Malonate degradation in Acinetobacter baylyi ADP1: operon organization and regulation by MdcR

Julie L. Stoudenmire,1 Alicia L. Schmidt,1 Melissa P. Turnen-Velasquez,1 Kathryn T. Elliott,2 Nicole S. Laniohan,1,3 S. Walker Whitley,† Nickolaus R. Galloway,1,3 Melesse Nune,3‡ Michael West,3 § Cory Momany,3 Ellen L. Neidle1,* and Anna C. Karls1,*

Abstract
Transcriptional regulators in the LysR or GntR families are typically encoded in the genomic neighbourhood of bacterial genes for malonate degradation. While these arrangements have been evaluated using bioinformatics methods, experimental studies demonstrating co-transcription of predicted operons were lacking. Here, transcriptional regulation was characterized for a cluster of mdc genes that enable a soil bacterium, Acinetobacter baylyi ADP1, to use malonate as a carbon source. Despite previous assumptions that the mdc-gene set forms one operon, our studies revealed distinct promoters in two different regions of a nine-gene cluster. Furthermore, a single promoter is insufficient to account for transcription of mdcR, a regulatory gene that is convergent to other mdc genes. MdcR, a LysR-type transcriptional regulator, was shown to bind specifically to a site where it can activate mdc-gene transcription. Although mdcR deletion prevented growth on malonate, a 1 nt substitution in the promoter of mdcA enabled MdcR-independent growth on this carbon source. Regulation was characterized by methods including transcriptional fusions, quantitative reverse transcription PCR, reverse transcription PCR, 5′-rapid amplification of cDNA ends and gel shift assays. Moreover, a new technique was developed for transcriptional characterization of low-copy mRNA by increasing the DNA copy number of specific chromosomal regions. MdcR was shown to respond to malonate, in the absence of its catabolism. These studies contribute to ongoing characterization of the structure and function of a set of 44 LysR-type transcriptional regulators in A. baylyi ADP1.

INTRODUCTION
Malonate in the environment results largely from its presence in plants. While it has long been known that some bacteria use malonate as a source of carbon and energy, recent studies have led to renewed attention to the role that this compound may play in diverse topics such as nitrogen fixation, virulence and bioremediation [1–5]. Malonate, which inhibits succinate dehydrogenase and affects lignolytic enzymes, can impact the success of applying microbial consortia for desired purposes such as bioremediation or lignin degradation [5–7]. To improve understanding of the regulated degradation of malonate, we investigated the transcriptional control of this important process in Acinetobacter baylyi ADP1, a genetically malleable soil bacterium [8].

Clustered genes for malonate degradation, present in diverse microbes, typically include genes for transcriptional regulators in the GntR or LysR families [9]. Several studies address transcription of these catabolic genes (often designated mdc or mat), yet most regulatory information about malonate metabolism is inferred from bioinformatics. In A. baylyi, the mdcR regulatory gene is convergent to genes involved in malonate metabolism (Fig. 1). This arrangement is similar to that in Klebsiella pneumoniae (Fig. S1, available in the online Supplementary Material), the only bacterium in which MdcR, a LysR-type transcriptional regulator, has...
been studied [10, 11]. *A. baylyi* and *K. pneumoniae*, both non-motile gamma-proteobacteria, have different arrangements of *mdc* genes. In *K. pneumoniae*, *mdcF* encodes a transporter belonging to a large family of mostly uncharacterized membrane proteins, classified in the Interpro database as the IPR004776 group [12]. This gene is not in *A. baylyi*, which instead has genes (*mdcLM*) that encode a differently classified two-subunit malonate/sodium symporter, belonging to Interpro groups IPR00469 and IPR004691 [13, 14]. In *K. pneumoniae*, the assumption that the *mdcABC-DEFGH* genes form an operon has not been verified. Moreover, while the comprehensive bioinformatic analysis of genes for malonate utilization demonstrates clustered gene organization, experimental evidence for co-transcription is lacking [9]. In this study, we investigated the transcription of *mdcABCDEGHLM* in *A. baylyi* to resolve the conflicting assumption that they form a single operon [9] and a different prediction that they form two operons [15]. This study contributes to a larger effort to understand MdcR, which has the potential to be used for metabolic engineering and as a biosensor to detect malonate [16].

In *Klebsiella* strains, there is often a small 11 nt overlap in the 3′ ends of convergent *mdcR* and *mdcH* genes. Additionally, there is a conserved gene upstream of *mdcR* that is likely to encode malate dehydrogenase (Fig. S1). In *K. pneumoniae*, a transcriptional start site was identified immediately upstream of *mdcR*. The adjacent dehydrogenase gene was not considered to be co-transcribed with the regulatory gene [11]. In *A. baylyi* ADP1, *mdcR* is convergent to a non-overlapping gene, *mdcM*. In this strain, the upstream neighbour of *mdcR* is ACIAD1763, a gene that is annotated as *srpH* to indicate that it may encode a serine O-acetyltransferase. There are two paralogues in ADP1, one also annotated as *srpH* (ACIAD1584) and another as *cysE* (ACIAD2091). Although *mdcR* and *srpH* do not co-localize in other *Acinetobacter* strains or known bacteria, we investigated whether their close proximity in ADP1 reflects co-transcription in this strain (Fig. S2).

In our initial attempts to characterize *mdcR* transcription, cDNA levels were too low for conclusive analysis. To address this problem, we developed a method to increase transcription by generating a multi-copy array of the *mdc*
region on the *A. baylyi* chromosome. This approach exploits the high efficiency of natural transformation and homologous recombination in *A. baylyi* to alter the copy number of genomic segments [17, 18]. This method facilitated the analysis of *mdcR* and *srpH* transcripts.

Studies of MdcR were conducted in conjunction with long-term studies of transcriptional regulation by LysR-type proteins, homologues that form the largest family of bacterial transcriptional regulators. Functional studies of MdcR contribute to comprehensive analyses of all 44 LysR-type regulators in *A. baylyi* [19]. The tractable genetic system of this strain makes it ideal for undergraduate researchers. Many studies in this report were first carried out as part of a college laboratory course at the University of Georgia that enables undergraduate students to conduct authentic research projects. The role of MdcR in malonate metabolism was demonstrated using mutational analysis, transcriptional fusions, transcript analysis and protein–DNA binding studies.

