Biochem. J. (1992) 287, 501–508 (Printed in Great Britain)

Studies on the binding of the *Escherichia coli* MelR transcription activator protein to operator sequences at the *MelAB* promoter

Richard CASWELL, Christine WEBSTER and Stephen BUSBY*
School of Biochemistry, University of Birmingham, P.O. Box 363, Birmingham B15 2TT, U.K.

Escherichia coli MelR protein binds to two sites located upstream of the *melAB* transcription start site. Although both sites are required for optimal melibiose-dependent expression from the *melAB* promoter, some MelR-dependent expression is found if the upstream site is deleted or if the spacing between the two sites is altered. Gel retardation assays have been exploited to study MelR binding to a DNA fragment carrying just the upstream site. Methylation interference analysis was used to identify one guanine (at −104) which is important for MelR binding. Mutational analysis confirmed the importance of this base and revealed a second position (at −110) where mutations interfere with *melAB* promoter activity. Experiments using potassium permanganate as a probe suggested that the DNA sequence around −110 adopts a distorted conformation. We propose that the mutation at −104 alters MelR binding by interfering with a direct contact, whereas the mutation at −110 primarily affects DNA conformation. The binding of purified MelR protein to a *melAB* promoter fragment carrying both binding sites has also been studied: binding results in four retarded bands in gel assays. Methylation interference experiments have been exploited to identify the binding sites occupied in each complex. Although both binding sites share a common 18 bp sequence, MelR binding to the more upstream site is stronger. We could find no evidence for co-operative interactions between MelR and RNA polymerase and no major effects of melibiose. Some evidence for melibiose-dependent distortion in complexes between MelR and the *melAB* promoter is discussed.

INTRODUCTION

The growth of *Escherichia coli* on melibiose requires the expression of the *melAB* operon, which encodes proteins essential for melibiose metabolism (Hanatani *et al.*, 1984). Expression of this operon is controlled by single promoter (p*melAB*) which is totally dependent on the *melR* gene encoding a transcription activator, MelR (Webster *et al.*, 1987, 1989). The upstream region of p*melAB* contains a perfect inverted repeat of 18 bp, with the two 18 bp elements separated by 20 bp. Using crude extracts from cells carrying plasmids encoding MelR as a source of MelR, Webster *et al.* (1989) showed that MelR protein binds to both of these 18 bp elements. In the preceding paper (Caswell *et al.*, 1992) we described the overproduction and purification of MelR protein. We have now exploited the purified protein to carry out a more detailed study of the interactions between MelR and p*melAB* DNA carrying either just one or both of the 18 bp operator sequences. Additionally, we have used recombinant DNA methods to create mutations throughout one of the 18 bp sequences in order to locate crucial bases involved in interactions with MelR.

EXPERIMENTAL

Strains, plasmids and gene manipulation

The *mel lac* host strain M182 was used throughout this work (Casadaban & Cohen, 1980). The starting plasmids used in this work (pUC9, pKK33, pRW2, pKK33/RW2 and pJW12) were described in Table 1 of the preceding paper (Caswell *et al.*, 1992).

All the new plasmids constructed during this study were derivatives of pKK33 and were made using standard recombinant DNA methodology (Maniatis *et al.*, 1982). Fig. 1 shows a diagram of the p*melAB* *Eco*RI–*Hin*dIII insert in pKK33 and the derivatives pKK-35, -34 and -36. pKK33 carries a *Hae*III–*Bgl*II fragment, covering the *melAB* promoter from base pair −136 to +21, cloned between the *Sma*I and *Bam*HI sites of pUC9 (see Table 1 and Fig. 1 of the preceding paper, Caswell *et al.*, 1992). To make pKK35, the *Eco*RI–*Sau*3A fragment from pKK33 covering MelR binding site 1 (base pairs −136 to −69) was cloned between the *Eco*RI and *Bam*HI sites of pUC9. To make pKK34, the *Sau*3A–*Hin*dIII fragment covering MelR binding site 2 (base pairs −73 to +21) was cloned between the *Bam*HI and *Hin*dIII sites of pUC9. To make pKK36, the *Eco*RI–*Hin*cII fragment from pKK35 covering MelR binding site 1 was ligated to the *Sma*I–*Hin*dIII fragment from pKK34 covering site 2. This results in a derivative of pKK33 with an 11 bp insert, 5′ CGTCGGGGATC 3′, in the *Sau*3A site between MelR binding sites 1 and 2 (bottom line of Fig. 1).

The point mutations described in Fig. 4 were selected after mutagenesis of pKK33 using doped oligodeoxynucleotides, as described by Jayaraman *et al.* (1989). Briefly, the *Eco*RI to *Sau*3A fragment of pKK33 was replaced with a pair of synthetic 'oligos' which had been 'doped' with contaminating bases from positions −110 to −91. A total of 50 of the resulting clones were purified and the nucleotide sequence of the *Eco*RI–*Hin*dIII fragment carrying p*melAB* was determined in each case: 32 of the clones carried mutations in MelR binding site 1; 23 clones carried single base changes at the 13 positions shown in Fig. 4; four of the clones carried the multiple changes shown at the bottom of Fig. 4; and five candidates contained gross rearrangements and were discarded. By convention, p*melAB* sequences are numbered with the transcription start taken as +1 (Webster *et al.*, 1987); point mutations at positions upstream of the start site are denoted by p*nx* where *x* is the new base on the top strand and *n* is the location of the change.

*Eco*RI–*Hin*dIII fragments were purified from plasmid pKK33 and mutated derivatives, and from pKK34, pKK35 and pKK36.

Abbreviations used: p*melAB*, *melAB* promoter; DTT, dithiothreitol; CRP, cyclic AMP receptor protein.
* To whom correspondence should be addressed.

