Overexpression, purification and characterization of the *Escherichia coli* MelR transcription activator protein

Richard CASWELL,* Jackie WILLIAMS,* Andrew LYDDIATT† and Stephen BUSBY*‡

Schools of *Biochemistry and †Chemical Engineering, University of Birmingham, P.O. Box 363, Birmingham B15 2TT, U.K.

The gene encoding *Escherichia coli* MelR protein has been cloned in the expression vector pJLA502. MelR has been overexpressed, substantially purified and shown to bind to DNA fragments carrying the *melAB* promoter. A truncated version of the *melR* gene, encoding the C-terminal half of MelR, was also cloned into pJLA502; the protein product of this truncated gene binds to the *melAB* promoter but was not overproduced. A number of amino acid substitutions were made in the recognition helices of two putative helix-turn-helix motifs in the C-terminal part of MelR, and the effects of these mutations on MelR-dependent transcription initiation at the *melAB* promoter have been measured.

INTRODUCTION

MelR protein in *Escherichia coli* is a transcription activator required for the expression of the *melAB* operon (Webster *et al.*, 1989), which encodes proteins essential for melibiose metabolism (Hanatani *et al.*, 1984). The *melR* and *melAB* genes are transcribed from divergent promoters (Webster *et al.*, 1987): expression from the *melR* promoter (*pmelR*) is dependent on the cyclic AMP receptor protein (Webster *et al.*, 1988), whereas the *melAB* promoter (*pmelAB*) is dependent on MelR binding at two sites located upstream of the *melAB* transcription start (Fig. 1) (Webster *et al.*, 1989). Interestingly, transcription initiation from *pmelAB* is totally dependent on the presence of melibiose in the growth medium, most likely because MelR activity is triggered by melibiose.

Nucleotide sequence analysis has shown that the *melR* gene is related to a number of other genes encoding transcription activators, including *araC* and *rhaR* from *E. coli*, *virF* from *Yersinia pestis*, *toxT* from *Vibrio cholerae*, *exsA* from *Pseudomonas aeruginosa* and *xylS* from *Pseudomonas putida* (Henikoff *et al.*, 1989; Ramos *et al.*, 1990; Frank & Iglewski, 1991; DiRita, 1992). MelR protein thus belongs to a family of transcription activators which may share common features with respect, for example, to DNA binding and mechanism of transcription activation. Since amino acids in the C-terminal part of AraC protein are important for DNA binding (Franclyn & Lee, 1988; Brunelle & Schleif, 1989; Menon & Lee, 1990), and since sequence similarities are particularly strong at the C-termini, it is likely that members of the AraC family share a common DNA-binding structure. Presumably, features at the N-terminal end of these proteins, where there are less sequence similarities, are important for triggering activity. Interestingly, both the MelR and AraC proteins bind to their cognate sites in both the presence and the absence of the cognate activator (arabinose and melibiose respectively) (Webster *et al.*, 1989; Lobell & Schleif, 1991).

Although a lot is known about AraC, related activators have been little studied *in vitro* and no structural information has been reported. In previous work we cloned the *melR* gene into multicopy plasmids but expression was poor, presumably because it carries poor transcription and translation start signals (Webster *et al.*, 1988, 1989). Here, we describe the overproduction and purification of MelR protein. We show that the purified protein

binds to sequences at *pmelAB* and we present evidence that the C-terminal portion of MelR protein carries the elements responsible for binding to DNA.

EXPERIMENTAL

Strains and plasmids

The *mel lac* host strain M182 previously described (Casadaban & Cohen, 1980) was used throughout this work. The plasmids used in this work are listed in Table 1. Standard methods for gene manipulation were used throughout (Maniatis *et al.*, 1982).

Cloning of *melR* into pJLA502 to give pRC1

The *EcoRI*–*HindIII* polylinker of pUC9 was replaced with a chemically synthesized *EcoRI*–*BglII*–*SmaI*–*BamHI*–*HindIII* linker. The *BglII*–*EcoRV* fragment from pKK3 covering *melR* (Fig. 1) was then cloned between the *BglII* and *SmaI* sites of this linker to give pAA300, which carries a *BamHI* site just downstream of the *melR* translation stop codon. pAA300 was linearized with *BglII*, and standard protocols employing Bal-31 exonuclease (Maniatis *et al.*, 1982) were used to delete progressively longer segments of sequence from upstream of *melR*. After re-ligation in the presence of *NcoI* linker (5' GCCATGGC 3', Amersham), the extent of deletions in recombinants were estimated by gel analysis of *NcoI*–*BamHI* restriction digests. A recombinant was selected in which the *NcoI* linker had ligated 20 bases upstream from the *melR* translation start site. DNA was

Fig. 1. Regulatory region of the *E. coli* melibiose operon

The regulatory region is depicted as 'tramlines' with the *melR* structural gene (906 bp) shaded black. The location of the transcription starts for *pmelR* and *pmelAB* are indicated by circles joined to wavy arrows showing the direction of transcription. The location of restriction sites exploited in this work are shown: *NcoI* and *BamHI* sites introduced during subcloning are shown in brackets. The two small horizontal arrows between the divergent transcription starts indicate the 18 bp inverted repeat that is the target for MelR binding.