### METHODS

#### Strains and growth conditions

*E. coli* plasmid hosts DH5α and TOP10 (Invitrogen) were grown in Luria–Bertani (LB) medium, also known as lysogeny broth (10 g Bacto tryptone, 5 g yeast extract and 5 g NaCl per litre) [20]. For protein expression, *E. coli* host BL21(DE3) RIPL (Agilent Technologies) was grown in an auto-induction medium [21]. *A. baylyi* strains, derived from the wild-type ADP1 (Table 1), were grown in LB or minimal medium.

#### Table 1. *A. baylyi* strains and plasmids

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant characteristics*</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADP1</td>
<td>Wild-type strain (BD413)</td>
<td>[22]</td>
</tr>
<tr>
<td>ACN1261</td>
<td>ΔmdcR::sacB::Km51261</td>
<td>This study</td>
</tr>
<tr>
<td>ACN1281</td>
<td>ΔmdcRS1281</td>
<td>This study</td>
</tr>
<tr>
<td>ACN1703</td>
<td>ΔmdcR::sacB::Km51261; mdcAS1705 (T→A, 1759.559); mdcAS1703 (T→C, 1759.427)</td>
<td>This study</td>
</tr>
<tr>
<td>ACN1704</td>
<td>ΔmdcR::sacB::Km51261; mdcAS1705 (T→A, 1759.559); mdcAS1704 (T→C, 1759.545)</td>
<td>This study</td>
</tr>
<tr>
<td>ACN1705</td>
<td>ΔmdcR::sacB::Km51261; mdcAS1705 (T→A, 1759.559)</td>
<td>This study</td>
</tr>
<tr>
<td>ACN1709</td>
<td>ΔmdcRS1281; mdcAS1705 (T→A, 1759.559)</td>
<td>This study</td>
</tr>
<tr>
<td>ACN1713</td>
<td>srpH::ISS1713</td>
<td>This study</td>
</tr>
<tr>
<td>ACN1714</td>
<td>mdcA::lacZ::Km51714; ΔmdcRS1281; mdcAS1705</td>
<td>This study</td>
</tr>
<tr>
<td>ACN1715</td>
<td>mdcA::lacZ::Km51714; ΔmdcRS1281</td>
<td>This study</td>
</tr>
<tr>
<td>ACN1728</td>
<td>mdcA::lacZ::Km51714; mdcAS1705</td>
<td>This study</td>
</tr>
<tr>
<td>ACN1729</td>
<td>mdcA::lacZ::Km51714</td>
<td>This study</td>
</tr>
<tr>
<td>ACN1770</td>
<td>srpH::ISS1713, multiple copies of mdc-region amplicon (linear pBAC1342 X ACN1713)</td>
<td>This study</td>
</tr>
<tr>
<td>ACN1772</td>
<td>mdcA::lacZ::Km51714, multiple copies of mdc-region amplicon (linear pBAC1342 X ACN1729)</td>
<td>This study</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Relevant characteristics*</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBAC1033</td>
<td>ApR; ΔmdcRS1281 (1766.508–1766.937; 1767.853–1768.571) in pUC19</td>
<td>This study</td>
</tr>
<tr>
<td>pBAC1052</td>
<td>ApR, KmR; ΔmdcR::sacB::Km51261; sacB::KmR cassette inserted in XhoI site of pBAC1033</td>
<td>This study</td>
</tr>
<tr>
<td>pBAC1256</td>
<td>ApR; mdcA region (1759.074–1759.997) in pUC19</td>
<td>This study</td>
</tr>
<tr>
<td>pBAC1275</td>
<td>ApR, KmR; srpH region (1767.763–1768.788) in pCR4-TOPO</td>
<td>This study</td>
</tr>
<tr>
<td>pBAC1277</td>
<td>ApR, KmR, SpR/SmR; srpH::ISS1713; insertion of IS1 in srpH (XbaI site 1768.282) of pBAC1275</td>
<td>This study</td>
</tr>
<tr>
<td>pBAC1285</td>
<td>ApR, KmR; mdcA::lacZ::Km51714; mdcA region (1759.074–1759.997) with reporter cassette inserted after 1759.642, see text for additional details</td>
<td>This study</td>
</tr>
<tr>
<td>pBAC1286</td>
<td>ApR, KmR; mdcA::lacZ::Km51714, promoter mutation T→A, position 1759.559 (only difference to pBAC1285)</td>
<td>This study</td>
</tr>
<tr>
<td>pBAC1342</td>
<td>ApR; DNA surrounding the mdc region (1767.369–1770.768; 1756.768–1758.403) in pUC18; DNA joined as depicted in Fig. 5</td>
<td>This study</td>
</tr>
<tr>
<td>pBAC1418</td>
<td>ApR; mdcR region (1767.855–1768.937) in pET28b(+);SapKO–CH.BspQI; used for MdcR purification</td>
<td>This study</td>
</tr>
<tr>
<td>pUC18 and pUC19</td>
<td>ApR; cloning vectors</td>
<td>[51]</td>
</tr>
<tr>
<td>pRMJ1</td>
<td>ApR, KmR; source of sacB::KmR cassette</td>
<td>[25]</td>
</tr>
<tr>
<td>pOK6</td>
<td>ApR; KmR; source of promoterless lacZ::KmR cassette</td>
<td>[27]</td>
</tr>
<tr>
<td>pCR4-TOPO</td>
<td>ApR; KmR; cloning vector</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>pUI1638</td>
<td>SpR/SmR; source of IS1 cassette</td>
<td>[26]</td>
</tr>
<tr>
<td>pET28b(+);SapKO–CH.BspQI</td>
<td>Expression vector used to make pBAC1418</td>
<td>[28]</td>
</tr>
</tbody>
</table>

*Underlined numbers correspond to positions on the *A. baylyi* ADP1 chromosome (GenBank CR543861).
medium with 10 mM succinate, 20 mM pyruvate or 20–40 mM malonate as a carbon source [22, 23]. When needed, antibiotics were added at final concentrations of 25 µg kanamycin (Km) ml⁻¹, 12.5 µg spectinomycin (Sp) and streptomycin (Sm) ml⁻¹, or 150 µg ampicillin (Ap) ml⁻¹. Bacteria were grown at 37°C.

