**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

**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 malI 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 malX transcript start. MalI binding at the malI promoter target is not required for repression of the malX promoter. Similarly, MalI binding at the malI promoter target is not required for repression of the malI. Although the malI and malX 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 malI 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′-GGTAAAAACGTTTATC-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 malY100 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 Δlac *E. coli* K-12 strain, M182, or its Δcrp derivative, as in Hollands et al. (2007). Inserts in prW50 were manipulated...
after PCR using the flanking primers D10520 (5'-CCCT 
GCGGTGCCCTCAAG-3') and D10527 (5'-GCAGGTGC 
TTGAACTGAGCCTGAAATCAGG-3') described in 
Lloyd et al. (2008). The shorter malX400 fragment was 
generated from malX100 by PCR using primer D10527 
together with D62262 (5'-GACGAATTCCGTTGCGTAT 
ATGTG-3'). Likewise, the shorter malI375 fragment was 
generated from malI100 by PCR using primer D10527 
together with D65378 (5'-GGAATTCCAAATTTAGTG 
GCCATAAACATC-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 
mall gene, together with the control empty derivative 
pACYC-ΔHN (Mitchell et al., 2007). The mall gene, 
共同其 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 
TAAGCTTCAAAAGTTTATCAAAATTTTAGTGA 
GCCATAAACATC-3') and D63434 (5'-TGGTGCATGCGCATG 
AGAGGATTAT TCGC-3'). The product was restricted with HindIII and 
and cloned into plasmid pACYC184 to generate plasmid pACYC-mall, which encodes mall 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-mall, 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 non-template strand. Activities of different malX 
and mall promoters cloned in pRW50 were deduced from 
measurements of β-galactosidase expression in M182 or its 
Δcrp derivative, carrying plasmid pACYC-mall or the 
control empty pACYC-ΔHN plasmid.

Results and discussion

Identification of the functional Mall-binding 
target at the mallX promoter

Figure 1 shows a diagram illustrating the malX-mall intergenic 
region with the transcription start sites for the malX 
and mall 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 mall transcription start. Figure 1 also 
shows the locations of two 16 base pair elements, suggested 
to be the operator targets for the Mall 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 
Δcrp derivative showed the mallX 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-
mall intergenic region, suggesting that the level of chromo-
somally encoded Mall is insufficient to repress the malX 
promoter significantly. Thus, to set up a system to measure
Mall-dependent repression of the malX promoter, we cloned the mall gene into plasmid pACYC184 to generate pACYC-mall. Measurements of β-galactosidase expression in M182 cells carrying pRW50 with the malX100 promoter show that the presence of pACYC-mall causes an ~30-fold reduction in expression, compared with the control with the empty pACYC-ΔHN plasmid (Table 1, upper panel). The experiment was then repeated with M182 cells carrying pRW50 with the malX400 promoter fragment, in which the mall 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 mall promoter activity nor repression by Mall 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-mall, appear as white Lac+ colonies. In contrast, if pACYC-mall is replaced with pACYC-ΔHN, colonies have a bright red, clear Lac+ appearance. Thus, to pinpoint the operator sequences essential for repression of the malX promoter by Mall, 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-mallII. We reasoned that such colonies

Table 1. Measurement of malX promoter activities

<table>
<thead>
<tr>
<th>Promoter fragment cloned in pRW50</th>
<th>Activity in M182 pACYC-ΔHN</th>
<th>Activity in M182 pACYC-mall</th>
<th>Repression ratio due to Mall</th>
</tr>
</thead>
<tbody>
<tr>
<td>mall100</td>
<td>1622 ± 170</td>
<td>51 ± 4</td>
<td>31.8</td>
</tr>
<tr>
<td>mall400</td>
<td>1735 ± 49</td>
<td>57 ± 3</td>
<td>30.4</td>
</tr>
<tr>
<td>mall400 – 24C</td>
<td>3657 ± 130</td>
<td>940 ± 50</td>
<td>3.9</td>
</tr>
<tr>
<td>mall400 – 22C</td>
<td>3452 ± 123</td>
<td>881 ± 126</td>
<td>3.9</td>
</tr>
<tr>
<td>mall400 – 18G</td>
<td>1131 ± 48</td>
<td>372 ± 12</td>
<td>3.0</td>
</tr>
<tr>
<td>mall400 – 17T</td>
<td>8332 ± 37</td>
<td>4925 ± 71</td>
<td>1.7</td>
</tr>
<tr>
<td>mall400 – 16A</td>
<td>2676 ± 7</td>
<td>1256 ± 10</td>
<td>2.1</td>
</tr>
<tr>
<td>mall400 – 15C</td>
<td>2312 ± 59</td>
<td>1063 ± 11</td>
<td>2.2</td>
</tr>
<tr>
<td>mall400 – 14A</td>
<td>6475 ± 52</td>
<td>2101 ± 82</td>
<td>3.1</td>
</tr>
<tr>
<td>mall400 – 14C</td>
<td>1895 ± 32</td>
<td>1097 ± 22</td>
<td>1.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Promoter fragment cloned in pRW50</th>
<th>Activity in M182 pACYC-ΔHN</th>
<th>Activity in M182 pACYC-mall</th>
<th>Repression ratio due to Mall</th>
</tr>
</thead>
<tbody>
<tr>
<td>mall100</td>
<td>2118 ± 63</td>
<td>138 ± 4</td>
<td>15.3</td>
</tr>
<tr>
<td>mall375</td>
<td>1575 ± 28</td>
<td>89 ± 6</td>
<td>17.6</td>
</tr>
<tr>
<td>mall375 + 5C</td>
<td>1728 ± 78</td>
<td>472 ± 14</td>
<td>3.7</td>
</tr>
<tr>
<td>mall375 + 8G</td>
<td>1990 ± 92</td>
<td>1137 ± 35</td>
<td>1.8</td>
</tr>
<tr>
<td>mall375 + 9G</td>
<td>1913 ± 141</td>
<td>744 ± 16</td>
<td>2.6</td>
</tr>
<tr>
<td>mall375 + 11A</td>
<td>2649 ± 191</td>
<td>1415 ± 77</td>
<td>1.9</td>
</tr>
<tr>
<td>mall375 + 12C</td>
<td>2277 ± 149</td>
<td>1196 ± 85</td>
<td>1.9</td>
</tr>
<tr>
<td>mall375 + 13C</td>
<td>2340 ± 54</td>
<td>1407 ± 18</td>
<td>1.7</td>
</tr>
<tr>
<td>mall375 + 16T</td>
<td>2923 ± 71</td>
<td>345 ± 17</td>
<td>8.5</td>
</tr>
<tr>
<td>mall375 – 49T</td>
<td>6023 ± 406</td>
<td>956 ± 25</td>
<td>6.3</td>
</tr>
</tbody>
</table>

