1518–1523 Nucleic Acids Research, 1995, Vol. 23, No. 9 

# The *Escherichia coli* MelR transcription activator: production of a stable fragment containing the DNA-binding domain

**Carmen M. Michán+, Stephen J. W. Busby and Eva I. Hyde***

School of Biochemistry, University of Birmingham, PO Box 363, Birmingham B15 2TT, UK

Received January 23, 1995; Revised and Accepted March 17, 1995

## ABSTRACT

**A set of nested deletions has been made in the *Escherichia coli melR* gene, encoding the *MelR* transcription activator protein. Expression of the resulting *melR* derivatives led to the production of nine MelR proteins with N-terminal deletions of different lengths. The properties of the shortened proteins have been studied both *in vivo* and *in vitro*. None of the truncated proteins activate transcription from the *E.coli melAB* promoter but three; MelR220, MelR183 and MelR173, inhibit activation of the *melAB* promoter by chromosomally-encoded full-length MelR. In gel retardation assays, both MelR183 and MelR173 clearly retard DNA fragments carrying the *melAB* promoter. MelR173 has been overproduced in a T7 expression system and shown to be stable *in vivo* for up to 2 h. DNAase I footprinting assays of partially purified protein show that it binds to the *melAB* promoter, protecting the same sites as the full-length protein. This fragment may be suitable for further structure/function studies of this class of transcription activator.**

## INTRODUCTION

The *Escherichia coli* MelR protein is a transcription activator that regulates the pathway for the degradation of the disaccharide melibiose (1,2). *In vivo*, in the presence of the effector melibiose, MelR activates transcription from the *mel*AB promoter (p*mel*AB), that controls expression of an α-galactosidase (encoded by *mel*A) and the melibiose transporter (encoded by *mel*B) (2). *In vitro*, MelR binds to two identical 18 bp sequences at p*mel*AB. These sequences are organised as an inverted repeat separated by 20 bp, located between bp –54 and –110 with respect to the start of transcription (2–4) (Fig. 1).

MelR belongs to the AraC/XylS family of prokaryote transcription regulators that includes at least 30 other proteins, most of them activators (5–7). The members of this family are most homologous in their C-terminal regions where two DNA-binding helix–turn–helix structures have been predicted (at positions 210–229 and 259–278 for MelR). It has been proposed that the N-terminal region is used in effector binding and dimerisation, and that the C-terminal region has sequence-specific DNA binding properties (6,7). This has been corroborated by mutational analysis of different members of the family, and from *in vitro* experiments with the AraC protein of *E.coli* (8–13). However, to date, there is no structural data for any member of the family.

We have previously expressed full length MelR from the $\lambda P_L$ and $P_R$ promoters (4) but the construct was not stable *in vivo* and the full length protein was poorly soluble, as other proteins in this family. The aim of this study was to produce a shortened MelR derivative that could recognise the DNA binding sites for MelR at p*mel*AB and could be stably overexpressed for further structure/function studies *in vitro*. In our previous work on MelR, we showed that the deletion of 16 or more amino acids from the C-terminus of MelR abolished DNA binding and, that MelR activity was affected by substitutions in the putative helix–turn–helix regions (2,4,14,15). Moreover we found that a MelR derivative containing 143 C-terminal amino acids bound weakly to the *mel*AB promoter, but was not expressed in sufficient quantity to be detected by SDS–PAGE (4). We therefore constructed a series of N-terminal deletions of the *MelR* gene and tested their DNA-binding properties *in vivo* and *in vitro* to find a short fragment that was stably expressed and still bound specifically to the *mel*AB promoter.

## MATERIALS AND METHODS

### Strains and plasmids

The *E.coli* K12 strains used in this work were M182 (Δ*lac mel*) (16), POP2094 (Δ*lac mel*$^+$) (17) and BL21(λDE3) (T7 RNA pol$^+$ F$^-$ *ompT* r$^{B-}$ m$^{B-}$) (18). Table 1 lists the different plasmids used.

### Recombinant DNA procedures

Standard protocols for DNA manipulation were employed (21).

The *Nco*I–*Bam*HI fragment carrying the *mel*R gene in pRC1, described by Caswell and co-workers (4), was cloned into the T7 expression vector pET9d (19) to give pCM117-303.

pCM117-314 was obtained by cloning the linker:

```
    NcoI PstI                    BstEII
5i-CATGGCTGCAGACACTGATACATTTGGTCACCT-3i
       CGACGTCTGTGACTATGTAAACCAGTGGAGTAC
```

* To whom correspondence should be addressed

+Present address: CSIC Estacion Experimental del Zaidin, C. Prof. Albareda 1, 18080, Granada, Spain

