100	1230 ± 52	58 ± 4	21.2
malI375	869 ± 98	44 ± 4	19.7
malI375+5C	1455 ± 72	237 ± 23	6.1
malI375+8G	1167 ± 25	487 ± 32	2.3
malI375+9G	1126 ± 23	300 ± 20	3.8
malI375+11A	1399 ± 48	414 ± 43	3.4
malI375+12C	1277 ± 38	397 ± 35	3.2
malI375+13C	998 ± 81	389 ± 41	2.5
malI375+16T	1574 ± 102	129 ± 19	12.2
malI375 – 49T	8798 ± 186	1239 ± 93	7.1

The second and third columns of the table list β-galactosidase activities (in Miller units) measured in the Δ*lac* strain M182 or its Δ*crp* derivative carrying pACYC-*malI* or control plasmid pACYC-Δ*HN*, together with different promoter::*lacZ* fusions cloned in pRW50. Cells were grown aerobically at 37 °C in Luria–Bertani medium containing 35 μg mL^{-1} tetracycline and 25 μg mL^{-1} chloramphenicol to the exponential phase ($OD_{650\,nm}$∼0.4). Each value is the mean ± 1 SD from at least three independent experiments. The upper section of the table lists the effects of different single mutations in the *malX* promoter from position −24 to −14. The lower two parts of the table list the effects of different mutations on *malI* promoter activity. The fourth column of the table lists the factor by which MalI represses expression in each case. Activity measurements were as in Lloyd *et al*. (2008).

would contain pRW50 carrying the *malX400* fragment with mutations that interfered with MalI binding. After screening over 2500 colonies, we identified eight different single-base changes that are shown in Fig. 2. Strikingly, all these substitutions fall in the 16 base pair sequence from position −24 to position −9 that had been suggested to be a target for MalI (Reidl *et al.*, 1989). Our result argues strongly that this sequence alone is necessary for MalI-dependent repression. The upper panel of Table 1 lists the effects of the different point mutations on *malX* promoter activity and MalI-dependent repression. Different mutations reduce repression from ∼30-fold to 1.7- to 3.9-fold. Interestingly, many of the base changes up- or downregulate the activity of the *malX* promoter in the absence of MalI. This is consistent with their location upstream of the −10 hexamer element (Fig. 2). Recall that many *E. coli* promoters carry weakly conserved promoter elements in this region that contribute to the overall promoter activity (Mitchell *et al.*, 2003).

Identification of the functional MalI-binding target at the *malI* promoter

Measurements of β-galactosidase expression in M182 cells carrying pRW50 with the *malI100* promoter show that the presence of pACYC-*malI* causes a sharp reduction in expression, compared with the control with the empty pACYC-Δ*HN* plasmid (Table 1, middle panel). To check whether the DNA site for MalI at the *malX* promoter plays any role in this repression, the experiment was repeated with pRW50 carrying the *malI375* promoter fragment, in which the *malI* promoter sequence upstream of the DNA site for CRP had been removed (illustrated in Fig. 1). The data in Table 1 show that the absence of the DNA site for MalI at the *malX* promoter does not compromise MalI-dependent repression of the *malI* promoter. However, *malI* promoter activity in the shorter *malI375* fragment is reduced by ∼25% compared with the *malI100* fragment. This was expected as we reported previously that upstream sequences are essential for optimal expression from the *malI* promoter (Lloyd *et al.*, 2008).

On MacConkey lactose indicator plates, colonies of M182 carrying pRW50 with either the *malI100* or the *malI375* promoter fragments together with pACYC-*malI* appear as white Lac$^-$ colonies. In contrast, if pACYC-*malI* is replaced with pACYC-Δ*HN*, colonies have a bright red clear Lac$^+$ appearance. Thus, we used error-prone PCR to generate a library of random mutations in the *malI375* promoter fragment and screened for mutations that resulted in pink or red colonies of cells containing *pACYC-malI*. After screening over 2500 colonies, we identified eight different single base changes shown in Fig. 2. Seven of the eight substitutions fall in the sequence from position +3 to position +18, which resembles the operator for MalI at the *malX* promoter, while the eighth is located at position −49.