Abbreviations used: *pmelR*, *melR* promoter; *pmelAB*, *melAB* promoter; DTT, dithiothreitol; PEI, poly(ethylene imine).

‡ To whom correspondence should be addressed.

Table 1. List of plasmids used

Name	Brief description	Origin and references
pUC9	Derivative of pBR322 carrying polylinker	Pharmacia (Vieira & Messing, 1982)
pJLA502	Vector for cloning foreign genes under P_L and P_R	Medac, Hamburg, Germany (Schauder <i>et al.</i> , 1987)
pKK3	Source of <i>mel</i> operon DNA	Webster <i>et al.</i> (1987)
pAA300	<i>Bgl</i> III– <i>Eco</i> RV fragment carrying <i>melR</i> cloned into pUC9	This work (Experimental section)
pRC1	<i>melR</i> cloned into pJLA502 with <i>atpE</i> translation start	This work (Fig. 2a)
pKK33	<i>Hae</i> III– <i>Bgl</i> III <i>pmelAB</i> fragment cloned in pUC9 polylinker	Webster <i>et al.</i> (1987, 1989)
pRC10	<i>Hae</i> III– <i>Bam</i> HI fragment covering C-terminal segment of <i>melR</i> cloned in pJLA502	This work (Experimental section)
pJW12	<i>Eco</i> RI– <i>Hind</i> III fragment carrying <i>melR</i> under the control of <i>galP2</i> cloned in pUC9	<i>Eco</i> RI– <i>Hind</i> III fragment from pAA207 (Webster <i>et al.</i> , 1989) cloned in pUC9
pJW13	Derivative of pJW12 carrying deletion in C-terminal end of <i>melR</i> , inactivating MelR	Webster <i>et al.</i> (1989)
pRW2	Broad-host-range <i>lac</i> expression vector	Lodge <i>et al.</i> (1990)
pKK33/RW2	<i>Eco</i> RI– <i>Hind</i> III fragment from pKK33 carrying <i>pmelAB</i> cloned into pRW2, with <i>pmelAB</i> controlling <i>lac</i> expression	This work (Experimental section)

purified from this recombinant and linearized with *Nco*I, and a second round of Bal-31 deletion was performed, using lower Bal-31 concentrations to make smaller deletions. After re-ligation in the presence of *Nco*I linker, a recombinant was found in which the ATG of the *Nco*I linker had replaced the *melR* ATG initiation codon. The *Nco*I–*Bam*HI fragment from this recombinant carrying *melR* was cloned into pJLA502 to give pRC1 (see Fig. 2).

Protein technology

Total protein was assayed using the Bio-Rad Protein Assay kit, which is based on the Coomassie Blue G250 dye-binding assay of Bradford (1976). Protein samples were analysed by PAGE in the presence of SDS using the Laemmli (1970) method, employing 5% stacking and 12.5% acrylamide resolving gels. Protein samples were mixed with an equal volume of standard Laemmli sample buffer and boiled for 5 min prior to loading onto gels. The gels were visualized after staining with Coomassie Blue R250. The percentage content of MelR in any sample was estimated after densitometry: stained gels were dried between cellulose sheets and were scanned with a Pharmacia LKB Ultrosan XL laser densitometer, and the results were evaluated using the Gelscan XL software package. Gels were calibrated with the high- and low-molecular-mass standard mixtures from Sigma (catalogue nos. SDS7 and SDS6H).