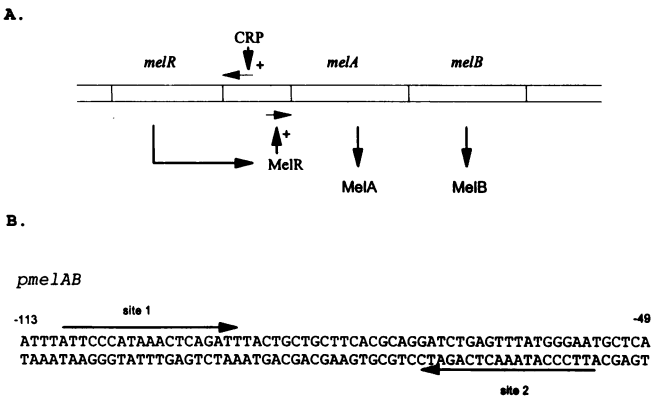


**Figure 1.** The *E.coli* melibiose operon. (**A**) Organization of the mel genes in the *E.coli* chromosome. The plus symbols represent positive regulation. Short horizontal arrows show the direction of transcription. (**B**) Sequence of the MelR-binding region at the melAB promoter. The position of the perfect inverted repeat is indicated by the arrows.

**Table 1.** Plasmids used in this work

| Plasmid | Description | Reference |
|---|---|---|
| pAA121-KK43 | pAA121 derivative with the *pmelAB* sequences from -136 to +21 with a GC to AT transversion at -73, creating a unique *BglII* site between MelR binding sites 1 and 2 cloned as a *EcoRI-HindIII* fragment | 15 |
| pRW5-KK43 | Idem but cloned into the broad host range *lac* expression vector pRW5 | 15 |
| pLysS | pACYC184 derivative which supplies low levels of T7 lysozyme | 19 |
| pET9d | Vector designed to allow the cloning of genes under the control of T7 transcription and translation signals | 19 |
| pRC1 | pJLA502 derivative containing the *melR* gene with an extra alanine in the second position cloned as a *NcoI-BamHI* fragment. This *melR* derivative is named *melR303* in this paper | 4 |
| pCM117-303 | pET9d derivative containing *melR303* from pRC1 as an *NcoI-BamHI* fragment | This work |
| pCM117-314 | Derivative of pCM117-303 with an *NcoI-PstI-BstEII* linker cloned into the *NcoI* site; this construction encodes a MelR derivative with 12 extra amino acids at its N-terminal end (*melR314*) | This work |
| pCM117-220 | Derivative of pCM117-314 deleted so that it encodes a truncated MelR with the C-terminal 220 amino acids under the control of the T7 promoter (*melR220*) | This work |
| pCM117-183 | Idem with the C-terminal 183 aa of MelR (*melR183*) | This work |
| pCM117-173 | Idem with 173 aa (*melR173*) | This work |
| pCM117-147 | Idem with 147 aa (*melR147*) | This work |
| pCM117-120 | Idem with 120 aa (*melR120*) | This work |
| pCM117-96 | Idem with 96 aa (*melR96*) | This work |
| pCM117-95 | Idem with 95 aa (*melR95*) | This work |
| pCM117-90 | Idem with 90 aa (*melR90*) | This work |
| pCM117-45 | Idem with 45 aa (*melR45*) | This work |
| pHH122 | Vector derived from pAA121 for *galE* fusions carrying a *HindIII-SphI-BamHI-SmaI* linker | 20 |
| pCM118 | pHH122 derivative with an *NcoI* linker inserted into the *HindIII* site | This work |
| pCM118-303 | pCM118 derivative with *melR303* cloned between the NcoI and BamHI sites giving a *galE* fusion under the *gal* promoter | This work |
| pCM118-314 | Idem with *melR314* | This work |
| pCM118-220 | Idem with *melR220* | This work |
| pCM118-183 | Idem with *melR183* | This work |
| pCM118-173 | Idem with *melR173* | This work |
| pCM118-147 | Idem with *melR147* | This work |
| pCM118-120 | Idem with *melR120* | This work |
| pCM118-96 | Idem with *melR96* | This work |
| pCM118-95 | Idem with *melR95* | This work |
| pCM118-90 | Idem with *melR90* | This work |
| pCM118-45 | Idem with *melR45* | This work |

into the *Nco*I site of pCM117-303. The orientation of the linker with the *Bst*EII site adjacent to the *mel*R coding region was used. This plasmid, pCM117-314, encodes a MelR derivative (MelR314) that has the sequence: Met-Ala-Ala-Asp-Thr-Asp-Thr-Phe-Gly-His-Leu added to the N-terminus of MelR303. Nested deletions of the MelR-coding region were performed using the Promega 'Erase-a-Base' system. pCM117-314 was digested with *Pst*I and *Bst*EII, and Exonuclease III was used to remove DNA unidirectionally from the 5′ protruding end (*Bst*EII), while protecting the adjacent *Nco*I site with the 3′-overhanging end (*Pst*I) (22). After treatment with S1 nuclease, Klenow enzyme and T4 DNA ligase, the DNA was transformed into *E.coli* M182.

pCM118 was constructed by cutting pHH122 (20) with *Hin*dIII, filling the *Hin*dIII ends with Klenow enzyme, and inserting an *Nco*I linker between the blunt ends. The resulting vector gave in-phase *gal*E–*mel*R fusions when full length *mel*R and truncated derivatives were introduced as *Nco*I–*Bam*HI fragments. The resulting fusions encode proteins with the GalE N-terminal four amino acids plus a seven amino acid linker, added to the N-terminus of the truncated MelR derivatives (i.e. Met-Arg-Val-Leu-Val-Thr-Ala-Cys-Glu-Ala-Ala-MelR).