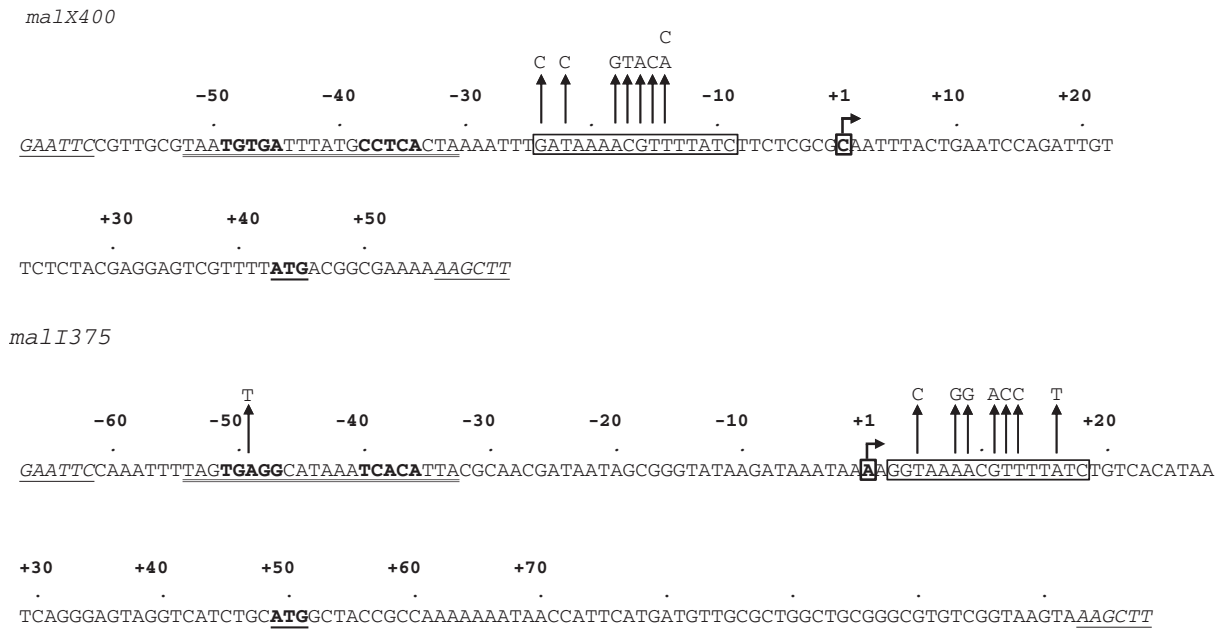

Fig. 2. Base sequence of *malX400* and *malI375* promoter fragments. The figure shows the sequence of the coding nontemplate strand of the *malX400* promoter fragment (upper part of the figure) and the *malI375* promoter fragment (lower part of the figure), from the upstream EcoRI site to the downstream HindIII site (both underlined). Each sequence is numbered from the respective transcript startpoint, which is boxed and marked +1. The shared DNA site for CRP is doubly underlined. The location and nature of each of the point mutations that reduced MalI-dependent repression is indicated and the two 16 base pair MalI-binding elements are highlighted by a box.

The middle panel of Table 1 lists the effects of the different point mutations on *malI* promoter activity and MalI-dependent repression. Different mutations reduce repression from ~17.5-fold to 1.7- to 8.5-fold. Strikingly, with the control pACYC-Δ*HN* plasmid, the +5C, +8G, +9G, +11A, +12C, +13C, and +16T mutations all cause small increases in β-galactosidase expression, while the −49T mutation causes a fourfold increase. The simplest explanation for these observations is that the −49T mutation considerably increases the intrinsic activity of the *malI* promoter, and that the reduction in MalI-dependent repression is a secondary consequence of the promoter being substantially stronger. In contrast, we suggest that the primary effect of the other seven substitutions is to interfere with MalI-dependent repression of the *malI* promoter, but that these changes also produce secondary effects, possibly by altering the structure at the 5′ end of the *malI* transcript.

The lower panel of Table 1 shows the results of an experiment to measure MalI-dependent repression of the *malI* promoter in a Δ*crp* background and the effects of the different mutations. Recall that, unlike the *malX* promoter, the *malI* promoter is active in the absence of CRP (Lloyd *et al.*, 2008). The results show that MalI-dependent repression is slightly greater in the absence of CRP, but each of the different mutations has a similar effect.