Fig. 2. Cloning and overexpression of MelR protein

(a) Cloning of *Nco*I–*Bam*HI fragment carrying *melR* into pJLA502. The circle is a sketch of pJLA502 showing the location of the pBR322 origin of replication (*ori*), the gene encoding β -lactamase (*amp*^r), the gene encoding the thermosensitive λ repressor (*Cl*^{ts857}), the λ promoters P_R and P_L , the *atpE* translation start (TIR) and the phage fd transcription terminator. The location of the *Nco*I and *Bam*HI sites used in the cloning of *melR* to give pRC1 are shown. The thickened, shaded part of the arc representing the cloned fragment indicates the MelR coding sequence. (b) Analysis of proteins made by pRC1. The figure shows an SDS/polyacrylamide gel stained with Coomassie Blue, calibrated with high-molecular-mass (lane 1) and low-molecular-mass (lane 2) protein standard mixtures. Clear lysates were made from M182 cells carrying pJLA502 or pRC1 and were analysed in lanes 3–6; in each case 6 μ g of total protein was loaded. The samples were as follows: lysate from M182 pJLA502 grown at 30 °C (lane 3), lysate from M182 pJLA502 grown at 43 °C (lane 4), lysate from M182 pRC1 grown at 30 °C (lane 5), and lysate from M182 pRC1 grown at 43 °C (lane 6).

Overproduction and purification of MelR

A 500 ml culture of M182 cells carrying pRC1 was grown with shaking at 30 °C to an A_{600} of 1.0 in L broth containing 80 μ g of ampicillin/ml. The culture was then transferred to a 43 °C shaking bath for a further 4 h. Cells were harvested by centrifugation, resuspended in 12 ml of 10 mM-Tris (pH 7.6)/1 mM-EDTA and lysed by French pressing. Phenylmethanesulphonyl fluoride (0.1 M) in acetone was then added to give a final concentration of 1 mM. The cell debris was removed by centri-

fugation at 30000 *g* for 30 min at 4 °C to give a clear lysate (see Table 2).

Nucleic acids were precipitated from the clear lysate by the addition of a solution of poly(ethylene imine) (PEI; from BASF; 1%, w/v, pH 8.0) to a 0.025% (w/v) final concentration. The sample was incubated at 4 °C for 15 min with occasional gentle stirring and the precipitate was removed by centrifugation for 10 min at 30000 *g* at 4 °C to give the PEI supernatant (Table 2).

(NH₄)₂SO₄ was dissolved in the PEI supernatant to 50% saturation and the sample was incubated at 4 °C for 15 min. The precipitated protein was harvested by centrifugation for 10 min at 10000 *g* at 4 °C and the supernatant was discarded. The pellet was redissolved in a minimum volume of column running buffer [50 mM-potassium phosphate, pH 7.6, 100 mM-NaCl, 0.1 mM-EDTA, 1 mM-dithiothreitol (DTT)] and dialysed against this buffer at 4 °C overnight.

The dialysed (NH₄)₂SO₄ pellet (4 ml, containing 80 mg of protein) was recovered and applied to a column containing 30 ml of Whatman phosphocellulose which had been pre-equilibrated in column running buffer. The column was rinsed successively with 60 ml of column running buffer, 120 ml of this buffer plus 0.4 M-NaCl and, finally, 120 ml of buffer plus 0.7 M-NaCl. Fractions of 6 ml were collected and fractions containing MelR protein were pooled. MelR was precipitated from these fractions by the addition of 75% (NH₄)₂SO₄ and centrifugation as before. Pellets were resuspended in small volumes of storage buffer and dialysed against this buffer [100 mM-potassium phosphate, pH 7.4, 50 mM-KCl, 1 mM-EDTA, 1 mM-DTT, 20% (v/v) glycerol]. Samples were stored at -70 °C.

Gel retardation assays

The *EcoRI*-*HindIII* fragment carrying *pmelAB* was purified from plasmid pKK33, treated with calf alkaline phosphatase and end-labelled with [γ -³²P]ATP. Samples of 10 μ l were prepared containing 1 nM labelled fragment in 40 mM-Tris, pH 7.9, 100 mM-KCl, 50 mM-NaCl, 5 mM-EDTA, 0.5 mM-DTT, 50 μ g of BSA/ml, 500 μ g of sheared salmon sperm DNA/ml plus 5 mM-melibiose. Samples of 5 μ l containing MelR in storage buffer were added to the labelled DNA fragment, and after 10 min of incubation at room temperature, 5 μ l of loading buffer [50% (v/v) glycerol, 0.2 mg of xylene cyanol FF/ml] was added and the sample was applied to the gel under tension. Gels contained 5% polyacrylamide and were run in standard TBE buffer at 50 V. DNA bands in gels were visualized by autoradiography.

Where crude extracts were used as a source of MelR protein (e.g. Fig. 6), 5 ml of culture was harvested by centrifugation. Cells were resuspended in 0.5 ml of MelR storage buffer [100 mM-potassium phosphate, pH 7.4, 50 mM-KCl, 1 mM-EDTA, 1 mM-DTT, 20% (v/v) glycerol] and disrupted by sonication. After removal of cell debris by centrifugation in an Eppendorf microfuge (10 min at high speed), the clear extracts were used directly or stored at -70 °C.