### *In vivo* assay of p*mel*AB activity modulated by truncated MelR proteins

*E.coli* strain POP2094 containing the plasmid pRW5-KK43 was transformed with the different *mel*R deletions cloned into pCM118 (pCM118-303, pCM118-314, pCM118-220, etc.). The resulting transformants were grown in minimal M9 medium ± melibiose, and β-galactosidase assays were performed as described by Caswell and co-workers (4).

### Preparation of crude extracts

*E.coli* M182 was transformed with plasmids containing the *mel*R deletions cloned into pCM118, and colonies were grown exponentially in LB media at 37°C. Three ml of each culture were harvested by centrifugation, cells were resuspended in 100 mM $K_2HPO_4$ pH 7.4, 50 mM KCl, 10% (v/v) glycerol, 1 mM EDTA and 1 mM DTT, and disrupted by sonication. Eight µl (10 mg/ml) PMSF was added and the samples were centrifuged for 10 min in a microfuge at 4°C. Aliquots of the supernatant (100 µl) were mixed with 40 µl 50% (v/v) glycerol and samples were stored at 4°C. The DNA-binding activity of samples prepared this way was stable for ≥ 10 days.

### Gel retardation assays

Labelling of the DNA fragments and gel retardation assays were performed as previously described (14,15). The concentration of DNA carrying p*mel*AB was typically 1 nM. Two different fragments were used: either KK43 *Eco*RI–*Hin*dIII that contains both MelR binding sites, or KK43 *Eco*RI–*Bgl*II with only one (15). The cell extracts and the DNA were incubated together at room temperature for 10 min. Samples were loaded under tension onto 7.5% polyacrylamide gels and electrophoresed at 12 V/cm in TBE buffer. Bands were detected by autoradiography.

### Pulse-labelling of proteins with [$^{35}$S]methionine

BL21(λDE3) [pLysS] transformed with the different pCM117 derivatives was grown at 37°C in M9 minimal medium supplemented with 1 ml/l of *E.coli* sulphur-free salts (23) until an $OD_{650}$ of 0.4 was reached. The synthesis of T7 RNA polymerase was then induced by the addition of 0.4 mM IPTG, and rifampicin

was added to 200 μg/ml to inhibit the *E.coli* RNA polymerase. After a further 2 h incubation, 200 μl aliquots were taken, gently mixed with 1 μl (13.3 μCi) of [$^{35}$S]methionine (specific activity 40–500 mCi/mmol) and incubated at 37°C for 10 min. In pulse–chase experiments 12.5 μl 0.4 mM unlabelled methionine was added and the sample was incubated for a further 30, 60 or 120 min. Cells were collected by centrifugation and kept at –20°C, prior to analysis on 12.5% SDS–PAGE gels and autoradiography, using standard protocols (24,25).

### Protein purification

Wild type MelR was purified as previously described (4). To purify MelR173, BL21(λDE3) [pLysS pCM117-173] was grown in LB at 37°C with aeration until exponential phase was reached. Cells were then induced with lactose (2 g/l) for 5 h and collected by centrifugation. Purification was initially performed as described by Caswell and co-workers (4) and MelR173 was assayed by gel retardation with DNA fragments containing p*mel*AB. Unexpectedly, unlike wild type MelR, MelR173 did not bind to phosphocellulose. The active fractions were pooled, and dialysed into 50 mM Tris–HCl pH 6.8, 100 mM NaCl, 0.1 mM EDTA and 1 ml/l β-mercaptoethanol. The dialysed preparation was loaded onto a DEAE–Sephadex column, and was eluted with a 0.1–2 M NaCl gradient. MelR eluted at ~300 mM NaCl. The active fractions were combined, giving a pool containing 1 mg/ml protein, where ~5% was MelR173 (estimated by the intensity of the corresponding band in a Coomassie stained PAGE). Glycerol was added to 25% and the preparation was stored at –20°C. Prior to footprinting assays, the semi-purified MelR173 sample was concentrated 5-fold by centrifugation using an Amicon Centricon 10 microconcentrator.

### DNAase I footprinting assays

The KK43 *Eco*RI–*Hin*dIII fragment containing p*mel*AB (15) was typically 1 nM, and was 5′-end labelled at either end. The conditions used were as described by Spassky and co-workers (26), but $MgCl_2$ was only added with the DNAase I enzyme. Salmon sperm DNA was added as indicated to compete with non-specific DNA-binding. The protein and DNA were incubated at room temperature for 10 min before the addition of DNAase I.

## RESULTS AND DISCUSSION

### Deletions of the N-terminal end of MelR

The starting point for this work was the *Nco*I–*Bam*HI fragment carrying the *mel*R gene described by Caswell and co-workers (4). This fragment has the initial ATG of the *mel*R open reading frame incorporated into the *Nco*I site, and encodes full-length wild type MelR with an extra alanine in the second position (MelR303). The *Bam*HI site is just downstream of the MelR termination codon. To control the expression of MelR303 stringently, this *Nco*I–*Bam*HI fragment was cloned into the T7 expression vector pET9d (19) to give pCM117-303. In this construction, *mel*R303 is under the control of a T7 promoter, and is expressed only in the presence of T7 RNA polymerase.