**Plasmid construction**

Standard methods [20] were used to make the plasmids, which are listed in Table 1. Oligonucleotides for PCR are listed in Table S1. All restriction and modification enzymes were from New England Biolabs (NEB), unless otherwise indicated. Some DNA pieces were joined by SOEing (splicing by overlap extension) PCR [24]. Genomic sequence coordinates for ADP1 correspond to GenBank (splicing by overlap extension) PCR [24]. Genomic indicated. Some DNA pieces were joined by SOEing were from New England Biolabs (NEB), unless otherwise listed in Table S1. All restriction and modification enzymes which are listed in Table 1. Oligonucleotides for PCR are Standard methods [20] were used to make the plasmids, Excised from pRMJ1 by SalI digestion [25].

To make an mdcA::::lacZ-KmR DNA, was obtained by selection for pBAC1285. ACN1281 was transformed by linearized pBAC1033 donor DNA to make a ΔmdcR mutant, ACN1281. Similarly, loss of sacB was used to select a mutant, ACN1709, after strain ACN1705 was transformed with linearized pBAC1033. ACN1705, described below, has a mdcA promoter mutation in addition to the sacB-KmR cassette. ACN1709 carries this promoter mutation and also an unmarked mdcR deletion.

The mdcA::::lacZ-KmR DNA was introduced into several strains by allelic replacement. ACN1281 was transformed by linearized pBAC1285 to generate ACN1715. ACN1281 was transformed by linearized pBAC1286 to introduce a promoter mutation together with mdcA::::lacZ-KmR to create ACN1714. Cell lysates with chromosomal DNA, made as described previously [30, 31], from ACN1715 and ACN1714 were used to transform ADP1. The resulting strains ACN1278 and ACN1279 have mdcA::::lacZ-KmR, wild-type mdcR and either a wild-type or mutated mdcA promoter, respectively. To introduce termination signals from the ßIS cassette upstream of mdcR, strain ADP1 was transformed with linearized pBAC1277. The resulting mutant, ACN1713, was selected by drug resistance (SmR/SmR).

**Increase in the copy number of specific chromosomal regions**

Engineered fragments can be used as donor DNA to create multiple copies of chromosomal regions [17, 18]. With this method, the donor DNA and chromosome recombine to form a precise duplication that can undergo further copy-number increase. This technique generated strains ACN1770 and ACN1772 after transformation of ACN1713 or ACN1729, respectively, with linearized pBAC1342 (digested with AatII). Multiple copies of a specific chromosomal region were selected by demanding resistance to high levels of antibiotics (final concentrations of 1 mg Sp and Sm ml⁻¹ for ACN1770; and 1 mg Km ml⁻¹ for ACN1772).

The copy number of the repeated chromosomal region was determined by quantitative (q) PCR. For the indicated A. baylyi strains, genomic DNA (1 ng: Quick Universal
Genomic Prep Kit, Zymo Research) and primers pairs (200 nM) for rpoA or mdcR (MTV274 and MTV275, or AA15 and AA16, respectively; Table S1) were used for qPCR with SYBR Select Master Mix (Applied Biosystems) as recommended by the manufacturer for standard mode in the StepOnePlus Real-Time PCR system (Applied Biosystems). The rpoA gene is single copy in the A. baylyi genome and serves as an internal control to evaluate the number of copies of mdcR. Primers and annealing conditions were optimized to give 90–100 % PCR efficiency and amplification of the single target was confirmed by melting-curve analysis. For calculating the copy number of the repeated chromosomal sequence, qPCR was performed with two biological replicates for each A. baylyi strain. Standard curves were generated using ADP1 genomic DNA (12.5 ng to 0.02 ng) for quantification of rpoA and mdcR; the ratio of the calculated concentrations for mdcR and rpoA is assumed to equal the ampiclon copy number.

Isolation of mdcA promoter mutations

Wild-type DNA was amplified by error-prone PCR with primers NSL99 and NSL100. Quantification of cDNA (rpoA, mdcA, mdcL, mdcR and srpH) indicated PCR efficiencies between 94 and 100 %. Transcript levels for the target genes were normalized to the rpoA transcript level, which is relatively constant in the wild-type and mutant A. baylyi strains grown under the indicated growth conditions.

The design of the reverse transcription PCR (RT-PCR) studies used to determine operon structure is described in detail in Results. cDNA was generated with gene-specific primers (see Results) in the SuperScript III first-strand synthesis system (Invitrogen) and used in PCR with LongAmp Taq (NEB) and primer pairs that are described in Results. PCR products were assessed by agarose gel electrophoresis.

β-Galactosidase assays to assess mdcA::lacZ transcriptional fusions

An mdcA::lacZ fusion replaced mdcA on the chromosome of several recipients. β-Galactosidase (LacZ) activity was used to evaluate transcription of the reporter gene. Liquid cultures were grown overnight on pyruvate with or without 10 mM malonate added as an effector. Cells were subcultured to an OD600 of 0.7 in fresh medium and lysed with chloroform and sodium dodecyl sulfate. LacZ was assayed as described previously [32].

MdcR-His5 purification

C-terminal penta-histidine tagged protein (MdcR-His5) was purified from E. coli BL21 (DE3) RPl (pBAC1418) as described in [33] with the following limited modifications. Cells were lysed in binding buffer containing 20 mM Tris, 500 mM NaCl, 25 mM imidazole, 10 % glycerol, 10 mM β-mercaptoethanol, pH 8; the final protein elution buffer (Q elution buffer) contained 20 mM Tris, 500 mM NaCl, 500 mM imidazole, 10 % glycerol, 10 mM β-mercaptoethanol, pH 9. MdcR-His5 protein had low solubility in the absence of imidazole and glycerol in the purification buffers. Samples were flash frozen in liquid nitrogen for storage at −80 °C until use.