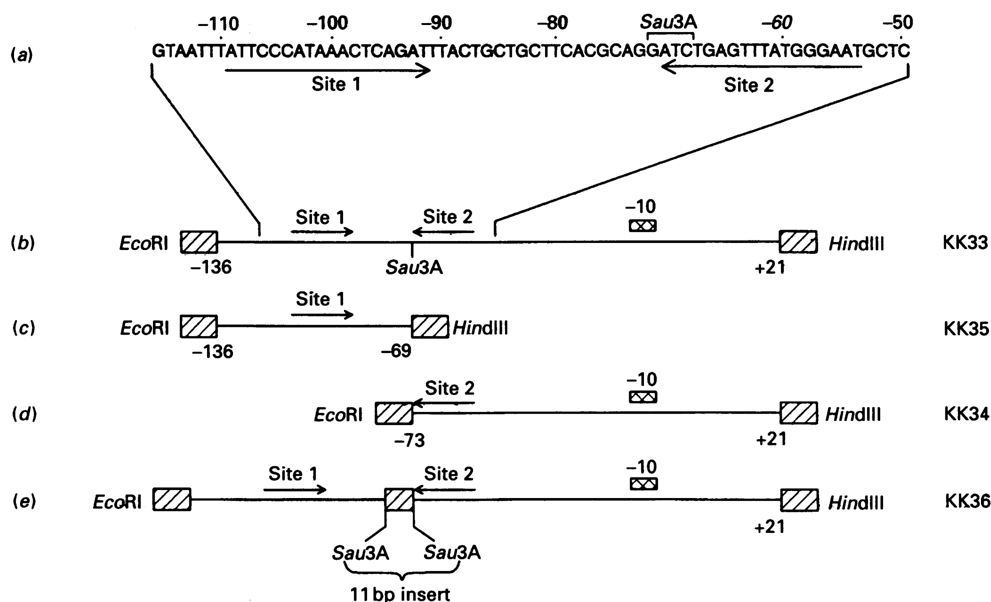

Fig. 1. Schematic representation of p*melAB* promoter fragments

Plasmid pKK33 carries an *Eco*RI–*Hin*dIII fragment covering sequence from −136 to +21 with respect to the *melAB* transcription start site, which is illustrated in (*b*). The *Eco*RI and *Hin*dIII linkers are illustrated as hatched boxes and the cross-hatched rectangle shows the position of the −10 hexamer sequence. The DNA sequence around MelR-binding sites 1 and 2 is expanded in (*a*) and the 18 bp inverted repeat elements are underlined with arrows. The *Sau*3A site used in the following constructions is indicated. (*c*), (*d*) and (*e*) illustrate the *Eco*RI–*Hin*dIII fragments in pKK35, pKK34 and pKK36 respectively, generated by exploiting the *Sau*3A site. MelR-binding sites 1 and 2 are denoted by arrows, and the 11 bp insert in pKK36 is shown as a box.

In experiments to monitor MelR binding, the fragments were labelled using [γ-^{32}P]ATP or [α-^{32}P]dATP. For measurements of p*melAB* activity *in vivo*, fragments were cloned into the broad-host-range *lac* expression vector pRW2.

Assay of p*melAB* activity *in vivo*

pRW2 derivatives containing *melAB* promoter inserts derived from pKK33, pKK34 or pKK36 were transformed into the *mel lac* strain M182, which had previously been transformed with pJW12, encoding *melR*. Note that pRW2 and pJW12 are compatible plasmids and can be maintained in media containing both ampicillin and tetracycline (Caswell *et al.*, 1992). Cells carrying both plasmids were grown in minimal medium plus or minus melibiose, and β-galactosidase assays were performed exactly as outlined in the previous paper (Caswell *et al.*, 1992). Small preparations of plasmids were made from each of the cultures to check that the copy number did not vary from one experiment to another. Since cells carrying pJW12 contain excess MelR protein, we can be sure that measured levels of β-galactosidase reflect p*melAB* activity.

Gel retardation assays and interference experiments

End-labelled *Eco*RI–*Hin*dIII fragments (typically 0.1 nM) were incubated with different concentrations of purified MelR and loaded on polyacrylamide gels exactly as described in the preceding paper (Caswell *et al.*, 1992). In some experiments 5 mM-melibiose was included in all buffers, whereas in others melibiose was omitted.

For dissociation rate assays, reactions were set up in the absence of carrier DNA at a MelR concentration (400 nM) sufficient to bind all the labelled fragment (1 nM). After incubation for 10 min, an excess of pKK35 DNA (equivalent to 200 nM fragment) was added; aliquots were removed at different times and loaded on to the gel under tension. As a control, MelR was first incubated with the competitor DNA and labelled fragment was then added. After electrophoresis the gel was developed by autoradiography using Fuji X-ray film. Half-times for the dissociation of complexes were deduced from semi-log plots after quantifying free and bound band intensities using a Pharmacia LKB Ultroscan XL laser densitometer.

For chemical interference experiments, labelled fragment was chemically modified with either dimethyl sulphate (for G residues) or hydrazine (for C+T residues) by the Maxam–Gilbert method (Maniatis *et al.*, 1982). After removal of dimethyl sulphate or hydrazine, the labelled fragment (1 nM) was incubated with 400 nM-MelR and loaded on to a polyacrylamide gel exactly as in the gel binding assays, except that wider wells (1 cm width) were employed. After electrophoresis and autoradiography, bands corresponding to free and bound fragment were located and cut out. DNA was purified from each band by standard electroelution procedures and treated with piperidine, again using the Maxam–Gilbert protocol. After removal of the piperidine, samples were loaded onto 6% sequencing polyacrylamide gels and, after electrophoresis and autoradiography, the positions of modification-induced cleavage were deduced.

Footprinting protocols

All binding reactions (except controls) contained 400 nM-MelR in a total volume of 20 μl in footprint buffer [20 mM-Tris pH 8.0, 100 mM-NaCl, 5 mM-$MgCl_2$, 0.1 mM-EDTA, 1 mM-dithiothreitol (DTT), 50 mg of BSA/ml, 5% (v/v) glycerol]. Melibiose (5 mM) and 50 nM RNA polymerase holoenzyme (from Northumbria Biologicals Limited) were included where indicated. Target DNA was typically 1 nM, and was 5′-end-labelled on the upper strand at the *Eco*RI site or on the lower strand at the *Hin*dIII site. The protocols of Spassky *et al.* (1984) and O'Halloran *et al.* (1989) were used for DNAase I and permanganate footprinting respectively exactly as in previous work from our laboratory (Chan *et al.*, 1990).