To make deletions in the melR gene, we exploited Henikoff's method, in which Exonuclease III is used to digest insert DNA from a 5′ protruding- or blunt-end restriction site (22). An

```
MNTDTFMCSS DEKQTRSPLS LYSEYQRMEI EFRAPHIMPT SHWHGQVEVN
                                                MelR220
VPFDGDVEYL INNEKVNINQ GHITLFWACT PHQLTDTGTC QSMAIFNLPM
                          MelR183       MelR173
HLFLSWPLDK DLINHVTHGM VIKSLATQQL SPFEVRRWQQ ELNSPNEQIR
        MelR147                               MelR120
QLAIDEIGLM LKRFSLSGWE PILVNKTSRT HKNSVSRHAQ FYVSQMLGFI
          MelR96 MelR90
AENYDQAITI NDVAEHVKLN ANYAMGIFQR VMQLTMKQYI TAMRINHVRA
                MelR95
LLSDTDKSIL DIALTAGFRS SSRFYSTFGK YVGMSPQQYR KLSQQRRQTF
          MelR45
PG
```

**Figure 2.** Sequences of the truncated MelR derivatives. The amino acid sequence of wild type MelR is shown in single letter code. The two predicted DNA-binding helix–turn–helix motifs are highlighted in bold characters. All the MelR derivatives have a short N-terminal linker (MA- or MRVLVTA-CEAA- in pCM117 or pCM118 derivatives, respectively), followed by the C-terminal part of MelR starting at the positions indicated by the arrows.

*Nco*I–*Pst*I–*Bst*EII linker was introduced into the *Nco*I site of pCM117-303 to produce pCM117-314, that encodes a MelR derivative with 12 extra amino acids at its N-terminus, MelR314. This plasmid was linearised with *Pst*I and *Bst*EII, and deletions were made as described in Materials and Methods. Thirty-two derivatives of pCM117-314, containing truncated *mel*R, were selected and sequenced. Nine of the deletions contained segments of *mel*R in phase with the translation start. Each of these plasmids encodes a different MelR C-terminal fragment under the control of the T7 promoter, with deletions from 33–85% of the full-length protein (Fig. 2). The MelR derivatives encoded by each plasmid are designated by the number of amino acids.

### *In vivo* effect of MelR derivatives on p*mel*AB

In order to determine the effects of the N-terminal deletions in MelR *in vivo*, we constructed plasmid pCM118, a vector that contains the *E.coli gal* operon promoter region and the *gal*E translation start, with unique *Nco*I and *Bam*HI sites just downstream of the *gal*E start codon. The different *Nco*I–*Bam*HI fragments carrying truncated *mel*R were transferred from the T7 vector into pCM118. This resulted in a series of *gal*E–*mel*R fusions expressed from the constitutive *gal*P2 promoter. These constructions express the different truncated MelR proteins with 11 extra amino acids at the N-terminus (the N-terminus of GalE plus a short linker).

In preliminary experiments in *E.coli* strain M182, we failed to detect activation of p*mel*AB with any of the truncated MelR proteins, suggesting that even the shortest deletion of 82 residues at the N-terminus of the protein, suppressed the activation function of MelR. We reasoned that if the truncated proteins could fold correctly, they might interfere with activation of p*mel*AB by full-length wild type MelR. To test this, we used the reporter plasmid pRW5-KK43, that has the *lacZ* gene under the control of p*mel*AB (pRW5 is a broad host range *lac* expression vector, and KK43 refers to a fragment carrying the *mel*AB promoter region; see Table 1). *E.coli* strain POP2094, containing wild type *mel*R in its chromosome, was transformed with pRW5-KK43 and the different pCM118-*mel*R derivatives. Transformants were grown in minimal medium in the presence or absence of melibiose and the resulting β-galactosidase activities were measured.

The results in Table 2 show that, in the absence of melibiose, none of the MelR derivatives have any effect on expression from

p*mel*AB. Melibiose triggers a 5-fold induction of p*mel*AB in POP2094 due to the chromosomally-encoded *mel*R protein. This induction is doubled by the introduction of plasmids encoding MelR303 and MelR314. However, induction was reduced by the plasmids encoding MelR220, MelR183 and MelR173, the most suppression being found with MelR173. In contrast, plasmids encoding shorter segments of MelR (MelR147, MelR120, MelR96, MelR95, MelR90 and MelR45) had only small effects on induction of p*mel*AB by the chromosomal *mel*R.