Electrophoretic mobility shift assay (EMSA) with MdcR-His5

To assess binding of purified MdcR-His5 to DNA, three PCR fragments in the mdcA operator-promoter region were generated using 6-carboxyfluorescein (6-FAM)-labelled primers. These primers, NSL91 and NSL94, were paired with each other or with NSL92 and NSL93. Protein-DNA binding reactions included 2 nM DNA fragment incubated with or without MdcR-His5 (20–320 nM) in buffer (0.1 mg BSA ml−1, 80 mM Tris-acetate pH 8, 100 mM KCl, 25 mM ammonium acetate, 5 mM magnesium acetate, 0.1 mM EDTA, 1 mM DTT) at 37 °C for 1 h. Some binding reactions included 1 mM malonate, as indicated in the text. Unlabelled PCR fragments were generated for specific (NSL93 and NSL80) and non-specific (MTV20 and MTV32) competition studies used to determine operon structure is described in detail in Results. cDNA was generated with gene-specific primers (see Results) in the SuperScript III first-strand synthesis system (Invitrogen) and used in PCR with LongAmp Taq (NEB) and primer pairs that are described in Results. PCR products were assessed by agarose gel electrophoresis.
assays; the competitor DNA was added at 0-, 1-, 10-, 100- and 150-fold molar excess relative to the 6-FAM-labelled target DNA fragment (PCR-generated with NSL91 and NSL94) in DNA binding reactions containing 60 nM MdcR-His6. All binding reactions were analysed on 6% polyacrylamide gels run in 1× TAE (40 mM Tris pH 7.6, 20 mM acetic acid, 1 mM EDTA), which contained 1 mM malonate when it was added to the binding reactions, and visualized using a Typhoon imager (GE Healthcare) with 488 nm excitation, 520 nm emission and 100 µm resolution. Equilibrium dissociation constants (Kd) were determined as described in [34].

**Computer prediction programs**

Software available online was used to evaluate genome organization, to predict promoter sequences and to predict MdcR-binding DNA sequences. These programs include Biocyc [35], Absynte [36], Pattern Locator [37], Motif Locator [38] and BPROM [39].

**RESULTS**

**MdcR-mediated regulation**

*A. baylyi* mutants without *mdcR* were made to test our prediction that MdcR activates transcription of the *mdc* genes by binding to a sequence (TTAA-N7-TTAA) matching the characteristics of DNA-binding sites that are recognized by LysR-type regulators [19]. This potential binding site was identified using the program Pattern Locator [37]. Our initial model was expanded in a comprehensive analysis of DNA sequences predicted to be involved in malonate metabolism in diverse bacteria. This bioinformatic approach identified a wide array of sequences likely to be regulated by MdcR homologues, and from these sequences a generally applicable bacterial DNA-binding site consensus was identified using the program Pattern Locator [37]. Our initial model was expanded in a comprehensive analysis of DNA sequences predicted to be involved in malonate metabolism in diverse bacteria. This bioinformatic approach identified a wide array of sequences likely to be regulated by MdcR homologues, and from these sequences a generally applicable bacterial DNA-binding site consensus was deduced for MdcR [9]. In strain ADP1, both predictions identify the same region upstream of *mdcA*, shown in Fig. 1. According to this model, the loss of MdcR would prevent activated transcription of *mdcA*, and any other co-transcribed *mdc* genes, thereby preventing growth on malonate as a sole carbon source. In newly constructed mutants, a sacB-Km cassette replaced *mdcR* (strain ACN1261) or *mdcR* was deleted (strain ACN1281). Unlike the wild-type, both mutants failed to grow with malonate as the carbon source despite growth on other carbon sources. This result is consistent with a model that MdcR activates the *mdcA* promoter (P*mdcA*) by binding upstream of the *mdcA* coding sequence [9, 19].

Despite attempts to isolate colonies that grow on malonate without *mdcR*, no spontaneous mutants emerged. Next, the *mdcA* promoter region was amplified by error-prone PCR, since PCR-generated DNA can introduce chromosomal mutations directly by allelic replacement in *A. baylyi* [40]. After a strain lacking *mdcR* (ACN1261) was transformed with mutated DNA, three mutants were selected by growth on malonate (ACN1703, ACN1704 and ACN1705). All carried the same mutation, as shown in Fig. 1. Although two of these three strains had an additional mutation (Table 1), we focused on the T→A substitution (allele designation *mdcA*51705). When this mutation was combined with an unmarked *△mdcR*, the resulting strain (ACN1709) grew on malonate, as did its counterpart with the sacB-Km cassette in place of *mdcR* (ACN1705). The doubling times for growth on 20 mM malonate were determined to be 58 ±8 min and 65±15 min for the wild-type strain, ADP1 and mutant, ACN1709, respectively.

**Transcriptional analysis**

To determine the position of the mutation relative to the start of *mdcA* transcription, we used 5′-RACE with malonate-grown ADP1 [41]. The identified start site, +1 in Fig. 1, revealed that the mutation lies in the −10 region of P*mdcA*. Plasmids were generated with the wild-type promoter (pBAC1285) and the mutated promoter, P*mdcA*51281 (pBAC1286), controlling a promoterless lacZ gene. The plasmid-borne transcriptional fusions replaced *mdcA* on the chromosome of strains with wild-type *mdcR* or a deletion, Δ*mdcR*51281 (Table 1). The lacZ gene, inserted into the *mdcA* coding sequence, prevents malonate from being degraded and allows its role as a non-metabolized effector to be investigated.

β-Galactosidase (LacZ) activity was used to assess *mdcA* transcription. In pyruvate-grown cultures of a strain that differed from the wild-type only by chromosomal insertion of the lacZ fusion (ACN1729), malonate was needed for high-level transcription controlled by P*mdcA* (Fig. 2). With this promoter, high-level gene expression also required *mdcR* (ACN1729 compared to ACN1715, Fig. 2). However, with the T→A promoter mutation that allowed MdcR-independent growth on malonate (P*mdcA*51705), *mdcR* was not needed to express *mdcA*:lacZ. High-level gene expression from P*mdcA*51705 was observed without malonate and was not significantly increased by addition of the effector (ACN1714, Fig. 2). However, in the presence of *mdcR*, malonate increased transcription from the mutated promoter, P*mdcA*51705. Without malonate, expression of the reporter gene was higher from P*mdcA*51705 than from the wild-type promoter, but it was lower than in the comparable strain lacking MdcR (ACN1728 versus ACN1714; Fig. 2). These results suggest that MdcR lowers the high-level transcription from P*mdcA*51705 when no effector is present. Nevertheless, this level is still elevated compared to that for P*mdcA* under the same conditions. Malonate caused MdcR-mediated increases in transcription for both promoters (P*mdcA* and P*mdcA*51705).