RESULTS

Two MelR-binding sites are needed for optimal melibiose-dependent expression from p*melAB*

The starting point for these experiments was the *Eco*RI–*Hin*dIII fragment from plasmid pKK33 carrying the *E. coli melAB* promoter region (Fig. 1). The nucleotide sequence upstream of the *melAB* transcription start site contains a perfect inverted repeat of two 18 bp elements separated by a 20 bp spacer. In previous studies, we showed that MelR binds to both these 18 bp elements (Webster *et al.*, 1989). To assess the importance of these sites in melibiose-dependent expression, we produced a series of pKK33 derivatives containing the upstream element (site 1; pKK35), the downstream element (site 2; pKK34) or an 11 bp insertion between the two sites (pKK36). To measure p*melAB* activity, *Eco*RI–*Hin*dIII fragments carrying the different promoter fragments were cloned into the broad-host-range *lac* expression vector (pRW2), the resulting plasmids were transferred to cells carrying the *melR* gene and melibiose-dependent β-galactosidase expression was measured.

We previously showed that expression from p*melAB* on the KK33 fragment is dependent on MelR (Webster *et al.*, 1987; Caswell *et al.*, 1992). The data in Table 1 confirm that expression is totally dependent on the presence of melibiose in the growth medium. Removal of the upstream MelR-binding site in the KK34 fragment results in a 4-fold reduction in p*melAB* activity. Interestingly, the insertion of 11 bp between the two MelR-binding sites (KK36) results in a similar reduction in activity. From these data, we conclude that one site is sufficient for some MelR binding, but that optimal p*melAB* activity requires two MelR-binding sites that are correctly juxtaposed.

Binding of purified MelR to a single site

Since one binding site appears to be sufficient for some MelR activity at p*melAB*, we first examined the binding of purified MelR to the single site in the *Eco*RI-*Hin*dIII fragment from pKK35 (corresponding to site 1; Fig. 1). Fig. 2(*a*) shows a titration of ^{32}P-labelled KK35 fragment with MelR, analysed by the gel retardation method. In this experiment, performed in the presence of melibiose, increasing the concentration of MelR protein resulted in the appearance of a single band corresponding to bound fragment. From these data we estimate that 6 nM-MelR is required to retard 50% of the labelled fragment under our conditions. The experiment was repeated in the absence of melibiose and the result appeared identical (not shown). Thus, within experimental error, according to this method, MelR binding to a single site is independent of melibiose. To check this conclusion, we then looked for effects of melibiose on dissociation rates of MelR–site 1 complexes. MelR (400 nM) was added to labelled fragment and, after equilibration, an excess of unlabelled pKK35 competitor DNA was added. Samples were taken at different times after addition of the competitor and loaded immediately on a polyacrylamide gel. Fig. 2(*b*) shows the relative amounts of bound and free fragment in the samples taken at different times. From the data we can conclude that, in these experiments, the half-time for the dissociation is 95 s, and this was not affected by the presence of melibiose in the samples.

Footprint analysis using DNAase I as a probe shows that MelR binding at site 1 protects bases from −85 to −113 (Webster *et al.*, 1989; present paper). To attempt to identify the bases involved in essential contacts with MelR, we performed binding interference assays. Briefly, end-labelled KK35 DNA was partially methylated using dimethyl sulphate, which preferentially modifies G residues under our conditions. An excess of MelR was then added and the sample was loaded on a polyacrylamide gel as in Fig. 2. Bands corresponding to free and bound fragment were cut out of the gel, and the DNA was purified and then treated with piperidine which cleaves at modified bases. Samples were then run on sequencing gels to determine the sites of cleavage. Preferential cleavage at particular positions in the DNA from the free band indicates that methylation at that position hinders or weakens MelR binding.

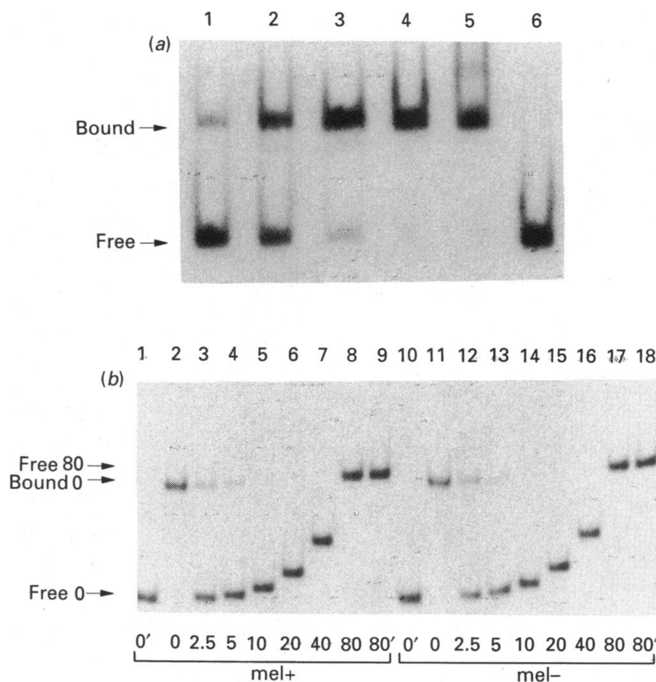

Fig. 2. Gel retardation analysis of MelR binding to site 1

(*a*) Labelled KK35 fragment (0.1 nM) was mixed with purified MelR (lane 1, 1.5 nM; lane 2, 6 nM; lane 3, 23 nM; lane 4, 93 nM; lane 5, 370 nM; lane 6, no protein) and free and bound fragments (indicated by arrows) were separated by electrophoresis. (*b*) Labelled KK35 fragment was incubated with 400 nM-MelR and then an excess of unlabelled fragment was added. Samples were then taken at the different times after addition of the competitor shown on the figure (min). The experiment was performed either with (mel+) or without (mel−) melibiose as indicated. For each experiment, the flanking lanes marked 0′ and 80′ show the migration of samples from a control experiment in which the competitor DNA was added to MelR before the labelled fragment (showing that the competitor DNA is sufficient to complex all the protein).