**Table 2.** *In vivo* effects of the truncated MelR on transcription from p*mel*AB

| | β-galactosidase activity (%) | |
|---|---|---|
| | – melibiose | + melibiose |
| Control cells (no added *mel*R) | 21 (±9) | 100 |
| *mel*R303 | 26 (±15) | 215 (±40) |
| *mel*R314 | 19 (±12) | 257 (±20) |
| *mel*R220 | 22 (±12) | 47 (±25) |
| *mel*R183 | 28 (±5) | 58 (±25) |
| *mel*R173 | 14 (±3) | 17 (±7) |
| *mel*R147 | 17 (±5) | 83 (±13) |
| *mel*R120 | 15 (±5) | 81 (±4) |
| *mel*R96 | 19 (±6) | 86 (±5) |
| *mel*R95 | 15 (±5) | 77 (±5) |
| *mel*R90 | 20 (±12) | 83 (±10) |
| *mel*R45 | 17 (±7) | 79 (±9) |

The activation/repression effect of the different melR derivatives was monitored by β-galactosidase assays on a p*mel*AB::*lac*Z fusion. The results are the mean of three independent experiments, which were standardised using as 100% the values obtained from the POP2094 cells only containing chromosomal *mel*R, grown in the presence of melibiose (~550 Miller units).

## *In vitro* binding of MelR derivatives to p*mel*AB

To test whether the truncated MelR proteins could recognise the MelR binding sites at p*mel*AB, gel retardation assays were performed *in vitro*, using crude extracts of *E.coli* M182 containing pCM118 derivatives encoding full-length or truncated MelR derivatives. Figure 3 shows a gel retardation experiment performed with the KK43 *Eco*RI–*Bgl*II p*mel*AB fragment that contains one MelR-binding site (15). As expected, extracts containing MelR303 and MelR314 gave clear retardation of this fragment. Of the truncated MelR derivatives, clear retarded bands were seen only with MelR183 and MelR173. Faint retarded bands were found with increased amounts of extracts containing MelR220 and MelR147, but only non-specific smearing was found with extracts from cells with pCM118 encoding shorter MelR derivatives (not shown). This shows that MelR183 and MelR173 bind to p*mel*AB and suggests that the *in vivo* effects on MelR-dependent activation of p*mel*AB observed with MelR220, MelR183 and MelR173 are due to competition between these proteins and the wild-type protein for the MelR sites on the operator. The effects of these proteins will depend both on the amount of protein in the cell and on the affinity of the derivative for the DNA. The shorter proteins, which show no effect either *in*

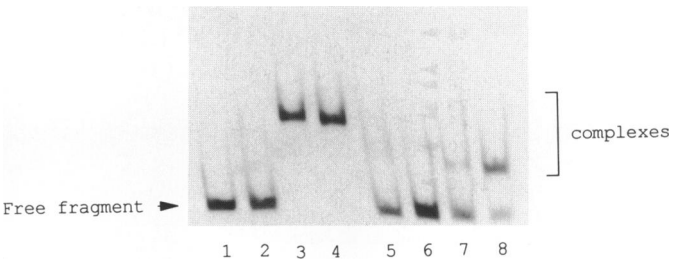


**Figure 3.** Binding of MelR derivatives to p*mel*AB measured by gel retardation assays. Labelled KK43 *Eco*RI–*Bgl*II fragment, containing MelR binding site 1 from pmelAB was incubated with protein extracts from M182 carrying different plasmids. Lane 1, labelled DNA only; lane 2, cell extract from M182 with pCM118 (vector only); lane 3, pCM118-303 (MelR); lane 4, pCM118-314 (MelR314); lane 5, pCM118-220 (MelR220); lane 6, pCM118-147 (MelR147); lane 7, pCM118-183 (MelR183); lane 8, pCM118-173 (MelR173). In each case 5 µl of cell extracts was used in a final reaction volume of 15 µl.

*vitro* or *in vivo*, are either not made in sufficient quantity, unstable, or do not compete efficiently for DNA binding.

To try to distinguish between these possibilities, the crude cell extracts were electrophoresed by SDS–PAGE and stained with Coomassie. However, the MelR proteins were produced in low amounts and are of similar molecular mass to other cell proteins and so could not be quantified. At present we do not have antibodies to MelR and so could not determine the amount of protein by Western blots; however, depending on the position of the major epitopes recognised by the antibodies, even this method may not be accurate for truncated derivatives.

## *In vivo* production and stability of MelR derivatives

In order to produce high levels of the derivatives, under stringent control, we used the pCM117 derivatives, encoding full-length or truncated MelR from a T7 promoter, transformed into *E.coli* strain BL21(λDE3) [pLysS]. To investigate the expression of different MelR derivatives by this system, the production and the size of the truncated proteins was studied by pulse labelling. The synthesis of T7 RNA polymerase was induced by IPTG and rifampicin was added to inhibit the host *E.coli* RNA polymerase, prior to pulse labelling with [$^{35}$S]methionine. All the MelR derivatives except MelR314 and MelR45 were clearly detected by autoradiography after SDS–PAGE and appeared to be the correct size. Results in Figure 4 show that the expression of different MelR derivatives varies greatly. Surprisingly, expression of MelR303, MelR314 and MelR220 was poor whilst substantially higher levels of MelR183, MelR173, MelR96, MelR95 and MelR90 were found. MelR147 and MelR120 were expressed at a lower level than MelR173 (not shown).