**mdcABCDEGH operon structure**

Bacterial *mdc* genes often overlap. In *A. baylyi* ADP1, overlap between the 3′ end of an upstream coding sequence and the 5′ end of its neighbour occurs with *mdcA/mdcB*, *mdcB/mdcC*, *mdcC/mdcD*, *mdcD/mdcE* and *mdcE/mdcG*. Only 12 nucleotides separate *mdcG* and *mdcH*. This arrangement suggests co-transcription of *mdcABCDEGH*. The coding sequences of *mdcH* and *mdcL* are separated by 47
nucleotides. Predictions in the Biocyc database suggest that the mdcLM genes form a distinct operon (Fig. S2) [15].

We used qRT-PCR for further evaluation. Wild-type cells transcribed mdcA at higher levels with malonate than pyruvate as the carbon source (an approximately 82-fold increase, data not shown). This result is consistent with the mdcA::lacZ studies in which malonate was used as an effector. Similar to mdcA, mdcL transcription was higher when cells were grown on malonate compared to pyruvate (data not shown).

In pyruvate-grown cells, transcription from the mutated promoter, P_{mdcA51705} (ACN1709), was approximately 50-fold higher for mdcA than from its native promoter, P_{mdcA} (ADP1) (Fig. 3). Similarly, this promoter mutation increased mdcL transcription significantly compared to the native promoter (Fig. 3). The ability of a single mutation at −10 of the mdcA promoter to considerably increase mdcL transcription indicates that this gene is expressed from P_{mdcA}.

We tested mdcL gene expression when transcription from P_{mdcA} is blocked. Using RT-PCR, mdcL transcription for a strain with elevated gene expression driven by P_{mdcA51705} (ACN1709) was compared to that in two strains with transcription terminated by a lacZ-KmR cassette (ACN1729 and ACN1772). As depicted in Fig. 4(a), ACN1729 and ACN1772 each have the lacZ-KmR cassette inserted in mdcA. The construction of ACN1772 and results with this strain are described below. The convergent orientation of genes in the lacZ-KmR cassette prevents extension of transcripts from P_{mdcA} into mdcL. With RNA from pyruvate-grown cultures, an mdcL-specific primer (NSL105, Table S1) generated cDNA that served as a PCR template. One primer set (NSL101 and NSL102) amplifies a 234 bp fragment corresponding to the 5′ end of mdcL (F_A, Fig. 4b). Another set (NSL101 and NSL105) amplifies a 556 bp fragment that extends through mdcH and mdcL (F_B, Fig. 4b). A third primer set (NSL102 and ALS46) does not amplify cDNA since one primer is outside the region of cDNA synthesis. This set serves as a negative control to detect residual genomic DNA in samples obtained by RNA isolation and reverse transcription (817 bp, F_C, Fig. 4b). All fragments were detected when genomic DNA was used as template (Fig. 4c). Fragments F_A and F_B, corresponding to mdcL transcripts, were observed when mdc-gene expression was controlled by P_{mdcA51705} (ACN1709, Fig. 4c). As appropriate for cDNA, the negative control using the third primer set failed to generate any F_C-sized fragments (Fig. 4c). These results confirm the co-transcription of mdcH and mdcL from P_{mdcA51705}.

If mdcL transcription were controlled solely by P_{mdcA}, then no mdcL transcription should be detected in strains with the chromosomal mdcA::lacZ fusion. With ACN1729 (P_{mdcA51705} mdcA::lacZ) cDNA as a template, fragments F_B and F_C were not detected. However, a small amount of fragment F_A was evident as a faint band (data not shown). These
results, which indicated an mdcL transcript that does not include mdcH, raised the possibility of another promoter upstream of mdcL. To boost confidence in this interpretation, we sought to increase the amount of this transcript.

**Tandemly repeated chromosomal regions increase mRNA and help reveal more than one promoter controlling mdcL**

In A. baylyi, tandem arrays of chromosomal segments can be generated via transformation with linear DNA that undergoes homologous recombination [17, 18]. We used this method to increase gene dosage of a chromosomal segment with mdcA::lacZ-KmR (in ACN1729; Fig. 5). The end points of this region (the amplicon) are precisely defined by the transforming DNA. Linear DNA from pBAC1342 generated an amplicon encompassing all mdc genes and also srpH. Transformants were selected for resistance to high concentrations of kanamycin, which requires multiple copies of the drug-resistance gene. Quantification by qPCR revealed an average of 29 amplicon copies per chromosome in ACN1770. Gene dosage was similarly increased in the mdcR region, since initial studies resulted in faint RT-PCR signals for mdcR transcripts (data not shown). We blocked transcripts originating upstream of srpH by disrupting it with a drug-resistance cassette (in ACN1713). Transcripts originating upstream of srpH should not extend into mdcR due to termination signals in this region (35 region) and TTGTA-TAAA (−10 region) [39].

**Multiple promoters also control mdcR**

Gene dosage was similarly increased in the mdcR region, since initial studies resulted in faint RT-PCR signals for mdcR transcripts (data not shown). We blocked transcripts originating from the upstream gene (srpH) by disrupting it with a drug-resistance cassette (in ACN1713). Transcripts originating upstream of srpH should not extend into mdcR due to termination signals in this region (35 region) and TTGTA-TAAA (−10 region) [39].

To investigate co-transcription of mdcR and srpH, a gene-specific mdcR primer (NSL103) was used to make cDNA transcriptions (data not shown). We blocked transcripts originating upstream of srpH by disrupting it with a drug-resistance cassette (in ACN1713). Transcripts originating upstream of srpH should not extend into mdcR due to termination signals in this region (35 region) and TTGTA-TAAA (−10 region) [39].