Table 1. Promoter activity of KK33, KK34 and KK36 fragments

β-Galactosidase activities were measured in extracts from M182 cells carrying the *lacZ* expression vector, pRW2, with the KK33, KK34 or KK36 promoter fragments. Host cells also carried pJW12, encoding the functional *melR* gene. Cells were grown in media plus or minus added melibiose. β-Galactosidase activities are expressed in standard Miller units, and each value is the average of at least three independent determinations.

p*melAB* insert	β-Galactosidase activity (units)	
	−Melibiose	+Melibiose
	20	20
KK33	25	400
KK34	25	95
KK36	25	90

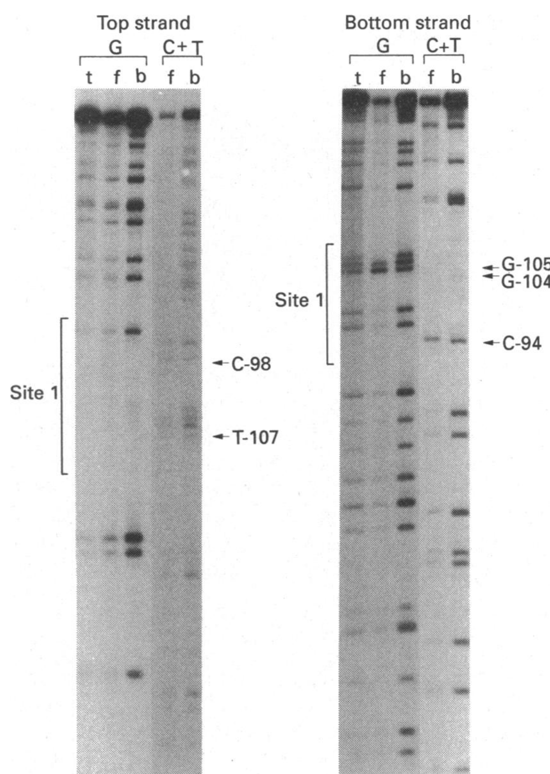

Fig. 3. Interference with MelR binding by chemical modification

KK35 DNA was end-labelled on the top or bottom strand and treated with dimethyl sulphate (G) or hydrazine (C+T). After gel retardation analysis using purified MelR protein, as in Fig. 2(*a*), free (f) and bound (b) fragment was purified and cleaved with piperidine to locate the sites of modification. The Figure shows analysis of the cleavage pattern using a sequence gel. As a control, chemically modified fragment that had been run through the gel retardation assay without MelR was included (t). The gel was calibrated by reference to the published sequence (Webster *et al.*, 1987) and the positions of a number of bases upstream of the *melAB* transcription start site are shown on both the top and bottom strands. The location of the 18 bp MelR-binding element at site 1 is indicated.

Table 2. Effects of different point mutations on p*melAB* activity

β-Galactosidase activities (in standard Miller units) were measured in extracts from M182 cells carrying different plasmids with the KK33 insert cloned in the *lacZ* expression vector pRW2. Plasmids carried different mutations in the KK33 insert. Host cells also carried pJW12, encoding the functional *melR* gene. Cells were grown in media plus or minus added melibiose.

p*melAB* derivative	β-Galactosidase activity (units)	
	−Melibiose	+Melibiose
Wild-type KK33 fragment	20	400
KK33 p104G	20	200
KK33 p110G	20	65
KK33 p96G/p104T	20	200
KK33 p94T/p97G	20	95
KK33 Δ95 p107G	20	95
KK34	20	90

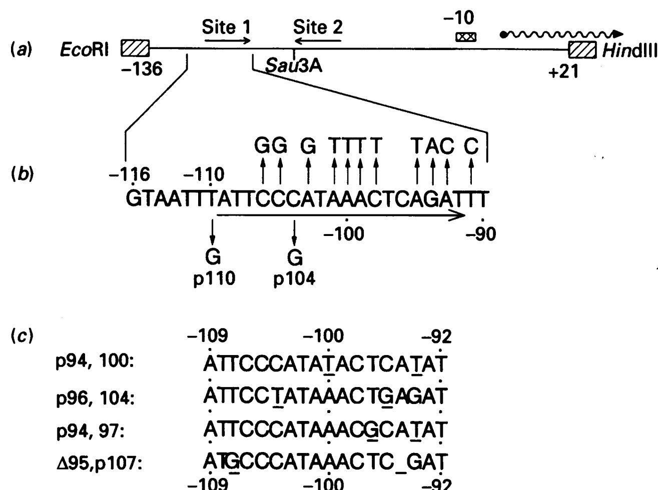

Fig. 4. Mutational analysis of MelR-binding site 1

The top line (*a*) is a scheme of the *melAB* promoter region, with the transcription start shown as a dot with a wavy line indicating the direction of transcription. The sequence of the upper strand around MelR-binding site 1 is shown in (*b*) together with the location and the nature of the single point mutations we isolated. The two changes that have effects on p*melAB* activity *in vivo* are shown below the sequence, and the eleven changes that had no measurable effect are shown above. (*c*) Upper strand sequences at binding site 1 of four derivatives each carrying two mutations; the positions of the base changes are underlined.