To investigate the reasons for the low levels of longer MelR derivatives compared to MelR173, a pulse chase experiment was performed. The amount of [$^{35}$S]methionine incorporated into MelR303 and MelR173 after induction of the T7 system was measured at different times after addition of an excess of unlabelled methionine. The results showed that both MelR303 and MelR173 are stable for up to 2 h, suggesting that, though poorly expressed, the longer MelR derivatives are stable. The experiment was repeated with two other MelR derivatives, MelR147 and MelR90; interestingly, pulse-labelled MelR147

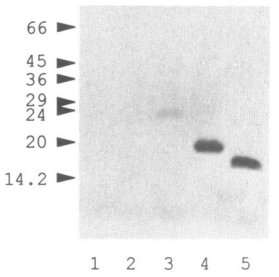


**Figure 4.** Expression of MelR and truncated derivatives. BL21(DE3) [pLysS] cells carrying different plasmids were selectively labelled with [$^{35}$S]methionine for 10 min after induction of T7 RNA polymerase (see Materials and Methods). The figure shows an autoradiogram of a 12.5% polyacrylamide gel after electrophoresis of the labelled proteins. The calibration was made with a mixture of proteins electrophoresed on the same gel and stained with Coomassie brilliant blue R250. The lanes correspond to cells containing: 1, pCM117-303; 2, pCM117-314; 3, pCM117-220; 4, pCM117-183; 5, pCM117-173.

was not detected after a 30 min chase with unlabelled methionine, whilst pulse-labelled MelR90 was stable.

## *In vitro* properties of MelR173

Since MelR173 could be stably expressed to a high level and gave a clear retarded band in gel retardation assays, we investigated its DNA-binding properties further. Figure 5A shows a comparison of a titration of a crude extract containing MelR173 (encoded by pCM118 in *E.coli* strain M182) with that containing full length MelR303, with the KK43 *Eco*RI–*Bgl*II fragment carrying a single MelR-binding site. MelR173 gives a single retarded band with this fragment. In contrast, at low concentrations of extract, MelR303 gives two retarded bands but at higher protein concentrations only the less mobile band is observed. Figure 5B shows similar titrations with the KK43 *Eco*RI–*Hin*dIII carrying both p*mel*AB MelR-binding sites. With this fragment, up to four different retarded bands are observed with MelR303, as found previously (4). The most mobile band was observed only at low concentrations of cell extract whereas the least mobile band was observed at high concentrations. With MelR173, only two retarded bands were seen over a wide range of concentrations of cell extract.

To investigate whether MelR173 was binding to the same sequences at p*mel*AB as full length MelR, DNAase I footprinting assays were performed, using the KK43 *Eco*RI–*Hin*dIII fragment carrying both MelR-binding sites. The source of protein was either 90% purified wild type MelR (MelR303) or the semipurified preparation of MelR173 (5% MelR173) from the T7 expression system described in Materials and Methods. Both MelR303 and MelR173 give clear protection of the fragment in the region of the two MelR-binding sites and protect the same bands, although some small differences in the relative intensities of bands are observed (Fig. 6). To obtain a clear footprint, a higher concentration of MelR173 than of MelR303 was needed. From titrations (not shown), we estimate that ~1 µM MelR173 is required for 50% protection in our conditions whilst 50% protection is afforded by 20–40 nM MelR303.