The second and third columns of the table list β-galactosidase activities (in Miller units) measured in the Δlac strain M182 or its Δcrp derivative carrying pACYC-mall or control plasmid pACYC-ΔHN, 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 (OD650nm=0.4). Each value is the mean ± 1SD 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 mall promoter activity. The fourth column of the table lists the factor by which Mall represses expression in each case. Activity measurements were as in Lloyd et al. (2008).
would contain pRW50 carrying the \textit{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 Mal-dependent repression. The upper panel of Table 1 lists the effects of the different point mutations on \textit{malX} promoter activity and Mal-dependent repression. Different mutations reduce repression from \textasciitilde 30-fold to 1.7- to 3.9-fold. Interestingly, many of the base changes up- or downregulate the activity of the \textit{malX} promoter in the absence of MalI. This is consistent with their location upstream of the –10 hexamer element (Fig. 2). Recall that many \textit{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 \textit{malI} promoter**

Measurements of $\beta$-galactosidase expression in M182 cells carrying pRW50 with the \textit{malI100} promoter show that the presence of pACYC-\textit{malI} causes a sharp reduction in expression, compared with the control with the empty pACYC-\textit{DHN} plasmid (Table 1, middle panel). To check whether the DNA site for MalI at the \textit{malX} promoter plays any role in this repression, the experiment was repeated with pRW50 carrying the \textit{malI375} promoter fragment, in which the \textit{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 \textit{malI} promoter does not compromise MalI-dependent repression of the \textit{malI} promoter. However, \textit{malI} promoter activity in the shorter \textit{malI375} fragment is reduced by \textasciitilde 25\% compared with the \textit{malI100} fragment. This was expected as we reported previously that upstream sequences are essential for optimal expression from the \textit{malI} promoter (Lloyd et al., 2008).

On MacConkey lactose indicator plates, colonies of M182 carrying pRW50 with either the \textit{malI100} or the \textit{malI375} promoter fragments together with pACYC-\textit{malI} appear as white Lac$^-$ colonies. In contrast, if pACYC-\textit{malI} is replaced with pACYC-\textit{DHN}, colonies have a bright red clear Lac$^+$ appearance. Thus, we used error-prone PCR to generate a library of random mutations in the \textit{malI375} promoter fragment and screened for mutations that resulted in pink or red colonies of cells containing pACYC-\textit{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 \textit{malX} promoter, while the eighth is located at position –49.

---

**Fig. 2.** Base sequence of \textit{malX400} and \textit{malI375} promoter fragments. The figure shows the sequence of the coding nontemplate strand of the \textit{malX400} promoter fragment (upper part of the figure) and the \textit{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 Malt-dependent repression. Different mutations reduce repression from ~17.5-fold to 1.7- to 8.5-fold. Strikingly, with the control pACYC-DH1 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 Malt-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 Malt-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 Malt-dependent repression of the mali promoter in a Δcrp 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 Malt-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 Malt, 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 216 base pair sequences, each containing an inverted repeat, that were both suggested to be targets for dimeric Malt. 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 Malt protein (G.S. Lloyd, unpublished data). Hence, we turned to a genetic approach by setting up an E. coli strain where Malt-dependent repression of the malXY or mali promoter yielded a clear phenotype, which was then used to screen for mutations that interfere with repression. Our results with the malXY promoter unambiguously identify the 16 base pair target from

![Base sequences upstream of the malX and mali 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 Mal binding (singly underlined). The lower part (b) of the figure similarly displays the mali 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).\n
position — 24 to position — 9 as the target for Mall 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 mall transcript start, appears to be the key target for Mall-dependent repression of the mall promoter, and the Mall operator site at the mall 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 mall and malX 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 mall and malX promoters share a DNA site for CRP, each has a separate and independent DNA site for Mall. The mall promoter Mall operator is located upstream of the transcript start and overlaps the upstream end of the −10 hexamer, while the mall promoter Mall 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 mall translation start sites in these genomes, and the comparison emphasizes how the precise locations of −10 elements and Mall 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 mall promoter, and this is consistent with Mall action at the mall promoter being autoregulatory.

The E. coli K-12 malX-mall 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 Mall. In contrast, although the divergent mall promoter resembles a Class II CRP-dependent promoter, it has adapted to ensure that the Mall repressor is always made. Thus, Mall-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′-TANα-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


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.


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