Construction of pRC10 carrying the C-terminal domain of MelR

The *HaeIII*-*BamHI* fragment (Fig. 1), covering the segment of *melR* corresponding to the C-terminal part of MelR, was purified from pRC1 (Fig. 2). It was then ligated into an *NcoI*-*BamHI* vector purified from pRC1, together with a short linker (made from the synthetic oligodeoxynucleotides 5' CATGGATATCGCAGG 3' and 5' CCTGCGATATC 3') to join the *HaeIII* and *NcoI* ends. Insertion of this linker gives pRC10, and results in the attachment of Leu-159 of *MelR* to the N-terminal sequence Met-Asp-Ile-Ala-Gly, with the initiating methionine under the control of the *atpE* translation start.

Site-directed mutagenesis of *melR*

As a starting point for these experiments, we subcloned the *EcoRI*-*HindIII* fragment from pAA207, described by Webster *et al.* (1989), into pUC9 to give pJW12. This fragment carries *melR* under the control of the constitutive promoter *galP2* and gives sufficient expression *in vivo* for optimal melibiose-dependent expression from *pmelAB*. The fragment was transferred to M13mp19 (Pharmacia) and single-stranded DNA was made by standard protocols. Mutations creating the changes listed in Table 3 were obtained by site-directed mutagenesis with synthetic oligodeoxynucleotides purchased from Alta Bioscience at the University of Birmingham. The Amersham Mutagenesis Kit (catalogue no. RPN 1523) was used, following the Amersham protocols exactly. After sequence analysis to check the mutations, fragments carrying mutated versions of *melR* were cloned back into pJW12.

Assay of MelR function *in vivo*

The *EcoRI*-*HindIII* fragment carrying *pmelAB* was purified from plasmid pKK33 and cloned into the broad-host-range *lac* expression vector, pRW2 (Table 1), to give pKK33/RW2. M182 cells carrying this plasmid were grown in media carrying 35 μ g of tetracycline/ml: the cells are Lac⁻, as *lacZ* expression is controlled by *pmelAB*, which requires MelR for activity. MelR was supplied by transforming M182 cells (pKK33/RW2) with pJW12 or mutated derivatives and plating transformants on media containing 35 μ g of tetracycline/ml plus 80 μ g of ampicillin/ml. Note that pRW2 and pJW12 derivatives carry different compatible replication origins and encode resistance to different antibiotics. β -Galactosidase expression in cells carrying pKK33/RW2 and pJW12 derivatives was assayed in cells grown in minimal medium with fructose as a carbon source, exactly as previously described (Busby & Dreyfus, 1983). The assay was performed at least three times on each transformant in medium either with or without 0.05% (w/v) melibiose.

RESULTS

Cloning and overproduction of MelR

Fig. 1 shows a map of the region of the *E. coli* chromosome containing *pmelAB*, *pmelR* and *melR*, which had previously been subcloned and sequenced (Webster *et al.*, 1987, 1988, 1989). To overproduce MelR protein we exploited plasmid pJLA502, developed by Schauder *et al.* (1987). This plasmid allows the cloning of genes under the control of the P_L and P_R promoters of bacteriophage λ ; pJLA502 also encodes a thermosensitive allele of λ repressor so that expression from P_L and P_R can be heat-induced (Fig. 2a). An additional feature is the strong *atpE* translation start signal which promotes translation initiation from the ATG codon in the unique *NcoI* site 5' CCATGG 3'.

The *melR* gene was cloned into pJLA502 on an *NcoI*-*BamHI* fragment to give plasmid pRC1 (Fig. 2a). To do this, we first inserted a *BamHI* linker at the *EcoRV* site just downstream of *melR* (Fig. 1) and then used Bal-31 exonuclease to create nested deletions from the upstream *BglIII* site (see the Experimental section for details of constructions). Bal-deleted DNA was ligated with *NcoI* linker (5' GCCATGGC 3') and resulting *NcoI*-*BamHI* fragments carrying *melR* were cloned into pJLA502. Gel analysis of the recombinants allowed the selection of one recombinant, pRC1, in which the ATG initiation codon of *melR* had been replaced by the ATG of the *NcoI* linker. This resulted in the replacement of the wild-type *melR* N-terminal sequence ATG AAT ACA GAT, encoding Met-Asn-Thr-Asp, by ATG GCG AAT ACA GAT, encoding Met-Ala-Asn-Thr-Asp. Thus, in pRC1, the *atpE* translation start signal directs translation of MelR with an additional alanine at the second codon.