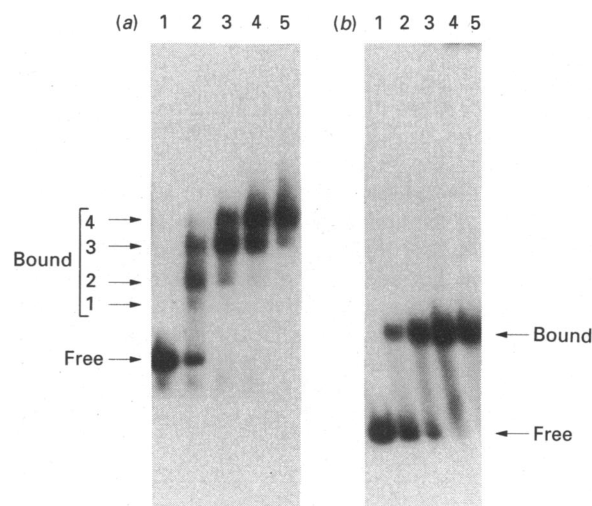

Fig. 5. Gel retardation analysis of MelR binding

(*a*) Labelled KK33 fragment (0.1 nM) was mixed with purified MelR (lane 1, no protein; lane 2, 6 nM; lane 3, 23 nM; lane 4, 93 nM; lane 5, 370 nM), and free and bound fragments (indicated by arrows) were separated by electrophoresis. (*b*) As in (*a*), but with KK35 fragment.

The results in Fig. 3 show that methylation of G residues on the top strand at site 1 (shown in Fig. 1) hardly affects the partition between free and bound. In contrast, fragments methylated on the bottom strand at G-104 are enriched in the free band, implying that MelR makes an important contact with this base. Interestingly, modification of G-105 on the lower strand also appears to weaken MelR binding. The experiment was repeated using DNA treated with hydrazine, which modifies C and T residues. From the results in Fig. 3 it is clear that the modifications had little detectable effects on MelR binding, except that changing bottom-strand C-94 caused a small reduction.

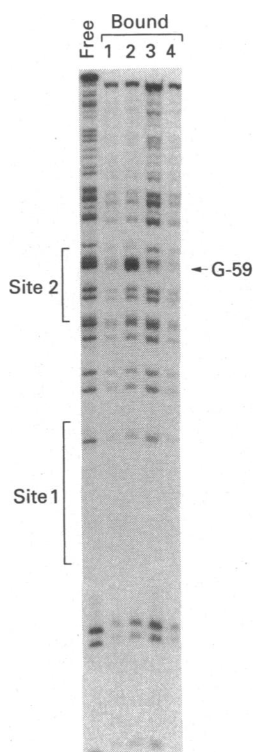

Fig. 6. Methylation interference with MelR binding to KK33 fragment

KK33 DNA was 5′ end-labelled on the top strand and treated with dimethyl sulphate. After gel retardation analysis using purified MelR protein, as in Fig. 5(*a*), DNA from the four bound complexes 1–4 was purified and cleaved with piperidine to locate the sites of modification. The Figure shows analysis of the cleavage pattern using a sequence gel. The sample loaded in the lane marked 'free' was run in the gel retardation assay without any MelR protein, cut out, cleaved with piperidine and then used to calibrate the sequence gel. The position corresponding to the top-strand G, 59 bp upstream of the *melAB* transcription start is identified with an arrow. The locations of MelR binding sites 1 and 2 are indicated.

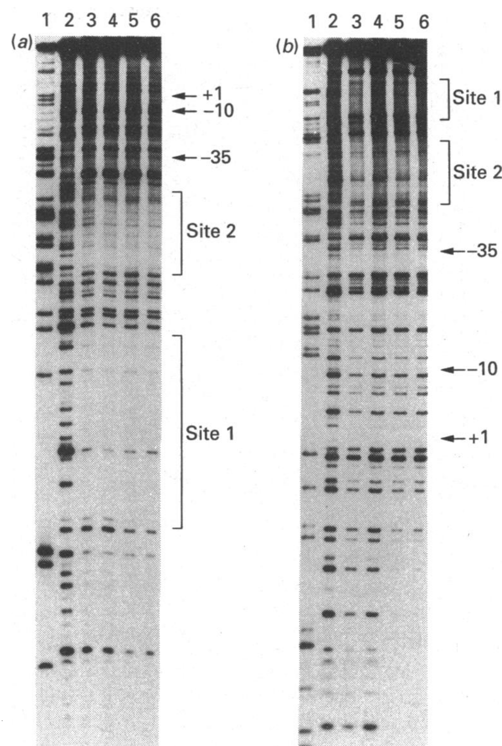

Fig. 7. DNAase I footprint analysis of complexes at p*melAB*

Autoradiograms of sequence gel analysis of DNAase I cleavage of the KK33 *Eco*RI–*Hin*dIII fragment 5′ end-labelled on the top (*a*) or bottom (*b*) strand. The experiments described by each lane are as follows: lane 1, G reaction calibration; lane 2, no protein added; lane 3, plus MelR but no melibiose; lane 4, plus MelR plus melibiose; lane 5, plus MelR plus RNA polymerase but no melibiose; lane 6, plus MelR plus RNA polymerase plus melibiose. The calibrations indicate positions with respect to the transcription start, and brackets indicate the positions of MelR-binding sites 1 and 2.

Mutational analysis of MelR-binding site 1

A complementary genetic approach was taken to identify base pairs important for MelR binding to site 1. Using 'doped' synthetic oligodeoxynucleotides we made a set of pKK33 derivatives carrying random mutations at different positions from base pairs −91 to −110, which contain MelR-binding site 1 (Fig. 4). Single point mutations were obtained at 13 different positions shown in Fig. 4(*b*). Additionally, three further derivatives contained two point mutations, and a fourth carried a single change plus a 1 bp deletion at position −95 (Fig. 4*c*).

To assess the effects of the different mutations on MelR binding, *Eco*RI–*Hin*dIII fragments from pKK33 carrying the different changes were cloned into pRW2 and melibiose/MelR-dependent β-galactosidase expression from p*melAB* was measured exactly as in the experiment described in Table 1. Surprisingly, only two out of the 13 single point mutations resulted in a measurable decrease in p*melAB* activity. These mutations were a CG to GC transversion at −104 (p104G) and a TA to GC transition at −110 (p110G), 1 bp upstream of the 18 bp sequence that is common to sites 1 and 2.