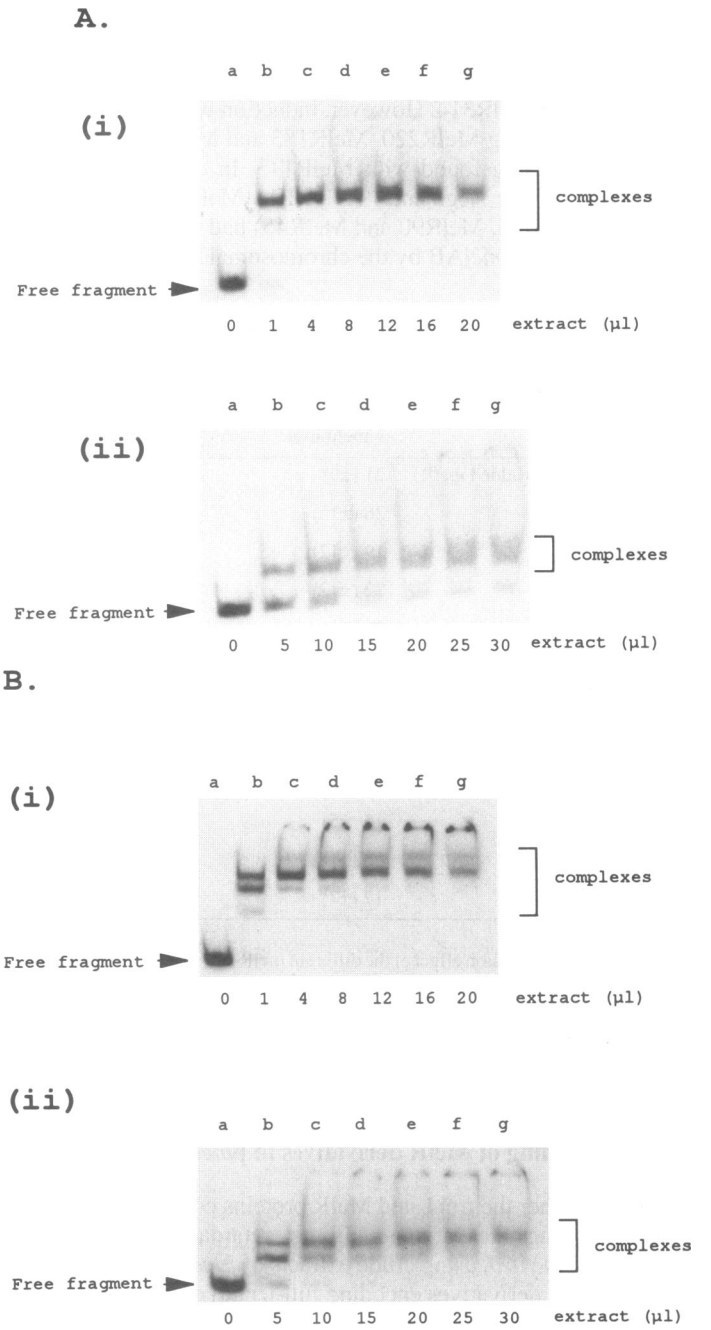


**Figure 5.** Titrations of p*mel*AB with cell extracts containing MelR303 and MelR173 using gel retardation assays. **(A)** Labelled *Eco*RI–*Bgl*II KK43 fragment containing site 1 of p*mel*AB was incubated with increasing amounts of cell extracts from (i) M182 pCM118-303 or (ii) pCM118-173. **(B)** Labelled *Eco*RI–*Hin*dIII KK43 fragment containing both MelR-binding sites 1 and 2 of p*mel*AB was incubated with increasing amounts of cell extracts from, (i) M182 pCM118-303 or (ii) pCM118-173. Different volumes of cell extract, as indicated in the figure, were incubated in a final reaction volume of 40 µl.

## Conclusions

Comparison of MelR with other proteins of the AraC/XylS family, suggests that the C-terminal half consists of a domain that is involved in specific binding to the *mel*AB promoter (5–7). This is consistent with the results of recent molecular genetic studies

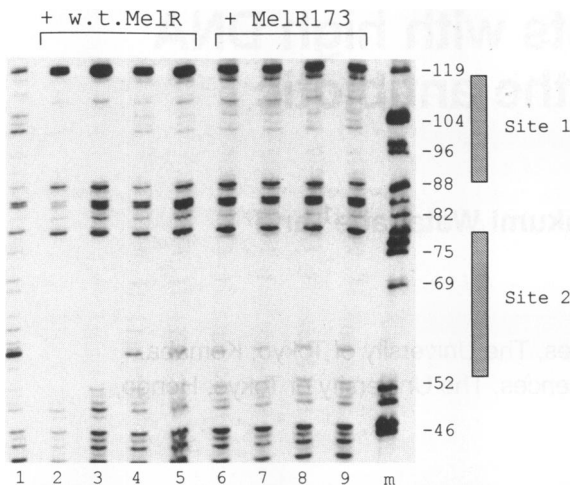


**Figure 6.** DNAase I footprinting of bound MelR303 and MelR173 at p*mel*AB. Labelled *Eco*RI–*Hin*dIII KK43 fragment containing both MelR-binding sites 1 and 2 of p*mel*AB was incubated with: no added protein (lane 1), wild type MelR (lanes 2–5) and MelR173 (lanes 6–9), with increasing concentrations of salmon sperm DNA to remove non-specific binding: 0 (lanes 2 and 6), 5 (lanes 3 and 7), 10 (Lanes 4 and 8) and 50 (lanes 5 and 9) ng/µl. Samples were treated with DNAase I and electrophoresis was performed on a 7.5% sequencing gel. Lane m is a G reaction ladder showing the positions of the bands relative to the p*mel*AB transcription start point.

of MelR (14,15) and other members of the AraC/XylS family (10,12,27–29). Here, we tested this directly, making a series of MelR deletion derivatives, to investigate whether a stable C-terminal DNA-binding domain could be isolated. Nine different N-terminally-truncated MelR derivatives were produced and their ability to interact with p*mel*AB, both *in vivo* and *in vitro*, was studied. Our results show that two of the derivatives, MelR183 and MelR173 bind to p*mel*AB, giving clear DNA bandshifts in gel retardation assays. Some binding was found with MelR147, consistent with previous studies of a 143 amino acid fragment (4), but no DNA binding was detected with any of the shorter MelR derivatives. DNAase I footprinting shows that MelR173 and full-length MelR occupy the same two sites at p*mel*AB, although the shortened MelR derivative binds with a ~40-fold lesser affinity. Our results provide important evidence that the C-terminal ~50% of MelR does fold independently forming a DNA-binding domain. Moreover, MelR173 can be stably expressed and may be suitable for more detailed structure/function studies.