Conclusions

Members of the LacI–GalR family of transcriptional repressors are usually functional as dimers, although in some cases, repression depends on the dimerization of dimers or interactions with other proteins, such as CRP (Weickert & Adhya, 1992; Valentin-Hansen *et al.*, 1996). Such repressors bind to inverted repeats at target sites and binding is modulated by a ligand (Weickert & Adhya, 1992; Swint-Kruse & Matthews, 2009). In the case of MalI, the ligand is unknown, but it is assumed that it must be related to the function of MalX and MalY, which, to date, is unknown. Reidl *et al.* (1989), who first discovered the *malI* gene, and the divergent *malXY* operon, identified two 16 base pair sequences, each containing an inverted repeat, that were both suggested to be targets for dimeric MalI. The aim of this work was to investigate these sequences and to determine if repression of the *malXY* and *malI* transcription units required one or both targets. In preliminary work, we attempted a biochemical approach, but we were unable to overexpress soluble functional MalI protein (G.S. Lloyd, unpublished data). Hence, we turned to a genetic approach by setting up an *E. coli* strain where MalI-dependent repression of the *malX* or *malI* promoter yielded a clear phenotype, which was then used to screen for mutations that interfere with repression. Our results with the *malX* promoter unambiguously identify the 16 base pair target from

(a) ***malX sequences***

```
K12      AAATTTGATAAAACGTTTTATCTTCTCGCGCAATTTACTGAATCCAGATTGTTCTCTACGAGGAGTCGTTTTATGACGGCG
O157     AAATTTGATAAAACGTTTTATCTTCTCGCGCAATTTACTGAATCCAGATTGTTCTCTACGAGGAGTCGTTTTATGACGGCG
APEC     AAATTTGATAAAACGTTTTATCTTCTCGCGCAATTTATTGAATCCAGATTGTTCTCTACGAGGAGTCGTTTTATGACGGCG
W3110    AAATTTGATAAAACGTTTTATCTTCTCGCGCAATTTACTGAATCCAGATTGTTCTCTACGAGGAGTCGTTTTATGACGGCG
UTI89    AAATTTGATAAAACGTTTTATCTTCTCGCGCAATTTATTGAATCCAGATTGTTCTCTACGAGGAGTCGTTTTATGACGGCG
CFT073   AAATTTGATAAAACGTTTTATCTTCTCGCGCAATTTATTGAATCCAGATTGTTCTCTACGAGGAGTCGTTTTATGACGGCG
301      AAATTTGATAAAACGTTTTATCTTCTCGCGCAATTTACTGAATCCAGATTGTTCTCTACGAGGAGTCGTTTTATGACGGCG
Sb227    AAATTTGATAAAACGTTTTATCTTCTCGCGCAATTTACTGAATCCAGATTGTTCTCTACGAGGAGTCGTTTTATGACGGCG
8401     AAATTTGATAAAACGTTTTATCTTCTCGCGCAATTTACTGAATCCAGATTGTTCTCTACGAGGAGTCGTTTTATGACGGCG
Sd197    AAATTTGATAAAACGTTTTATCTTCTCGCGCAATTTACTGAATCCAGATTGTTCTCTACGAGGAGTCGTTTTATGACGGCG
Ss046    AAATTTGATAAAACGTTTTATCTTCTCGCGCAATTTACTGAATCCAGATTGTTCTCTACGAGGAGTCGTTTTATGACGGCG
```

(b) ***malI sequences***

```
K12      TATAAGATAAATAAAAGGTAAAACGTTTTATCTGTCACATAATCAGGGAGTAGGTCATCTGCATGGCTACC
0157     TATAAGATAAATAAAAGGTAAAACGTTTTATCTGTCACAAAATCAGGGAGTAGGTCATCTGCATGGCTACC
APEC     TATAAGATAAATAAAAGGTAAAACGTTTTATCTGTCACATAATCAGGGAGTAGGTCATCTGCATGGCTACC
W3110    TATAAGATAAATAAAAGGTAAAACGTTTTATCTGTCACATAATCAGGGAGTAGGTCATCTGCATGGCTACC
UTI89    TATAAGATAAATAAAAGGTAAAACGTTTTATCTGTCACATAATCAGGGAGTAGGTCATCTGCATGGCTACC
CFT073   TATAAGATAAATAAAAGGTAAAACGTTTTATCTGTCACATAATCAGGGAGTAGGTCATCTGCATGGCTACC
301      TATAAGATAAATAAAAGGTAAAACGTTTTATCTGTCACATAATCAGGGAGTAGGTCATCTGCATGGCTACC
Sb227    TATAAGATAAATAAAAGGTAAAACGTTTTATCTGTCACATAATCAGGGAGTAGGTCATCTGCATGGCTACC
8401     TATAAGATAAATAAAAGGTAAAACGTTTTATCTGTCACATAATCAGGGAGTAGGTCATCTGCATGGCTACC
Sd197    TATAAGATAAATAAAAGGTAAAACGTTTTATCTGTCACATAATCAGGGAGTAGGTCATCTGCATGGCTACC
```