**Fig. 4.** Co-transcription analysis of mdcA/mdcL. (a) Key features of strains used in analysis of mdcA/mdcL co-transcription are depicted including the T→A mutation (●) in the mdcA promoter of ACN1709 and the chromosomal lacZ-KmR insertion in mdcA for ACN1729 and ACN1772. For ACN1772, the box delineates the chromosomal region that is tandemly repeated (amplicon) and the red and blue lines correspond to the ends of the linear DNA fragment used to generate the repeated chromosomal region (as depicted in Fig. 5). (b) A gene-specific primer for mdcL was used to make cDNA from RNA (dashed arrow). PCR primers can generate fragments (open boxes) of size 0.23 kb (F<sub>A</sub>), 0.56 kb (F<sub>B</sub>) or 0.82 kb (F<sub>C</sub>). (c) Genomic DNA (gDNA) or cDNA was used as template for PCR, and primer sets could generate the fragments listed above each lane of the agarose gel. DNA molecular weight markers (2 log ladder; NEB) are in the first lane. The gel shown is representative of results seen for RT-PCR analysis of RNA from at least two biological replicates.
from pyruvate-grown cultures. As depicted in Fig. 6(b), primer sets were designed to produce srpH-region fragments: 169 bp F_D (primers AA16 and NSL107), 816 bp F_E (primers AA16 and ALS45), 381 bp F_F (primers NSL104 and NSL107) and 873 bp F_G (primers NSL108 and AA18). All fragments were produced with wild-type genomic template DNA (Fig. 6c). F_E was not detected with any cDNA template since one primer for this fragment anneals to DNA outside the region of cDNA synthesis Fig. 6(b). Fragments F_D, F_F and F_G were detected with wild-type cDNA. Similar results were observed with cDNA from ACN1772, a strain with multiple copies of the wild-type srpH-mdcR region Fig. 6(c). Detection of the largest fragment (F_G) indicated that srpH and mdcR are co-transcribed.

If a promoter for srpH (P_{srpH}) solely accounts for mdcR transcription, then disruption of srpH by \( \Delta S \) should prevent growth on malonate by blocking mdcR transcription. However, an srpH-disrupted mutant able to grow on malonate (data not shown). Furthermore, fragments F_D and F_G corresponding to mdcR transcripts were detected in cDNA from ACN1770, the mutant carrying multiple copies of the region with \( \Delta S \)-disrupted srpH (Fig. 6c). These fragments suggest another promoter for mdcR located within srpH between the \( \Delta S \)-cassette insertion and the end of fragment F_C. Computational sequence analysis predicted one transcriptional start site in this region, 139 nt upstream of the mdcR coding sequence, with the following promoter sequence: TTTACG (−35 region) and CCGTATCAT (−10 region) [39].

The relative levels of mdcR transcripts originating from promoters upstream and downstream of \( \Delta S \) were assessed with qRT-PCR. With the wild-type srpH configuration, measured cDNA levels for mdcR represent a combination of transcripts initiating from both promoters. If all transcripts from the 5’ end of srpH extend through mdcR as an operon, and if additional mdcR transcripts initiate from a proximal promoter, then the ratio of srpH/mdcR would be less than unity. However, in strain ACN1772, with a wild-type srpH configuration, this ratio is approximately unity (Fig. 6d). This result suggests that the level of mdcR transcription initiating from the proximal promoter is very low or that there are srpH transcripts that do not extend into mdcR. With \( \Delta S \) blocking transcription from P_{srpH} in ACN1770, the detected mdcR transcript levels dropped significantly (ACN1772 in Fig. 6d). This drop resulted in a more than 20-fold increase in the srpH/mdcR cDNA ratio when srpH cDNA was assessed with primers upstream of the drug-marker insertion (Fig. 6d). Therefore, it appears that an srpH-mdcR operon accounts for the majority of mdcR transcription.

**MdcR-DNA binding sequences**

The identification of multiple sites of transcriptional initiation in the region of *A. baylyi* mdc genes raised questions about whether MdcR binds near these or other promoter regions elsewhere on the chromosome. A previous analysis of DNA sequences likely to be involved in malonate metabolism in diverse bacteria led to the identification of a conserved sequence predicted to bind MdcR near mdcA [9]. This comprehensive study, based on large-scale sequence analysis, generated a computationally derived position weight matrix that could indicate the likelihood of DNA sequence patterns to bind MdcR. We used this information to investigate how similar *A. baylyi* sequences are to the consensus. The previously derived position weight matrix
was as input for the Motif Locator software to search for pattern matches throughout the ADP1 genome [38].

A. baylyi genomic sequences with the best scores for potential MdcR-binding sites were evaluated further (Fig. S3). The ADP1 sequence with the highest score in this analysis lies upstream of mdcA (underlined text in Fig. 1). This site corresponds to the position that we previously predicted for MdcR-mediated regulation [38]. We had searched for dyad symmetry surrounding a T-N$_1$A sequence that is often conserved in LysR-type binding sites (Fig. S3). The only sequences identified by Motif Locator that had evident significance for MdcR-mediated regulation corresponded to the DNA upstream of mdcA, which is underlined in Fig. 1, and its reverse complement.

This potential binding motif, which had previously been compared and aligned in many different bacteria [9], was aligned with comparable sequences from a few additional microbes that either had not been included in the previous study or that had some feature of interest (Fig. 7a). For example, the sequence from Pseudogulbenkiana sp. NH8B was chosen because it had an atypical genetic arrangement for the mdc genes (Fig. 7b). When sequences upstream of mdcA were aligned, we discovered an additional region of conserved sequence, not previously observed, that overlaps the −35 region of P$_{mdcA}$ (Fig. 7a, highlighted in blue).