Fig. 3. Elution of protein from a phosphocellulose column

The protein absorption (left ordinate) of different samples taken during the purification of MelR protein is represented as a function of elution volume from the column (which is fully described in the Experimental section). The salt concentration (0–0.7 M-NaCl; right ordinate) in the buffer used to eluate the column is depicted as a broken line. The brackets marked M1 and M2 denote the fractions which were combined to give MelR pool 1 and MelR pool 2 respectively.

Fig. 4. Gel analysis of fractions from the phosphocellulose column

SDS/polyacrylamide gel stained with Coomassie Blue, calibrated with low-molecular-mass protein standard mixture (lane 1). The gel was loaded as follows: 6 µg of dialysed (NH₄)₂SO₄ pellet (lane 2), 6 µg of flow-through proteins (lane 3), 1 µg of MelR pool 1 (lane 4) and 1 µg of MelR pool 2 (lane 5).

Table 2. Purification of MelR protein

Values are from one preparation. The preparation was performed independently three times with good reproducibility.

Sample	Total protein (mg)	MelR (%)	Total MelR (mg)	Yield (%)
Cleared lysate	165	8	13.2	100
PEI supernatant	138	8	11.0	83
Dialysed (NH ₄) ₂ SO ₄ pellet	80	13	10.4	78
MelR pool 1	1.4	85	1.2	9
MelR pool 2	1.2	50	0.6	4.5

Fig. 5. Gel retardation assay of MelR throughout purification

Autoradiogram of gel analysis of the mobility of the labelled KK33 fragment carrying *pmelAB* with two MelR-binding sites. Samples loaded on the gel contained the following additions to each lane: 1, sample buffer only; 2, MelR pool 1, 0.35 µg of protein, 0.3 µg of MelR, final concn. 1 µM; 3, MelR pool 1, 0.09 µg of protein, 0.07 µg of MelR, final concn. 0.25 µM; 4, MelR pool 2, 0.56 µg of protein, 0.3 µg of MelR, final concn. 1 µM; 5, MelR pool 2, 0.14 µg of protein, 0.07 µg of MelR, final concn. 0.25 µM; 6, dialysed (NH₄)₂SO₄ pellet, 2.8 µg of protein, 0.3 µg of MelR, final concn. 1 µM; 7, dialysed (NH₄)₂SO₄, 0.7 µg of protein, 0.07 µg of MelR, final concn. 0.25 µM; 8, cleared lysate, 3.38 µg of protein, 0.3 µg of MelR, final concn. 1 µM; 9, cleared lysate, 0.84 µg of protein, 0.07 µg of MelR, final concn. 0.25 µM. The position of migration of the unbound fragment is marked 'Free', and the locations of the four complexes (see Caswell *et al.*, 1992) between MelR and *pmelAB* DNA are numbered 1–4.

Fig. 2(b) shows gel analysis of proteins in clear lysates made from strain M182 carrying either pJLA502 or pRC1 grown at either 30 °C or 43 °C. The gel shows the appearance of a substantial band due to pRC1 in the cells grown at 43 °C, which represents around 8% of the total protein. The size of this band corresponds well with the predicted molecular mass of MelR protein (34927 Da: Webster *et al.*, 1987).

Purification of MelR protein

A purification protocol was devised for MelR starting with a 500 ml culture of M182 cells carrying pRC1. After heat induction and preparation of a clear lysate, nucleic acids were precipitated with PEI. MelR was then precipitated with (NH₄)₂SO₄ and the pellet was resuspended, dialysed and applied to a phosphocellulose column. Fig. 3 shows a typical elution profile from the column. Electrophoresis in polyacrylamide gels was used to monitor the purification of MelR. Fig. 4 shows the gel analysis of proteins in the fractions from the phosphocellulose column. The gel shows that MelR, which is at least 85% pure, elutes from the column in buffer containing 0.7 M-NaCl. Table 2 shows that the purification gives milligram quantities of MelR, with a yield of around 10%.

Although Fig. 4 shows that the 'purified' MelR is the expected size, it was possible that we had inadvertently purified another protein of similar size. To exclude this possibility, we investigated the DNA-binding properties of the purified protein using gel retardation assays. Using crude extracts as a source of protein, we previously showed that MelR specifically binds to two sites at *pmelAB* (Webster *et al.*, 1989). Fig. 5 shows experiments in which samples from different stages of the purification were added to labelled fragment covering the *melAB* promoter. We used a fragment from pKK33, which carries the *HaeIII*–*BglIII* sequence covering *pmelAB* (Fig. 1) cloned between *EcoRI* and *HindIII*

