The aldA gene of Escherichia coli is under the control of at least three transcriptional regulators

Ana Limón, Elena Hidalgo† and Juan Aguilar

Expression studies on the aldA gene encoding aldehyde dehydrogenase in Escherichia coli showed induction by two types of molecule (hydroxylaldehydes and 2-oxoglutarate), carbon catabolite repression and respiration dependence. Promoter deletion analysis showed that the proximal operator, which includes inducer-regulator complex and catabolite repression protein (Crp) recognition sites, was necessary for induction by either type of inducer, and that full induction by aldehydes required the cooperation of distal operator sequences beyond position −119. Interactions of the regulator protein with the −59 to −6 fragment were shown by DNA mobility shift assays. Fusions of different deletions of the aldA promoter to lacZ indicated that a Crp site proximal to the transcriptional start point (top) was functional in the cAMP-dependent catabolite repression of this system, whereas a distal control site was likely to operate in a CAMP-independent catabolite repression. DNA mobility shift and footprint analyses showed that only the tsp proximal site was bound by pure Crp with a $K_d$ of $5.4 \times 10^{-7}$ M. As shown by an Arc-defective strain, the aldA gene seems to be repressed by the Arc system under anaerobiosis, displaying its physiological full induction and activity in the presence of oxygen.

**Keywords**: aldehyde dehydrogenase, aldA, regulation, transcription, Escherichia coli

### INTRODUCTION

The gene aldA maps at min 32 (Baldomá & Aguilar, 1988; Chen et al., 1987) of the Escherichia coli chromosome, and has been reported to encode an aldehyde dehydrogenase (ALDH) of 479 amino acids (Hidalgo et al., 1991) functioning on a broad spectrum of substrates. Initially it was assigned the role of oxidizing to L-lactate the L-lactaldehyde (Sridhara & Wu, 1969) formed in the metabolism of L-fucose (Cocks et al., 1974) and L-rhamnose (Baldomá & Aguilar, 1987). Further studies demonstrated that it was also involved in the oxidation of glycolaldehyde formed in the metabolism of different pentoses (LeBlanc & Mortlock, 1971) and in the metabolism of ethylene glycol in mutant cells adapted to grow on this non-natural carbon source (Boronat et al., 1983). In fact, it was shown that the enzyme oxidized other hydroxyaldehydes, including L-glyceraldehyde and the α-oxoaldehyde methylglyoxal (Baldomá & Aguilar, 1987).

The ALDH activity is induced not only by growth on L-fucose, L-rhamnose or D-arabinose, through the formation of intermediate hydroxyaldehydes, but also by growth in the presence of glutamate or amino acids yielding glutamate, with the exception of proline (Quintilla et al., 1991). Induction by these amino acids requires glutamate accumulation. 4-Aminobutyric acid is also able to induce ALDH through its transamination to glutamate. Glutamate induction was shown to be due to its conversion to 2-oxoglutarate, the tricarboxylic acid cycle intermediate with which glutamate is in equilibrium. Consistently with that, growth on 2-oxoglutarate also induces the aldA gene. Both types of inducer, aldehydes and 2-oxoglutarate, share the same
regulatory protein (Quintilla et al., 1991), so far uncharacterized, whose gene has not yet been located. Induction by either of these two compounds, which occurs only aerobically, is repressed by glucose. This catabolite repression is only partially relieved by cAMP, in accordance with the presence of an additional cAMP-independent catabolite repression by carbon sources other than glucose, such as glycerol (Quintilla et al., 1991). Thus the information available indicates that this enzyme is rather non-specific, participates in several metabolic pathways, and responds to multiple regulatory signals.

In this report, we analyse the control of expression of the aldA gene and characterize three different regulatory functions in its promoter: one controlled by the regulator-inducer complex, another by the catabolite repression protein (Crp)–cAMP complex, and the third by the Arc anaerobic repressor.