**Experimental assessment of MdcR-DNA binding predictions**

EMSA assays assessed MdcR-His$_6$-DNA binding using three overlapping 6-FAM-labelled PCR fragments in the mdcA

---

**Fig. 6.** Transcriptional analysis of mdcR. (a) Key features of the strains used in the transcriptional analysis of mdcR are depicted including the IS$_S$ insertion in srpH for ACN1713 and ACN1770 and the amplicons of ACN1770 and ACN1772 (boxed regions with amplicon copy numbers). The red and blue lines correspond to the ends of the linear DNA fragment used to generate the duplicated chromosomal regions (as depicted in Fig. 5). (b) A gene-specific primer for mdcR was used to make cDNA from RNA (dashed arrow). PCR primers can generate fragments of size 0.17 kb (F$_D$), 0.82 kb (F$_E$), 0.38 kb (F$_G$) or 0.87 kb (F$_F$). (c) Genomic DNA (gDNA) or cDNA was used as template for PCR, and primer sets could generate the fragments listed above each lane of the agarose gel. DNA molecular weight markers (2 log ladder; NEB) are in the first lane. The gel shown is representative of results seen for RT-PCR analysis of RNA from three biological replicates. (d) qRT-PCR was used to compare transcript levels of mdcR (grey bars) and srpH (white bars), both relative to the internal rpoA control. Bars are averaged data with standard error for at least three biological replicates.
The predicted MdcR-binding site is in the middle of fragment I and towards the end of fragment III but is not on fragment II (Fig. 8). MdcR-His$_{5}$-dependent retardation of DNA migration occurred with fragments I and III, but not fragment II, in the presence and absence of the effector malonate (Fig. 8b). The equilibrium dissociation constant, $K_d$, was calculated from EMSAs for MdcR-His$_{5}$ binding to fragment I in the absence of malonate ($K_d$ of 25 ±4 nM) and in the presence of malonate ($K_d$ of 54±7 nM) (Fig. 8c). An unlabelled PCR-generated fragment that contains the predicted MdcR binding site competed with the 6-FAM-labelled fragment I for binding to MdcR-His$_{5}$. Fragment I binding to MdcR-His$_{5}$ was significantly reduced in the presence of 10-fold molar excess specific competitor DNA relative to the 6-FAM-labelled fragment I, and was eliminated at 100-fold molar excess of specific competitor DNA (Fig. 8d). Specificity of the MdcR-DNA binding was confirmed with a PCR-generated DNA fragment from a different genomic region. This nonspecific competitor DNA did not reduce MdcR-His$_{5}$ binding to fragment I, even at 150-fold molar excess (Fig. 8d).

**DISCUSSION**

Although LysR-family transcriptional regulators are more prevalent than any other type of prokaryotic transcriptional regulator, many remain uncharacterized due to problems with protein solubility [19, 44, 45]. In *A. baylyi*, MdcR, a member of this family, is encoded in a convergent orientation relative to the genes it regulates. Although this convergent arrangement is common in bacteria, different mdcR configurations were identified (Fig. 7). In one example, mdcR is upstream of and divergently transcribed from its target genes, a typical organization for lysR-type genes (*Pseudogulbenkiania* sp., Fig. 7) In another case, mdcR lies upstream of and in the same orientation as mdcLM (*Polaromonas naphthalenivorans*, Fig. 7). The significance of these different genetic organizations is unclear.

LysR-type regulators recognize a T-N$_{11}$-A pattern with some dyad symmetry [46]. This pattern corresponds to interactions with two DNA-binding domains, each at the N-terminal region of one subunit of an oligomer [43]. To account for oligomeric binding of a LysR-type regulator, typically as a homotetramer, two or more ‘T-N$_{11}$-A’ patterns are usually present, sometimes with slight sequence

![Fig. 7. Conserved sequences and genetic arrangements in the vicinity of mdcA in different bacteria.](image-url)
variation [43, 47]. However, our bioinformatic analysis revealed a single pattern of this type associated with MdcR binding (TTAA-N_7-TTAA). This pattern is conserved in diverse sequences upstream of mdcA within a longer consensus sequence (Fig. 7) [9, 19]. The specific binding of MdcR-His_5 to the mdcA promoter region in EMSAs (Fig. 8) likely results from two MdcR DNA-binding domains interacting with the perfect dyad symmetry of the site (ATTAACTT-N_1-AAGTTAAT). Furthermore, we discovered a previously unobserved conserved sequence that corresponds to the P_{mdcA}^1 region of P_{mdcA} for ADP1, TT(A or G)TTGA (Fig. 7). It is possible that the DNA-binding domain of a subunit of an MdcR oligomer binds to this region. In the absence of malonate, MdcR binding to the promoter could repress P_{mdcA}. Such repression by MdcR is suggested by the decreased transcription of mdcA::lacZ from the mutated promoter, P_{mdcA51705}, when MdcR is present without its effector. Either the addition of malonate or the removal of the mdcR gene alleviates this repression (Fig. 2). A model to explain these observed effects on mdcA transcription is based on well-characterized LysR-type regulators, such as BenM [43, 47]. According to this model, malonate would bind to MdcR and result in conformational changes that shift the oligomer to a form that contacts RNA polymerase and activates transcription. While the EMSA results indicate that malonate changes the K_d of MdcR for the P_{mdcA} region, the data reported here cannot answer questions concerning specific interactions between oligomeric forms of MdcR with DNA or RNA polymerase.

**Operon organization for the mdcA, B, C, D, E, G, H, L and M genes**

We demonstrated that MdcR activates transcription of mdcA and that malonate, in the absence of catabolism, serves as an effector for MdcR (Fig. 2). While spontaneous regulatory mutants can sometimes be selected [48], isolation of a mutation in the −10 region of P_{mdcA} required...
mutagenesis. Nevertheless, the PmdcA1705 mutation helped clarify regulation by alleviating the requirement of MdcR and malonate for high-level mdcA transcription (Figs 2, 3b). This mutation also indicated that mdcA and mdcL are co-expressed from PmdcA because it caused high-level mdcL transcription (Fig. 3b).