The experiment in Table 1 showed that total deletion of MelR-binding site 1 causes a 4-fold reduction in melibiose-dependent expression from p*melAB*. Thus any mutation in the KK33 promoter fragment that abolishes MelR binding to site 1 will reduce p*melAB* expression to that found with the KK34 fragment (Fig. 1 and Table 2). By this criterion, the combination of the deletion at −95 plus a mutation at −107 (Δ95 p107G) prevented MelR binding, as did p94T/p97G. The double mutations p96G/p104T lowered p*melAB* activity 2-fold, whereas p94T/p100T had no effect. The data in Table 2 show that the single mutation p104G results in the same 2-fold decrease in expression as p96G/p104T. It is striking that no single point mutation reduced expression to that found with the pKK34 fragment (carrying site 2 only), except for p110G.

Binding of MelR to fragments carrying site 1 and site 2

Using gel retardation assays, we compared the binding of purified MelR to the *Eco*RI–*Hin*dIII fragments from pKK33 and pKK35. The KK33 fragment carries both MelR binding sites, whereas the KK35 fragment carries just site 1 (Fig. 1). The results in Fig. 5 show that, whereas increasing concentrations of MelR resulted in a single retarded band with the KK35 fragment, a series of four retarded bands (denoted bound 1–4 in Fig. 7*a*) of decreasing mobility appeared with the KK33 fragment. The appearance of these bands was not affected by the presence or absence of melibiose (results not shown). Bands 1 and 2 appear to be intermediates, presumably corresponding to DNA molecules with just one site occupied.

To determine the nature of the different bound complexes with the KK33 fragment, we exploited the observation that methylation of bottom strand G-104, and by inference top strand G-59, reduced MelR binding. Purified fragments end-labelled on either the upper or the lower strand were methylated with

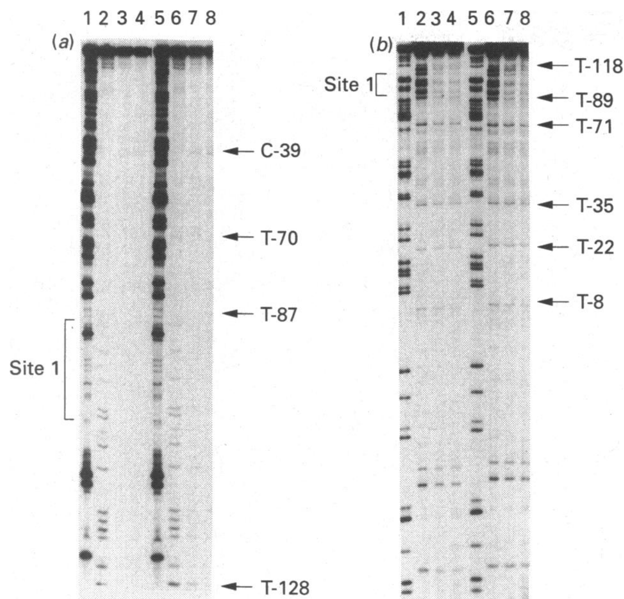

Fig. 8. Permanganate footprint analysis of complexes at p*melAB*

Autoradiograms of sequence gel analysis of 5′ end-labelled KK33 *Eco*RI–*Hin*dIII fragments after incubation with different proteins and treatment with permanganate and piperidine. (*a*) Top strand, (*b*) bottom strand. Lanes 1 and 5 are G reaction sequence calibrations. Permanganate treatments of fragments were as follows: lanes 2 and 6, fragment only with no added protein; lanes 3 and 7, fragment plus MelR protein; lanes 4 and 8, fragment plus MelR plus RNA polymerase. Melibiose was included in the samples used in lanes 6, 7 and 8, but was absent from the samples used in lanes 2, 3 and 4.

dimethyl sulphate as before, and used in gel retardation assays. After autoradiography of the gel, the bands corresponding to free DNA and retarded complexes 1–4 were cut out and analysed to find the sites of methylation. Fig. 6 shows the result with fragment labelled on the top strand. The band resulting from methylation of G-59 was enhanced in the DNA from bound band 2, but not in the other complexes. The clear implication of this is that bound bands 1, 3 and 4 carry MelR bound at G-59 (i.e. bound at site 2), but bound band 2 carries MelR just at site 1 (i.e. methylation of G-59 does not affect this complex). In a parallel experiment (results not shown) with fragment labelled on the lower strand we found that methylation of lower-strand G-104 was enhanced in bound band 1 but not in the other complexes, implying that MelR is bound at site 1 (including G-104) in complexes 2, 3 and 4, but not in complex 1.

Taken together, the results in Figs. 5 and 6 suggest that bound band 1 is due to MelR binding at site 2, bound band 2 is due to MelR binding at site 1, and that bands 3 and 4 carry MelR at both sites. The relative intensities of bound bands 1 and 2 at low concentrations of MelR show that the MelR binding to site 2 is weaker than to site 1, despite both containing the same 18 bp core sequence.

Footprint analysis of MelR binding to KK33 fragments

Fig. 7 shows the result of DNAase I footprinting experiments to investigate the binding of purified MelR to the KK33 fragment either in the presence or the absence of RNA polymerase. Fig. 8 shows a similar experiment in which permanganate was used to probe for unwinding or distortion in the DNA duplex. We used conditions that had been previously used for the formation of ternary complexes between the *E. coli* cyclic AMP receptor protein, RNA polymerase and *gal* or *lac* promoter DNA (Spassky *et al.*, 1984). The results clearly show that MelR binds to both site 1 and site 2 in the presence or absence of melibiose; at lower concentrations of MelR a greater occupancy of site 1 can be observed (results not shown). However, this experiment provides no evidence for any co-operative interactions between MelR and RNA polymerase, and the principal effect of RNA polymerase is to cause protection of sequences downstream of +1. In the experiments with permanganate (Fig. 8), no RNA polymerase-induced unwinding of sequences around the p*melAB* transcription start site was detected. However, on the top strand the modification of C-39 was clearly induced by the presence of MelR and melibiose, suggesting some distortion of this region.

A surprising feature of the permanganate footprint is the preferential attack of bases just upstream of MelR-binding site 1 in the absence of any MelR protein or RNA polymerase (lanes 2 and 6 of Figs. 8*a* and 8*b*). This modification was suppressed by the addition of MelR, suggesting that the region just upstream of MelR-binding site 1 adopts a distorted conformation that allows base modification by permanganate.