**METHODS**

**Bacterial strains and plasmids.** The bacterial strains and plasmids used in this study are listed in Table 1.

**Growth conditions and preparation of cell extracts.** Cells were grown aerobically on LB or minimal medium as described previously (Boronat et al., 1981). In this report, we analyse the control of expression of the aldA gene and characterize three different regulatory functions in its promoter: one controlled by the regulator-inducer complex, another by the catabolite repression protein (Crp)–cAMP complex, and the third by the Arc anaerobic repressor.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant genotype</th>
<th>Source/reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECL1</td>
<td>HfrC phoA8 relA1 tonA22 T2R (α)</td>
<td>Lin (1976)</td>
</tr>
<tr>
<td>DHL52F</td>
<td>880 aclZ AM15 (ΔlacZYA–argF) U169 hsdR17(λm59) supE44 recA1 endA1 gyrA96 thi-1 relA1</td>
<td>BRI.</td>
</tr>
<tr>
<td>MC1061</td>
<td>hsdR mcrB araD139 Δ(lacJAB未曾–lacI)7679 galU galK rpsL thi 4316</td>
<td>Meissner et al. (1987)</td>
</tr>
<tr>
<td>RY1000</td>
<td>araD139 Δ(lacIPOZYA) U169 rpsL (str) thi mot recA59</td>
<td>M. Aldea, University of Lleida, Spain</td>
</tr>
<tr>
<td>J111</td>
<td>ECL40 recA1 slrA2: Tn10</td>
<td>Hidalgo et al. (1991)</td>
</tr>
<tr>
<td>ECL368</td>
<td>ECL547 araA131</td>
<td>Mozola &amp; Nash (1993)</td>
</tr>
<tr>
<td>K37</td>
<td>strA galK K2 su+ α− F−</td>
<td>Granston &amp; Nash (1993)</td>
</tr>
<tr>
<td>HN1491</td>
<td>K37 himA::cat</td>
<td></td>
</tr>
<tr>
<td>K2704</td>
<td>K37 himD3::cat</td>
<td></td>
</tr>
</tbody>
</table>
The linear DNA was protected from digestion by the deletion DNA fragments into plasmid pRS550, as described by Simons system, while the 5'-end marked the direction of progress of restriction enzyme and then a 5'-overhanging end was formed by digestion with EcoRI (see Fig. 1). In this way, the 3'-end of the inserts and their junctions were verified by nucleotide sequencing.

Transcriptional fusions were constructed by inserting the DNA fragments into plasmid pRS550, as described by Simons et al. (1987). The plasmids containing the desired deletion were digested with PstI and the resulting inserts (from 400 to 700 bp) were purified. The vector pRS550 was linearized by digestion with BamHI, blunted by treatment with Klenow polymerase and ligated with the inserts by T4 ligase. The direction of the inserts and their junctions were verified by nucleotide sequencing. Plasmid pRS550 carries a cryptic lac operon and genes that confer resistance to both kanamycin and ampicillin. After introduction of the recombinant plasmids into the streptomycin-resistant strain MC1061, blue colonies were isolated on X-Gal plates containing ampicillin, were dried under vacuum and the bands were visualized by autoradiography. Physical mapping and sequence of the aldA promoter

RNA preparations and Northern blot experiments. For preparation of total RNA, cells of a 25 ml culture grown to OD600 0.5 were collected by centrifugation at 5000 g for 10 min and processed as described by Belasco et al. (1985). Northern blot hybridization was performed with each RNA sample (10 μg) following the procedure described previously by Moralez et al. (1993). The probe was 32P-labelled by the random primed method (Sambrook et al., 1989) using as template a 300 bp (PstI-BamHI) aldA internal fragment (see Hidalgo et al., 1991).

DNA mobility shift assay. We adopted the method described by Nunoshiba et al. (1992) with minor modifications for the DNA mobility shift assay. The DNA fragments were end-labelled by filling in with the Klenow fragment of DNA polymerase and [γ-32P]dATP. For the DNA-binding reaction, mixtures (20 μl) contained 10 mM Tris/HCl (pH 7.5), 75 mM KCl, 2 mM diethiothreitol, 10% (v/v) glycerol, 2 μg/ml (dl-dC), 10 fmol labelled fragments and protein extracts, purified Crp protein (a gift from S. Adhya, NIH, USA) and/or RNA polymerase from E. coli (purchased from Pharmacia and diluted in reaction buffer with 1 mM Mg2+)] in the amounts indicated in the experiments. The binding reactions were performed at 25 °C for 15 min and samples were electrophoresed in 5% (w/v) polyacrylamide gels in running buffer (450 mM Tris; 450 mM boric acid, pH 8.3; 1 mM EDTA) for 1-2 h at 350 V.

DNase I footprinting. Fragment 2 (Fig. 3) was PCR-amplified using as template plasmid Bluescript in which fragment 2 had been previously cloned (pBS-F2); as primers T3 and T7 oligonucleotides, one of which had been labelled at the 5'-end with [γ-32P]dATP (1×108 Bq mmol-1; Amersham); and T4 polynucleotide kinase. The PCR product was purified by non-denaturing PAGE and EtBr (Schleicher & Schuell) columns. The Crp-DNA binding reactions were performed as described above. After incubation, MgCl2 and CaCl2 were added to final concentrations of 1.25 mM and 5 mM, respectively. The mixtures were treated with DNase I (5 ng DNase I (Sigma) in 50 mM Tris/HCl, pH 7.5, 10 mM MgSO4, 1 mM diethiothreitol and 50%, w/v, glycerol) at 25 °C for 2 min. The reactions were stopped by addition of 5 μg yeast tRNA ml-1 in 270 mM ammonium acetate, 70% (v/v) ethanol. After precipitation, the DNA pellet was resuspended in 5 μl formamide loading buffer (Sambrook et al., 1989) and the samples were subjected to electrophoresis through an 8% (w/v) denaturing sequencing gel, drying and autoradiography.