Fig. 6. Gel retardation assay for the function of truncated MelR

Autoradiogram of gel analysis of the mobility of the labelled KK33 fragment as in Fig. 5. Samples loaded on the gel include crude lysates containing 5 µg of protein, prepared from cells as follows: 1, M182 cells carrying pJLA502 grown at 30 °C; 2, M182 carrying pJLA502 grown at 45 °C; 3, M182 carrying pRC1 grown at 30 °C; 4, M182 carrying pRC1 grown at 43 °C; 5, M182 carrying pRC10 grown at 30 °C; 6, M182 carrying pRC10 grown at 43 °C. The position of migration of the unbound fragment is marked 'Free'.

linkers. Labelled KK33 fragment was incubated with MelR pool 1, MelR pool 2, the dialysed (NH₄)₂SO₄ pellet or the clear lysate and then loaded on a non-denaturing polyacrylamide gel. In each case the concentration of total protein was adjusted so that the total amount of MelR was constant. Because MelR binds independently to the two binding sites in the KK33 fragment, up to four retarded bands are observed, the intensities of which vary according to the MelR concentration (Caswell *et al.*, 1992). Fig.

5 shows that the pattern of band retardation by purified MelR is similar to that seen in the crude extracts, proving that, at least according to this criterion, we had purified the correct protein. Interestingly, purified MelR retarded the labelled DNA fragments less than did equivalent amounts of MelR in the crude extract; although it is possible that some helper factor is being purified away, it is more likely that some MelR DNA-binding activity is lost during the purification.

Subcloning the C-terminal domain of MelR

The available evidence suggests that proteins of the MelR/AraC class bind to DNA via a C-terminal DNA-binding domain (Francklyn & Lee, 1988; Brunelle & Schleif, 1989; Menon & Lee, 1990; Ramos *et al.*, 1990). To attempt to show this for MelR, we cloned a segment of *melR*, corresponding to the C-terminal portion of the protein, into pJLA502 to give pRC10. In pRC10, MelR sequence from Leu-159 to the C-terminal Gly-302 is attached to the N-terminal sequence Met-Asp-Ile-Ala-Gly, which is controlled by the *atpE* translation start and the λ P_L and P_R promoters. However, after polyacrylamide gel analysis of proteins in crude extracts made from heat-induced M182 cells carrying pRC10, we found no evidence for overproduction of the C-terminal domain of MelR (results not shown). Despite this, evidence for the accumulation of some truncated MelR protein comes from gel retardation assays. Labelled KK33 fragment carrying *pmelAB* was incubated with different clear lysates and protein-induced shifts in band migration were measured by electrophoresis. Fig. 6 shows that extracts from M182 cells carrying pRC10 grown at 43 °C contained some factor that retarded the *pmelAB* fragment. This factor was not present in cells grown at 30 °C and was also absent from extracts of cells carrying the control plasmid pJLA502. Interestingly, whereas extracts containing full-length MelR gave multiple bands in the assay, the truncated MelR gave a single band.

Fig. 7. Location of helix-turn-helix motifs in MelR

The horizontal bar represents the 302-amino-acid MelR protein with the cross-hatched region representing the sequence from Val-184 to Gly-302 which displays particularly strong similarities to other proteins of the AraC class. The location and sequence of the two 20-amino-acid segments of MelR that fit best with matrices predicting helix-turn-helix motifs are shown. Note that the sequences that fit best with helix-turn-helix predictions in other members of the AraC family fall at similar positions (Henikoff *et al.*, 1989). The first six amino acids of the putative 'recognition helix' have been numbered at each motif and the location and nature of the mutations that we created are shown by arrows.

Table 3. Effects of mutations in *melR* on expression from *p_{melAB}*

The Table shows β -galactosidase activities measured in M182 cells carrying pKK33/RW2, with *p_{melAB}* controlling expression of the *lacZ* gene. Cells also contained *melR* carrying different mutations introduced on pJW12 or pJW13. β -Galactosidase activities are expressed in standard Miller (1972) units. Each value is the average of at least three independent determinations.