A number of other points emerge from our study. First, the expression of the truncated MelR derivatives in the T7 system differ markedly from one case to another. Interestingly, the longer MelR derivatives (including the starting wild type) were expressed at lower levels than the shortened derivatives, although full length MelR is stable in this system. The low level of the longer derivatives may be due either to poor transcription, instability of the mRNA or to poor translation. Preliminary analysis of the DNA sequence shows no MelR site in the gene that would lead to autoregulation, and the RNA shows no elements of secondary structure that might attenuate translation. Further experiments are needed to examine the amount of mRNA and its stability.

Secondly, none of the truncated MelR derivatives were competent to activate transcription at p*mel*AB, suggesting that the N-terminal region carries functions for activation. The shorter deletions may interfere with inducer binding or subsequent triggering by melibiose. Comparison of p*mel*AB binding by full-length MelR and MelR173 (Fig. 5) suggests that interactions between bound MelR molecules are also suppressed by N-terminal deletion, as multiple retarded bands are observed with the full-length MelR but not with the truncated derivative. These multiple bands are observed with purified MelR and are assumed to come from two protein molecules binding to a single site, due to protein–protein interactions. The reduced affinity of the truncated MelR for DNA may be in part due to loss of these cooperative contacts. A similar situation is found with AraC, where the dimerisation motif was located in its N-terminal region (12). However, in contrast to MelR, N-terminal deletions in AraC give truncated proteins competent for inducer-independent transcription activation (11).

## ACKNOWLEDGEMENTS

The authors thank Drs Hylary Trayer and Tom Young for help with protein purification. CMM was funded by the Spanish Ministry of Education, and SJWB is a Royal Society EPA Fund Research Fellow. This work was funded by the UK Science and Engineering Research Council with Project Grant GRF 75322.

## REFERENCES

1 Webster, C., Kempsell, K., Booth, I. and Busby, S. (1987) *Gene* **59**, 253–263.
2 Webster, C., Gardner, L. and Busby, S. (1989) *Gene* **83**, 207–213.
3 Webster, C., Gaston, K. and Busby, S. (1988) *Gene* **68**, 297–305.
4 Caswell, R., Williams, J., Lyddiatt, A. and Busby, S. (1992) *Biochem. J.* **287**, 493–499.
5 Henikoff, S., Wallace, J. and Brown, J. (1989) *Methods Enzymol.* **183**, 111–132.
6 Ramos, J.L., Rojo, F., Zhou, L. and Timmis, K.N. (1990) *Nucleic Acids Res.*, **18**, 2149–2152.
7 Gallegos, M.T., Michán, C. and Ramos, J.L. (1993) *Nucleic Acids Res.*, **21**, 807–810.
8 Francklyn, C. and Lee, N. (1988) *J. Biol. Chem.* **263**, 4400–4407.
9 Brunelle, A. and Schleif, R.F. (1989) *J. Mol. Biol.* **209**, 607–622.
10 Zhou, L., Timmis, K.N. and Ramos, J.L. (1990) *J. Bacteriol.* **172**, 3707–3710.
11 Menon, K. and Lee, N. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 3708–3712.
12 Bustos, S.A. and Schleif, R.F. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 5638–5642.
13 Carra, J.F. and Schleif, R.F. (1993) *Nucleic Acids Res.* **21**, 435–438.
14 Caswell, R., Webster, C. and Busby, S. (1992) *Biochem. J.* **287**, 501–508.
15 Williams, J., Michán, C., Webster, C. and Busby, S. (1994) *Biochem. J.* **300**, 757–763.
16 Casadaban, M. and Cohen, S. (1980) *J. Mol. Biol.* **138**, 179–207.
17 Gaston, K., Chan, B., Kolb, A., Fox, J. and Busby, S. (1988) *Biochem. J.* **235**, 809–918.
18 Grodberg, J. and Dunn, J.J. (1988) *J. Bacteriol.* **170**, 1245–1253.
19 Studier, F.W. (1991) *J. Mol. Biol.* **219**, 37–44.
20 Hussain, H, Grove, J., Griffiths, L., Busby, S. and Cole, J. (1994) *Mol. Microbiol.* **12**, 153–163.
21 Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
22 Henikoff, S. (1984) *Gene* **28**, 351–359.
23 Cole, J.A., Coleman, K.J., Compton, B.E., Kavanagh, B.M. and Keevil, C.W. (1974) *J. Gen. Microbiol.* **85**, 11–22.
24 Laemmli, U.K. (1970) *Nature* **227**, 680–685.
25 Anderson, B.L., Berry, R.W. and Telser, A. (1983) *Anal. Biochem.* **132**, 365–375.
26 Spassky, A., Busby, S. and Buc, H. (1984) *EMBO J.* **3**, 43–50.
27 Ramos, J.L., Michán, C., Rojo, F., Dwyer, D. and Timmis, K.N. (1990) *J. Mol. Biol.* **211**, 373–382.
28 Michán, C., Kessler, B., de Lorenzo, V., Timmis, K.N. and Ramos, J.L. (1992) *Mol. Gen. Genet.* **235**, 406–412.
29 Clarke, P., Lee, J.H., Burke, K. and Wilcox, G. (1992) *Gene* **117**, 31–37.