RESULTS

Physical mapping and sequence of the aldA promoter

The physical map and sequence of a 1 kb fragment upstream of the ATG start codon of the aldA structural gene previously cloned and sequenced (Hidalgo et al., 1991) was determined. Fig. 1 displays the physical map of clone pALD1.1 encompassing this region, and part of its sequence, which includes 185 nucleotides upstream from the transcriptional start point (tp; position 1) previously established by Hidalgo et al. (1991). Analysis of the sequence revealed several possible regulatory elements able to participate in the control of aldA expression. A -10 hexamer (GTGTAAT) was identified with four nucleotides matching the consensus and an extended TGN sequence directly upstream from these four nucleotides (Kumar et al., 1994), whereas no clear -35 box was apparent. Among the several inverted repeats with high energy of stabilization found, there was one centred at -67, which contains only the upstream half of the Crp consensus. Also an Arc consensus as described by Lynch & Lin (1996) overlaps with the -10 hexamer. Finally, three IHF recognition consensus sequences (Nash & Granston, 1991) were apparent at positions 7 to -6, -55 to -67 and -130 to -142.

Control of aldA gene expression is at the transcriptional level

Total RNA was prepared from cells of the ALDH-deficient strain JA111 transformed with plasmid pALD1.1 or from strain ECL1 grown under inducing and non-inducing conditions and analysed by Northern blotting with an aldA internal labelled probe. The probe hybridized to a transcript of 1.5 kb when the RNA
A. LIMÓN, E. HIDALGO and J. AGUILAR

Physical map of plasmid pALD1.1 with the restriction sites and the location of the aldA open reading frame (thick arrow) in the insert (thin line) cloned in the vector Bluescript (thick line). The sequence of the promoter region between positions -185 and +45 is shown in the expansion below. The ATG initiation codon is at the 3'-end of the sequence, the Shine–Dalgarno (SD) ribosome recognition site is boxed and the tsp is shown by a black arrowhead. The inverted repeat of highest energy of stabilization is indicated by a thin arrow. The -10 box is labelled with the extended sequence in italics. The sequences matching the consensus for Crp are in bold, the extent of the Crp footprinting is between square brackets, the IHF sequences are underlined and the sequence matching the ArcA-binding consensus is shown by asterisks. Numbers indicate the nucleotide position, taking tsp as position +1.

Control of aldA gene expression by the activator

Results presented in Fig. 3 showed that full induction by growth on rhamnose requires up to position -170, while full induction by growth on casein acid hydrolysate requires sequence only to position -93, suggesting a different mechanism for each of the two inducing conditions. Constructions including sequences upstream of position -193 reduced the full induction by casein acid hydrolysate by 40%, to values similar to those of wild-type aldA promoter function (pALD1.1). Expression of aldA was completely abolished by anaerobiosis under all of the conditions tested. Finally, for all deletions, growth on rhamnose plus glucose led to loss of induction, while growth on rhamnose plus glycerol gave partial loss of the induction displayed by the deleted promoters containing the distal control region (up to position -170). The construction with the -93
Regulation of the \textit{aldA} gene of \textit{E. coli}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{Expression of the \textit{aldA} gene from pALD1.1 and deletion analysis of the \textit{aldA} promoter region. A graphic representation of the insert (thick line) cloned in the Bluescript vector (not shown) is presented in the middle part. \textit{tsp} (+1), the \textit{−10} box (−10), the Shine-Dalgarno box (SD) and the translation initiation codon ATG (43) of the \textit{aldA} structural gene (ALD) indicated by the arrow are marked on the insert. Putative binding sites for Crp, ArcA, IHF and the upstream regulatory site (URS), as well as fragments f1, f2, f3 and f5 used for retardation experiments, are presented in the upper part by boxes and lines with numbers indicating their limits. Deleted constructions of the promoter region that were cloned into the same vector and assayed for ALDH activity are shown below by solid arrows. Numbers next to these arrows indicate the 3'-end of the deleted DNA represented by the open bars. The specific ALDH activity associated with each construction is indicated on the right. ALDH activities were assayed from transformed cells grown under the following conditions: aerobically on rhamnose (Rha + O$_2$), casein acid hydrolysate (Caa + O$_2$), rhamnose plus glucose (Rha + Glu + O$_2$) and rhamnose plus glycerol (Rha + Gly + O$_2$), or anaerobically on rhamnose (Rha − O$_2$). The bottom line displays the activities obtained with the entire insert contained in plasmid pALD1.1.}
\end{figure}

fragment, which is not inducible by rhamnose, presented some induction either by glycerol or glucose. The reason for this unexpected induction in this particular construction is unknown.