Fig. 3. Base sequences upstream of the *malX* and *mall* genes in different strains. The upper part (a) of the figure identifies the *malX* translation start (doubly underlined) and shows the upstream sequences in bacterial genome sequences taken from the XBASE database (Chaudhuri *et al.*, 2008). Sequences are aligned to show the conservation of positioning of putative −10 hexamer elements (shaded box) and 18 base pair DNA sites for Mall binding (singly underlined). The lower part (b) of the figure similarly displays the *mall* translation start and upstream sequences. The listed sequences are taken from the genome sequences of *Escherichia coli* K-12 (K12), *E. coli* O157:H7 EDL933 (O157), *E. coli* APEC O1 (APEC), *E. coli* W3110 (W3110), *E. coli* UTI89 (UTI89), *E. coli* CFT073 (CFT073), *Shigella flexneri* 2a str.301 (301), *Shigella boydii* Sb227 (Sb227), *S. flexneri* 5 str.8401 (8401), *Shigella dysenteriae* Sd197 (Sd197), and *Shigella sonnei* Ss046 (Ss046).

position −24 to position −9 as the target for MalI binding and show that the second 16 base pair element, which is located upstream (Fig. 1), plays little or no role. In contrast, this second element, which is located from position +3 to position +18, downstream of the *malI* transcript start, appears to be the key target for MalI-dependent repression of the *malI* promoter, and the MalI operator site at the *malX* promoter plays little or no role. This repression appears to be independent of CRP. Indeed, repression in the absence of CRP appears to be slightly stronger than in its presence (Table 1).

The divergent *malX* and *malI* promoters share a common DNA site for CRP. As for other divergent bacterial promoters that share an activator-binding site, activation in one direction is largely independent of activation in the opposite direction and this is likely to be due to the low frequency of initiation at most promoters (El-Robh & Busby, 2002). Although the *malX* and *malI* promoters share a DNA site for CRP, each has a separate and independent DNA site for MalI. The *malX* promoter MalI operator is located upstream of the transcript start and overlaps the upstream end of the −10 hexamer, while the *malI* promoter MalI operator is located downstream of the transcript start. This organization is well conserved in the genomes of different strains of *E. coli* and related *Shigella*. Figure 3 shows a comparison of the base sequences upstream of the *malX* and *malI* translation start sites in these genomes, and the comparison emphasizes how the precise locations of −10 elements and MalI operator sequences have been maintained. This provides yet another example of how efficient repression can result from a repressor interacting at different locations at a bacterial promoter (Rojo, 2001; Barnard *et al*., 2004). Interestingly, repression is marginally greater at the *malX* promoter than at the *malI* promoter, and this is consistent with MalI action at the *malI* promoter being autoregulatory.

The *E. coli* K-12 *malX-malI* intergenic regulatory region provides a simple example of 'evolution and tinkering' (Jacob, 1977). The *malX* promoter is an unremarkable CRP-dependent promoter that resembles scores of Class II promoters (Busby & Ebright, 1999) and it can be shut off by MalI. In contrast, although the divergent *malI* promoter resembles a Class II CRP-dependent promoter, it has adapted to ensure that the MalI repressor is always made. Thus, MalI-dependent repression is marginally less efficient compared with the *malX* promoter, the dependence on CRP is relaxed by the DNA site for CRP being located at position −43.5, and the promoter carries seven repeats of a 5′-TAN_8-3′ motif, to facilitate RNA polymerase recruitment (Lloyd *et al*., 2008).

Acknowledgements

This work was funded by a Wellcome Trust program grant. We thank undergraduate project students, Clare Mensley, James Fuller, and Maria Jesus Pina, for some of the constructions.