Plasmid/mutation in <i>melR</i>	β -Galactosidase activity (units)	
	– Melibiose	+ Melibiose
pJW13/deletion of <i>melR</i> C-terminus	24	20
pJW12/wild-type <i>melR</i>	24	480
pJW12/ <i>melR</i>	24	560
Asn-222 → Ala		
pJW12/ <i>melR</i>	24	250
Ser-271 → Ala		
pJW12/ <i>melR</i>	24	140
Ser-272 → Asn		
pJW12/ <i>melR</i>	24	460
Arg-269 → Gly		
pJW2/ <i>melR</i> stop codon at 270	24	32

Mutations in the C-terminal domain of MelR

Sequence comparisons show that MelR, and other members of the AraC class of proteins, contain two potential helix-turn-helix motifs in the C-terminal half of the sequence (Henikoff *et al.*, 1989; Tobin & Schleif, 1990). In MelR protein these correspond to residues 210–229 (helix-turn-helix 1) and 259–278 (helix-turn-helix 2), which are illustrated schematically in Fig. 7. To attempt to confirm that the C-terminal part of MelR is involved in DNA binding and to examine the role, if any, of these putative helix-turn-helix motifs, we used site-directed mutagenesis to create changes at key residues in these sequences. X-ray analysis of a number of DNA-binding proteins has shown that it is the second helix (the recognition helix) of helix-turn-helix motifs that interacts directly with specific sequences (Steitz, 1990). Since position 2 of the recognition helix is usually directly involved in contacts with bases, we made the changes Asn-222 → Ala and Ser-271 → Ala, which alter the amino acids at this position in both of the putative recognition helices (Fig. 7).

To determine the effects of these mutations, we measured MelR-dependent transcription from *p_{melAB}* using plasmid pKK33/RW2, a derivative of a broad-host-range *lac* expression vector, in which *lac* expression is controlled by *p_{melAB}*. M182 cells carrying pKK33/RW2 were transformed with pJW12 derivatives containing wild-type or mutated *melR*. As a control we used pJW13, a derivative of pJW12 carrying a deletion that removes the C-terminus of MelR (see Tables 1 and 3). Note that, in pJW12, *melR* is constitutively expressed from the *galP2* promoter and there is no autoregulation of MelR expression (Webster *et al.*, 1988). In the M182 host strain used in this experiment, expression from *p_{melAB}* is totally dependent on MelR supplied by pJW12. We previously showed that expression from *p_{melAB}* is not affected by large changes in *melR* gene dosage (Webster *et al.*, 1987, 1989); thus it is unlikely that our measurements are distorted by small changes in MelR expression resulting from mutations in pJW12.

The data in Table 3 show that MelR activity *in vivo* is not affected by the change Asn-222 → Ala in helix-turn-helix 1, but is halved by the change Ser-271 → Ala in helix-turn-helix 2. Since this suggests that the second motif might be more important in DNA binding, we created two further mutations in this region,

Ser-272 → Asn and Arg-269 → Gly. Additionally, as a control, we converted the codon for Ser-270 into a stop codon. The data in Table 3 show that the change Ser-272 → Asn reduces MelR activity by 75 %, the change Arg-269 → Gly has no effect, and the introduction of a stop codon at position 270 suppresses melibiose-dependent transcription activation from *p_{melAB}*.

DISCUSSION

The *melR* gene was discovered accidentally (Webster *et al.*, 1987) in a segment of DNA essential for expression from the promoter region immediately upstream of the *melA* gene (Shimamoto *et al.*, 1984; Liljestrom & Liljestrom, 1987). The observation that the *melAB* promoter is positively rather than negatively controlled (Webster *et al.*, 1987) strongly suggested that *melR* encodes a transcription activator, and this was corroborated by the discovery of sequence similarities between MelR and the well-characterized arabinose operon activator, AraC (Henikoff *et al.*, 1989). In previous work, Webster *et al.* (1989) showed that crude extracts from cells carrying MelR-encoding plasmids contain a factor that binds specifically to two sites located upstream of the *melAB* transcription start; however, MelR was not overexpressed in these cells, and we could not detect a protein band corresponding to MelR in electrophoresis experiments. In the present work we have successfully overexpressed and substantially purified MelR, and verified that the gene encodes a protein of the expected size. Moreover, the protein binds to specific sites at *p_{melAB}* (Fig. 5; see also Caswell *et al.*, 1992). This work should provide a base for future *in vitro* studies of melR binding to its target. Notwithstanding, a number of problems are apparent. Firstly, the overexpression of MelR is not outstanding and the yield of our purification protocol places severe limitations on the amount of pure protein that can be obtained. Secondly, there is some evidence for loss of activity of the protein during the purification; it is not clear whether this is due to the intrinsic instability of the protein, as with AraC protein, or to the loss of some essential auxiliary factor. In this regard it is worth noting that, although MelR-dependent stimulation of gene expression *in vivo* is dependent on the addition of melibiose to the medium, the binding of purified MelR at *p_{melAB}* is not dependent on melibiose *in vitro* (Webster *et al.*, 1989; Caswell *et al.*, 1992). Moreover, we were unable to demonstrate MelR-dependent stimulation of run-off transcription from *p_{melAB}* *in vitro* (R. Caswell & S. Busby, unpublished work), raising the possibility that some factor is missing.