Another promoter was discovered when mdcL was expressed despite the termination of transcription from PmdcA. RT-PCR localized a second transcriptional start site to a region within 322 nt upstream of the mdcL coding sequence (Fig. 4). A single likely promoter in this region was predicted for transcription initiation within the 47 nt region separating the coding sequences of mdcH and mdcL. The software used for this prediction, BPROM, is most accurate within a localized region near the transcriptional start site and is estimated to be 80% accurate for sigma 70 type promoters [39]. This software exactly predicted the transcriptional start site that we identified by 5'-RACE for mdcA (Fig. 1).

Transcription from this second promoter could allow the MdcLM transport proteins to be synthesized when cells are grown under non-inducing conditions. Thus, malonate could enter the cell to increase transcription from PmdcA. Transcription of mdcLM distinct from mdcA occurs in other bacteria. As shown for Polaramonas (Fig. 7), the convergent orientation of mdcL to other mdc genes requires separate regulation of genes for transport and catabolism.

**Operon organization for mdcR and srpH**

The proximity of mdcR and srpH, separated by 38 nt, suggested their cotranscription (Fig. S2). However, after termination of transcription from PspH by interposon insertion, the resulting strain (ACN1713) remained able to use malonate as sole carbon source, which requires mdcR transcription. RT-PCR analysis confirmed a transcription start site for mdcR within srpH, downstream of the IS8 insertion (Fig. 6c). However, mdcR was also expressed from a promoter upstream of IS8 as part of an srpH-mdcR operon. qRT-PCR indicated that most of mdcR gene expression results from co-transcription with srpH (Fig. 6d). Annotation of this srpH gene (ACIAD1763) suggests it encodes an acetyltransferase, although this possibility has not been investigated. The significance of an srpH-mdcR operon is not evident, and database searches failed to reveal similar gene organization in any other microbe.

The transcription of mdcR differs between A. baylyi and K. pneumoniae CG43-17, wherein MdcR was shown to bind upstream of mdcR to a palindromic sequence that does not match the consensus binding sequence upstream of mdcA [11]. We found no similar palindromic upstream of mdcR or srpH in ADP1, and our attempts to detect specific binding of MdcR to these regions (or near mdcL) by EMSA were unsuccessful (data not shown). Strong matches to the consensus binding sequence upstream of mdcA were not detected elsewhere in the mdc region or elsewhere in the ADP1 chromosome where MdcR-binding would be likely. Many LysR-type proteins are negatively auto-regulated, often resulting from protein–DNA interactions near a promoter for the regulatory gene and its divergent target. In A. baylyi, the bioinformatic and EMSA analyses provide no evidence that MdcR directly regulates its own transcription.

**Increased chromosomal gene dosage: a novel method for enhanced transcriptional analysis**

To improve detection of low-level transcripts, we generated tandem arrays of chromosomal segments by using our previously developed methods [17, 18]. We increased transcript levels of mdcL and mdcR while maintaining the native chromosomal promoters and genomic context. Strains were selected with approximately 20–30 copies of the chromosomal mdc region. Transcriptional analysis of these amplified strains by RT-PCR methods yielded the same conclusions as studies of their single-copy counterparts (data not shown). However, the higher chromosomal gene dosage resulted in RT-PCR products that were more abundant, easier to detect and gave greater confidence that the PCR products accurately reflected cDNA structure. The similar application of this method to other weakly expressed genes in A. baylyi should facilitate transcriptional characterization. This approach is applicable to other bacteria, although modifications might be needed to address microbial differences in transformation and recombination capabilities.

**Multiple promoters for mdc genes**

Despite the increased application of high-throughput techniques for comprehensive transcriptome analysis in prokaryotes, it remains difficult to predict a suboperon [49]. This term describes mdcLM, wherein an internal transcriptional start site allows a subset of genes within an operon to be differentially expressed. Experimental data, as we provide for mdcL and mdcR, contribute to improved understanding of the complexity of operon organization in bacteria [50]. These multiple promoter regions are depicted in Fig. S3(c). Furthermore, experimental data for specific gene sets, such as mdc gene clusters, contribute to gene-expression databases and to improved accuracy of computational methods for operon prediction.

**Concluding remarks**

These studies of MdcR in A. baylyi ADP1 provide the foundation for in-depth structure–function investigations of LysR-type transcriptional regulators. There is no single regulatory model that can accurately describe all members of this family [45]. Atypical aspects of MdcR-mediated regulation raise questions about the interactions of this regulator with RNA polymerase and with DNA. As demonstrated here, the protein can be purified from A. baylyi and we can now investigate its oligomeric structure and how its mode of transcriptional activation reflects the atypical sequence patterns upstream of mdcA.

**Funding information**

We acknowledge the following funding sources: NSF grants to E.L.N. (MCB-1361188, DEB-1556541, and MCB-1615365) supported A.L.S.,
M. P. T. V., N. S. L. and S. W. W.; NSF grant to A. C. K. (MCB-1051175); NSF grant to C. M. and E. L. N. (I02-1024108) supported N. R. G., M. N. and M. W. University of Georgia Franklin College of Arts and Sciences and Department of Microbiology provided laboratory course funding and support for J. L. S., K. T. E., M. P. T. and N. S. L.; College of New Jersey School of Science and Department of Biology support for K. T. E.

Acknowledgements

We acknowledge the contributions to this work by undergraduate students in MIBO4600L: Laura Alexander, Brittany Babin, Lindsey Cheek, Stephanie Elam, Glen Feller, Alex Howell, Courtney Matheny, Shaleen Ragha, Spencer Short, Park Skelton, Windy Soenarto, Cameron Wagenbrenner, Erena Weathers and Leah Williams. These students helped to create the mdcR and mdcA mutant and reporter strains of A. baylyi, performed the initial experiments to map the start site for the mdcA operon transcript and determine operon structure for mdcLM and mdcR-srpH, and performed the first EMSAs with MdcR-His<sub>6</sub>. We thank Cassandra Bartlett for assistance with construction and confirmation of A. baylyi mutant strains. We appreciate Jennifer Tsang’s teaching assistance. We thank Jan Mazek for his assistance with the bioinformatic analysis of the A. baylyi genome for MdcR binding sites.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References


Edited by: R. E. Parales and W. Achouak

Five reasons to publish your next article with a Microbiology Society journal

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as ‘excellent’ or ‘very good’.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at microbiologyresearch.org.