DISCUSSION

Sequence-specific binding by MelR protein

MelR is a sequence-specific DNA-binding protein which is essential for melibiose-dependent induction of transcription from the *melAB* promoter. The core of the recognition sequence appears to be an 18 bp element, two copies of which are present at p*melAB*; the two elements are arranged as inverted repeats separated by 20 bp. Interestingly, the related proteins AraC and RhaR also bind to tandem elements of similar length (20 bp for RhaR and 18 bp for AraC; Hendrickson & Schleif, 1985; Hamilton & Lee, 1988; Tobin & Schleif, 1990). At the *E. coli rhaRS* promoter, the two RhaR-binding elements are organized in an inverted repeat separated by 17 bp, whereas at the *araBAD* promoter the two 18 bp AraC-binding elements are arranged as direct repeats. A striking common feature of MelR, RhaR and AraC is that they all bind to their cognate operators at target promoters in both the presence and the absence of inducer.

Our results both *in vivo* and *in vitro* show that a single 18 bp binding element is sufficient for specific binding of MelR, and that this is sufficient for some transcription activation. From parallels with AraC and RhaR, and because the 18 bp sequence contains no inverted or direct repeats, it is likely that each 18 bp element is recognized by one MelR subunit. This raises the question of how a protein 'sees' such an extended sequence. One possibility is that each subunit contains two helix-turn-helix motifs and each motif is responsible for specific contacts with one part of the 18 bp element. Although this cannot be ruled out, at present there is no firm evidence for this (see Caswell *et al.*, 1992), and MelR (and presumably other members of the AraC family) could contain a different type of DNA-binding motif. For this reason we attempted to locate at least some of the bases with which MelR makes close contacts. Since the 18 bp sequence is repeated perfectly at p*melAB*, we supposed that changes at most positions would alter MelR binding sufficiently to affect activation at p*melAB*. We obtained single point mutations at 11 of the positions in the 18 bp element but, surprisingly, only one of these changes, at position −104, reduced MelR-dependent activation (Fig. 4 and Table 2). Gel retardation assays using purified MelR and DNA fragments carrying the different mutations confirmed that MelR binding is reduced by the p104C mutation, but not by the other single changes in the 18 bp element (J. Williams, C. Webster & S. Busby, unpublished work). The chemical interference experiment shown in Fig. 3 provided evidence that methylation of G-104 reduces MelR binding, most

likely because the base is directly involved in the MelR–DNA contact (Siebenlist *et al.*, 1980). The remarkable concurrence between the mutational analysis and chemical modification studies provides clear evidence that the base pair at −104 is important for MelR binding to site 1. The interference experiment in Fig. 6 strongly suggests that the symmetric base pair at −59 is important for MelR–site-2 interactions.

Clearly, the important contacts determining MelR-binding specificity are unlikely to be limited to one base pair and it is probable that some of the bases missed in the random mutagenesis of the 18 bp element are also crucial. Presumably, contacts are spread throughout the 18 bp element; there is some evidence for the importance of base pairs at −94 and −97 from the effects of different combinations of mutations and from the hydrazine modification experiments (Fig. 3). However, the most striking result concerns the effects of the p110G mutation, 1 bp upstream of the 18 bp repeated element. We had deliberately extended the mutagenesis to include base pairs at −91 and −110, immediately downstream and upstream respectively of the 18 bp element, expecting mutations at these positions to have no effect and to be our controls (Fig. 4 and Experimental section). In fact, the mutation at −110 had a greater effect than any of the other changes that we studied. Recall that eliminating MelR binding to site 1 reduces melibiose-induced p*melAB* expression to that found with the pKK34 fragment carrying just site 2 (Tables 1 and 2). The results in Table 2 show that the p110G mutation reduces the expression to below that found with the KK34 fragment. Gel retardation studies show that purified MelR protein binds to fragment carrying p110G but results in a smeared bound band with anomalous mobility (J. Williams, C. Webster & S. Busby, unpublished work). The simplest explanation for this is that the p110G mutation does not stop MelR binding to site 1, but results in a conformation that hinders the action of melR at site 2. There is both direct and indirect evidence for involvement of the sequence just upstream of site 1 in some kind of anomalous conformation. Firstly, despite the fact that sites 1 and 2 have identical 18 bp sequences, MelR binds to site 1 better than to site 2 (Figs. 5 and 7); perhaps this is facilitated by the sequences just upstream of site 1. Secondly, thymines just upstream of site 1 appear especially susceptible to attack by potassium permanganate (Fig. 8), a reagent that only modifies bases in single-stranded or distorted DNA (Borowiec *et al.*, 1987). Finally, we noticed that the 11 bp sequence at the upstream end of site 1 [5′ TAATTTATTCC 3′ from positions −115 to −105 (Fig. 4)] is identical to the sequence from base pairs −52 to −42 at the *E. coli galP*1 promoter. This is located at the upstream end of the binding site for the cyclic AMP receptor protein, which activates *galP*1 (Taniguchi & de Crombrugghe, 1983). During transcription initiation at *galP*1, it is known that this sequence is involved in bending and wrapping (Busby & Buc, 1987; Schultz *et al.*, 1991). Further, we have previously shown that this region in *galP*1 is susceptible to attack by permanganate (Chan *et al.*, 1990). A plausible scenario is that MelR binding to site 1 is followed by some kind of bending, or at least a conformation change, in upstream sequences which is necessary for transcription activation. Note that there is increasing evidence that such distortions are important for the activity of a large number of promoters (Travers, 1991, and references therein). We propose that the p110G mutation disrupts this change.