References

Barnard A, Wolfe A & Busby SJ (2004) Regulation at complex bacterial promoters: how bacteria use different promoter organisations to produce different regulatory outcomes. *Curr Opin Microbiol* **7**: 102–108.

Barne KA, Bown JA, Busby SJ & Minchin SD (1997) Region 2.5 of the *Escherichia coli* RNA polymerase sigma 70 subunit is responsible for the recognition of the extended −10 motif at promoters. *EMBO J* **16**: 4034–4040.

Busby S & Ebright RH (1999) Transcription activation by catabolite activator protein (CAP). *J Mol Biol* **293**: 199–213.

Chaudhuri RR, Loman NJ, Snyder LA, Bailey CM, Stekel DJ & Pallen MJ (2008) xBASE2: a comprehensive resource for comparative bacterial genomics. *Nucleic Acids Res* **36**: D543–D546.

Clausen T, Schlegel A, Peist R, Schneider E, Steegborn C, Chang YS, Haase A, Bourenkov GP, Bartunik HD & Boos W (2000) X-ray structure of MalY from *Escherichia coli*: a pyridoxal 5′-phosphate-dependent enzyme acting as a modulator in *mal* gene expression. *EMBO J* **19**: 831–842.

El-Robh M & Busby S (2002) The *Escherichia coli* cyclic AMP receptor protein bound at a single target can activate transcription initiation at divergent promoters: a systematic study that exploits new promoter probe plasmids. *Biochem J* **368**: 835–844.

Hollands K, Busby SJ & Lloyd GS (2007) New targets for the cyclic AMP receptor protein in the *E. coli* K-12 genome. *FEMS Microbiol Lett* **274**: 89–94.

Jacob F (1977) Evolution and tinkering. *Science* **196**: 161–166.

Lloyd GS, Hollands K, Godfrey RE & Busby SJ (2008) Transcription initiation in the *Escherichia coli* K-12 *malI-malX* intergenic region and the role of the cyclic AMP receptor protein. *FEMS Microbiol Lett* **288**: 250–257.

Lodge J, Fear J, Busby S, Gunasekaran P-S & Kamini N-R (1992) Broad host-range plasmids carrying the *Escherichia coli* lactose and galactose operons. *FEMS Microbiol Lett* **95**: 271–276.

Mitchell JE, Zheng D, Busby S & Minchin S (2003) Identification and analysis of 'extended –10' promoters in *Escherichia coli*. *Nucleic Acids Res* **31**: 4689–4695.

Mitchell JE, Oshima T, Piper S *et al*. (2007) The *Escherichia coli* regulator of sigma 70 protein, Rsd, can up-regulate some stress-dependent promoters by sequestering Sigma 70. *J Bacteriol* **189**: 3489–3495.

Reidl J & Boos W (1991) The *malX malY* operon of *Escherichia coli* encodes a novel Enzyme II of the phosphotransferase system recognising glucose and maltose and an enzyme abolishing the endogenous induction of the maltose system. *J Bacteriol* **173**: 4862–4876.

Reidl J, Romisch K, Ehrmann M & Boos W (1989) MalI, a novel protein involved in the regulation of the maltose system of

Escherichia coli, is highly homologous to the repressor proteins GalR, CytR and LacI. *J Bacteriol* **171**: 4888–4899.

Rojo F (2001) Mechanisms of transcriptional repression. *Curr Opin Microbiol* **4**: 145–151.

Swint-Kruse L & Matthews KS (2009) Allostery in the LacI/GalR family: variations on a theme. *Curr Opin Microbiol* **12**: 129–137.

Valentin-Hansen P, Søgaard-Andersen L & Pedersen H (1996) A flexible partnership: the CytR anti-activator and the cAMP-CRP activator protein, comrades in transcription control. *Mol Microbiol* **20**: 461–466.

Weickert M & Adhya S (1992) A family of bacterial regulators homologous to Gal and Lac repressors. *J Biol Chem* **267**: 15869–15874.

West D, Williams R, Rhodius V, Bell A, Sharma N, Zou C, Fujita N, Ishihama A & Busby S (1993) Interactions between the *E. coli* cAMP receptor protein and RNA polymerase at Class II promoters. *Mol Microbiol* **10**: 789–797.

Zdych E, Peist R, Reidl J & Boos W (1995) MalY of *Escherichia coli* is an enzyme with the activity of a beta C-S lyase (cystathionase). *J Bacteriol* **177**: 5035–5039.