Genetic analysis strongly suggests that the C-terminal part of AraC protein is responsible for DNA sequence recognition and that at least one of the two potential helix-turn-helix motifs is involved (Brunelle & Schleif, 1989; Menon & Lee, 1990). The striking sequence similarities between MelR and other members of the AraC class of proteins provide circumstantial evidence that one or other of the two potential helix-turn-helix motifs in MelR may also be involved in binding. Previously we showed that deletion of 16 amino acid residues from the C-terminus of MelR was sufficient to stop MelR binding to *p_{melAB}* in gel retardation assays (Webster *et al.*, 1989). Here we demonstrate that a segment of protein covering the C-terminal half of MelR is sufficient for some specific binding, although titrations showed that binding was weaker than with the full-length protein (R. Caswell & J. Woodward, unpublished work). Our results with amino acid substitutions in MelR are also consistent with the suggestion that helix-turn-helix 2 is involved in DNA recognition. Changes at positions 2 and 3 of the recognition helix reduce DNA binding, but the change Arg-269 → Gly in the turn (Fig. 7) has little effect. Surprisingly, the change at position 2 of the recognition helix in motif 1 had no effect. These are the first

reported single amino acid changes in MelR, and although their effects do not provide a direct proof that motif 2 is involved in DNA binding, they do raise the possibility that MelR differs from AraC, in which changes in the helix-turn-helix corresponding to motif 1 have clearer effects than changes in the sequence corresponding to motif 2 (Brunelle & Schleif, 1989). Further investigations will require the purification of mutated proteins and careful measurements of binding to cognate sites.

We are most grateful to John McCarthy for advice on the use of pJLA502 and to Eva Hyde for help in designing the MelR purification protocol. pRC10 was constructed by Jeremy Woodward as part of an undergraduate practical project. This work was funded by the U.K. Science and Engineering Research Council via the University of Birmingham Biochemical Engineering Centre and Project Grant Number GRF 75322.

REFERENCES

- Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248–254
- Brunelle, A. & Schleif, R. (1989) *J. Mol. Biol.* **209**, 607–622
- Busby, S. & Dreyfus, M. (1983) *Gene* **21**, 121–131
- Casadaban, M. & Cohen, S. (1980) *J. Mol. Biol.* **138**, 179–207
- Caswell, R., Webster, C. & Busby, S. (1992) *Biochem. J.* **287**, 501–508
- DiRita, V. (1992) *Mol. Microbiol.* **6**, 452–458
- Francklyn, C. & Lee, N. (1988) *J. Biol. Chem.* **263**, 4400–4407
- Frank, D. & Iglewski, B. (1991) *J. Bacteriol.* **173**, 6460–6468
- Hanatani, M., Yazyu, H., Shiota-Niiya, S., Moriyama, Y., Kanazawa, H., Futai, M. & Tsuchiya, T. (1984) *J. Biol. Chem.* **259**, 1807–1812
- Henikoff, S., Wallace, J. & Brown, J. (1989) *Methods Enzymol.* **183**, 111–132
- Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685
- Liljestrom, P. L. & Liljestrom, P. (1987) *Nucleic Acids Res.* **15**, 2213–2220
- Lobell, R. & Schleif, R. (1991) *J. Mol. Biol.* **218**, 45–54
- Lodge, J., Williams, R., Bell, A., Chan, B. & Busby, S. (1990) *FEMS Microbiol. Lett.* **67**, 221–226
- Maniatis, T., Fritsch, E. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Menon, K. & Lee, N. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 3708–3712
- Miller, J. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Ramos, J., Rojo, F., Zhou, L. & Timmis, K. (1990) *Nucleic Acids Res.* **18**, 2149–2152
- Schauder, B., Blocker, H., Frank, R. & McCarthy, J. (1987) *Gene* **52**, 279–283
- Shimamoto, T., Yazyu, H., Futai, M. & Tsuchiya, T. (1984) *Biochem. Biophys. Res. Commun.* **121**, 41–46
- Steitz, T. (1990) *Q. Rev. Biophys.* **23**, 205–280
- Tobin, J. & Schleif, R. (1990) *J. Mol. Biol.* **211**, 75–89
- Vieira, J. & Messing, J. (1982) *Gene* **19**, 259–268
- Webster, C., Kempell, K., Booth, I. & Busby, S. (1987) *Gene* **59**, 253–263
- Webster, C., Gaston, K. & Busby, S. (1988) *Gene* **68**, 297–305
- Webster, C., Gardner, L. & Busby, S. (1989) *Gene* **83**, 207–213

Received 28 February 1992/16 April 1992; accepted 24 April 1992