Binding of the activator protein to this sequence was proven by gel retardation experiments with $^3$P-labelled fragment 5 (see f5 in Fig. 3). The electrophoretic mobility of fragment 5 (−6 to −59) was retarded by crude extracts of cells of strain ECL1 grown aerobically on rhamnose (Fig. 4a). To determine the sequence specificity of binding, 100-fold excess of unlabelled DNA fragment was mixed as the competitor with the labelled fragment. The major retarded band was markedly reduced by addition of the unlabelled fragment. However, there was no competition by poly(dI-dC). Furthermore, the major band was found to be proportional to the concentration of protein in the binding reaction. The retarded band with lower mobility (band labelled N) corresponded to RNA polymerase−fragment complex as indicated by the enhancement of this band upon addition of pure RNA polymerase to the mixture (not shown).

Binding of the regulator to the DNA fragment containing the \textit{aldA} control region was further characterized by gel retardation by studying the effect of addition of inducer molecules to extracts of cells grown under either inducing (rhamnose) or non-inducing (glucose) conditions (Fig. 4b). Extracts of glucose-grown cells displayed a retarded band (Fig. 4b, lane 3) due to endogenous 2-oxoglutarate formed in the metabolism. Addition of glycolaldehyde increased this major retarded band R, and revealed a fainter retarded band G, probably due to a multimeric complex (Fig. 4b, lanes 4 and 5). Addition of 2-oxoglutarate with or without glycolaldehyde strongly increased the same major retarded band R, but did not reveal the G band (Fig. 4b, lanes 6 and 7), presumably due to regulator capture in the R band.

\textbf{Crp–CAMP control of \textit{aldA} gene expression}

To assess the putative function of the Crp site centred at −60 and that of the upstream regulatory site (URS in Fig. 3) centred at −134, three deletions were fused to the
**Fig. 4.** Electrophoretic mobility shift of aldA promoter DNA fragment f5 by extracts of strain ECL1 grown aerobically on rhamnose (a). The effect of the addition of the inducers glycolaldehyde and 2-oxoglutarate to extracts of strain ECL1 grown aerobically on rhamnose or glucose is also presented (b). Free (D), regulator-bound (R), regulator-bound in the presence of glycolaldehyde (G) and RNA-polymerase-bound (N) bands after autoradiography are indicated. The composition of the binding reaction run in every lane is indicated at the base: f5*, 10 fmol labelled fragment f5; f5, 20 ng unlabelled fragment 5; poly(dI-dC), 2 pg unlabelled synthetic oligonucleotide; extract (Rha), 1 or 2 pg, as indicated, crude protein extract of cells grown aerobically on rhamnose; extract (Glu), 1 pg crude protein extract of cells grown aerobically on glucose; and glycolaldehyde or 2-oxoglutarate, 50 mM of these inducer molecules.

 lacZ gene to generate plasmids pRS550 Δ−35, pRS550 Δ−119 and pRS550 Δ−287. As described in Methods, all fusions were assayed as single-copy lysogens in strain RYC1000 to avoid potential artifacts due to multiple copies of the fusions in each cell.

 Cultures of the lysogens on rhamnose plus glucose medium were grown to exponential phase and then reinoculated into the same medium with or without 5 mM cAMP. Samples were taken every hour to analyse the β-galactosidase activity. Fig. 5 shows that strain RYC1000::ARS Δ−35 displayed no activity while RYC1000::ARS Δ−119 or RYC1000::ARS Δ−287 recovered high levels of β-galactosidase activity after 4–5 h incubation in the presence of cAMP, indicating that at least the Crp site between positions −47 to −74 was functional. Nevertheless, these activities were about half those obtained by growth on rhamnose (not shown), which is consistent with the previously described cAMP-independent catabolite repression (Quintilla et al., 1991).