Melibiose-dependent transcription activation and organization of regulatory elements

Expression from the *melAB* promoter *in vivo* is totally dependent on the inclusion of melibiose in the medium, and yet, *in vitro*, we could demonstrate no effects of melibiose on MelR interactions with target DNA. From our knowledge of the *ara* and *rha* systems, it is perhaps not surprising that MelR binds to its target in both the absence and the presence of melibiose. However, our inability to demonstrate MelR-dependent transcription initiation at p*melAB in vitro* suggests that there is something missing in our biochemistry. With the AraC protein, the inducer arabinose triggers a redistribution of AraC subunits between different binding sites (Lobell & Schleif, 1990). We could find little evidence for any such alternance with MelR, although at low concentrations the occupancy of site 2 compared with site 1 was slightly increased by melibiose (R. Caswell, unpublished work). In any case, it is unlikely that a melibiose-induced redistribution between sites is crucial for activation, since the residual expression *in vivo* from p*melAB* carrying just binding site 2 is still totally melibiose-dependent. In the case of the *E. coli* MerR protein, another transcription factor (unrelated to MelR) that binds in both the presence and the absence of inducer, activation is due to a subtle inducer-dependent distortion in the DNA (Frantz & O'Halloran, 1990) that alters the geometry of the spacer region at the target promoter (Parkhill & Brown, 1990; Ansari *et al.*, 1992). The appearance of a clear MelR-dependent melibiose-inducible distortion at −39 suggests that MelR may act by a similar mechanism of inducing distortion. Perhaps the induced stress is linked to the distortions in the DNA around site 1. The simplest scenario would be to imagine that RNA polymerase and melibiose/MelR together form a nucleoprotein complex, possibly similar to the type envisaged by Raibaud (1989) at the *E. coli malB* promoters.

Binding of MelR to sites 1 and 2 at p*melAB* is hardly, if at all, co-operative. Gel binding assays show the appearance of four retarded bands, two of which appear to have MelR protein bound at just one of the sites (Figs. 7 and 8). MelR appears to occupy both sites in the two most retarded complexes but, as yet, we do not understand the differences between these (bands 3 and 4 in Fig. 5). The classic textbook view of transcription factors is that binding to target sites at promoters is triggered by ligand-induced conformational changes. Clearly MelR, along with AraC, RhaR and MerR proteins, behave differently in that they must be permanently bound at their targets, and the ligand-induced changes result in the complex somehow switching into an active mode. Permanently anchoring such proteins at their targets cuts out any delays or inefficiencies in target location. However, this mode of operation is only possible for proteins, such as MelR, that act at a very small number of targets. Interestingly, MelR synthesis is totally dependent on the cyclic AMP receptor protein (CRP), a transcription activator which acts at a large number of promoters (Webster *et al.*, 1988). Binding of CRP to target sites is triggered by intracellular cyclic AMP levels which are controlled by the degree of glucose starvation (Ullmann & Danchin, 1983). Interestingly, CRP-binding sites at different promoters vary in their affinity for cyclic AMP–CRP, the consequence of this being that tight binding sites are filled at low cyclic AMP concentrations, whereas weaker sites fill only when intracellular cyclic AMP reaches higher levels (Kolb *et al.*, 1983; Gaston *et al.*, 1989). Since the CRP-binding site at the *melR* promoter is one of the weakest to be characterized to date (Webster *et al.*, 1988), this promoter is activated only under conditions of the most severe metabolic stress, i.e. when cyclic AMP levels are at their highest. This suggests that melibiose is a low-priority energy source. Presumably MelR protein is made only when cells are subjected to the most extreme stresses. MelR then attaches to the binding site at p*melAB* and the cell is then ready should melibiose appear in the growth medium. This scenario suggests that the melibiose operon has evolved to permit *E. coli* to marshall the available resources with the maximum economy.

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

Ansari, A., Chael, M. & O'Halloran, T. (1992) Nature (London) **355**, 87–89
Borowiec, J., Zhang, L., Sasse-Dwight, S. & Gralla, J. (1987) J. Mol. Biol. **196**, 101–111
Busby, S. & Buc, H. (1987) Microbiol. Sci. **4**, 371–375
Casadaban, M. & Cohen, S. (1980) J. Mol. Biol. **138**, 179–207
Caswell, R., Williams, J., Lyddiatt, A. & Busby, S. (1992) Biochem. J. **287**, 493–499
Chan, B., Spassky, A. & Busby, S. (1990) Biochem. J. **270**, 141–148
Frantz, B. & O'Halloran, T. (1990) Biochemistry **29**, 4747–4751
Gaston, K., Kolb, A. & Busby, S. (1989) Biochem. J. **261**, 649–653
Hamilton, E. & Lee, N. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 1749–1753
Hanatani, M., Yazyu, H., Shiota-Niiya, S., Moriyama, Y., Kanazawa, H., Futai, M. & Tsuchiya, T. (1984) J. Biol. Chem. **259**, 1807–1812
Hendrickson, W. & Schleif, R. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 3129–3133
Jayaraman, P.-S., Cole, J. & Busby, S. (1989) Nucleic Acids Res. **17**, 135–145
Kolb, A., Spassky, A., Chapon, C., Blazy, B. & Buc, H. (1983) Nucleic Acids Res. **11**, 7833–7851
Lobell, R. & Schleif, R. (1990) Science **250**, 528–532
Maniatis, T., Fritsch, E. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
O'Halloran, T., Frantz, B., Shin, M., Ralston, D. & Wright, J. (1989) Cell **56**, 119–129
Parkhill, J. & Brown, N. (1990) Nucleic Acids Res. **18**, 5157–5162
Raibaud, O. (1989) Mol. Microbiol. **3**, 455–458
Schultz, S., Shields, G. & Steitz, T. (1991) Science **253**, 1001–1007
Siebenlist, U., Simpson, R. & Gilbert, W. (1980) Cell **20**, 269–281
Spassky, A., Busby, S. & Buc, H. (1984) EMBO J. **3**, 43–50
Taniguchi, T. & de Crombrugghe, B. (1983) Nucleic Acids Res. **11**, 5165–5170
Tobin, J. & Schleif, R. (1990) J. Mol. Biol. **211**, 75–89
Travers, A. (1991) Curr. Opin. Struct. Biol. **1**, 114–122
Ullmann, A. & Danchin, A. (1983) Adv. Cyclic Nucleotide Res. **15**, 1–46
Webster, C., Kempsell, 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