 The response to Crp of the tsp distal control site was further ruled out in favour of the proximal copy of the Crp sequence by gel retardation experiments with 32P-labelled fragments that contained either of the Crp-like boxes. It was demonstrated that the electrophoretic mobility of fragment 3 (−111 to −276) was not affected by incubation with pure Crp in the presence of 20 μM cAMP. Furthermore, a synthetic fragment (−121 to −150) with the distal Crp site was not retarded, indicating that it did not bind Crp–cAMP pure complex (not shown). In contrast, the electrophoretic mobility of fragment 2 (−3 to −184) was significantly retarded by incubation with Crp–cAMP complex (Fig. 6, lane 2). Specificity of the sequence was ascertained by adding poly(dI-dC) (Fig. 6, lane 3) or unlabelled fragment 2 (Fig. 6, lane 6). Addition of pure RNA polymerase permitted the identification of the retarded band (N) corresponding to the DNA–RNA polymerase complex formed with fragment 2 in the absence (Fig. 6, lane 4) or in the presence (Fig. 6, lane 5) of the Crp–cAMP complex.

 The dissociation constant of the Crp–cAMP complex for fragment 2 was determined by incubating 0.5 ng labelled fragment 2 with concentrations of Crp ranging from 0.2 to 1.6 μM in the presence of 20 μM cAMP. The intensity of retarded bands was measured by densitometry and plotted against the concentration of Crp present in the reaction mixture (Fig. 7). A K_d of 5·4 × 10^{-7} M was determined in this way.

 The region of the Crp binding was determined more precisely by DNase I protection. The 5'– and 3'–end-labelled Crp sequences were partially digested with DNase I in the presence or absence of Crp–cAMP complex and the digest was analysed by electrophoresis.
Regulation of the aldA gene of *E. coli*

Fig. 5. Time-course of β-galactosidase activity of strains RYC1000::ARS Δ−35 ( ▼, ■) and RYC1000::ARS Δ−119 ( ▲, ) grown on rhamnose plus glucose in the absence (open symbols) or presence (filled symbols) of 5 mM cAMP. A rhamnose plus glucose-grown inoculum was used to start 200 ml of each of the three cultures, and 10 ml samples were taken at 1 h intervals for determination of β-galactosidase activities.

Fig. 6. Electrophoretic mobility shift of aldA promoter DNA fragment f2 by Crp. Free (D), Crp–cAMP-bound (C) and RNA-polymerase-bound (N) bands after autoradiography are indicated. The composition of the binding reaction run in every lane is indicated at the base: f2*, 10 fmol labelled fragment f2; f2, 20 ng unlabelled fragment f2; poly(dl-dC), 2 µg unlabelled synthetic oligonucleotide; Crp, 20 pmol pure Crp; cAMP, 400 pmol cAMP; and RNApol, 500 fmol pure RNA polymerase.

Aerobic/anaerobic control of aldA gene expression by Arc

The description of the ALDH encoded by aldA as a predominantly aerobic enzyme led us to study the control of its expression in the presence or absence of oxygen. Since a 10-fold increase in enzyme activity was observed in response to oxygen availability, we examined whether the ArcA regulatory protein was involved in this control (Iuchi & Lin, 1988). A strain containing a mutation in arcA (strain ECL968) as well as its parental strain (ECL547) were made lac-negative by mutagenesis with diethyl sulphate (Roth, 1970), and transformed with plasmid pRS550 Δ−287 containing lacZ fused to the aldA promoter. These strains were analysed following aerobic or anaerobic growth on rhamnose minimal medium. The expression of the aldA–lacZ fusion in the arcA mutant strain grown under anaerobic conditions (18 700 units of β-galactosidase activity) rose to the level of the wild-type parental strain, reaching the same level obtained in both strains under aerobic conditions (19375 units in the wild-type strain and 20750 units in the mutant strain). These findings suggest that ArcA acts as a negative regulator of aldA gene expression under anaerobiosis. The consensus for ArcA recognition, TGTTAATTAA, as determined by Lynch & Lin (1996), was indeed found between positions −13 and −4 of the aldA promoter with a conservation of 10/10 nucleotides.

Binding of IHF to the aldA promoter

The promoter sequence presented above displayed three copies of a rather conserved IHF recognition sequence (positions 7 to −6, −55 to −67 and −130 to −142). Binding of the IHF, formed by himA and himD (himD) gene products, to these sequences was studied by gel retardation experiments with 32P-labelled fragments 1, 2 or 3 containing the IHF recognition sequences (Fig. 1). Electrophoretic mobility of any of the three fragments was significantly retarded after incubation with 100, 50 or 30 ng IHF purified protein (not shown). Specificity of the sequence was ascertained by adding the nonspecific competitor poly(dl-dC) or the specific competitor unlabelled fragment. However, a himA mutant (strain HN1491) and a himD (himD) mutant (strain K2704), previously made lac-negative by mutagenesis with diethyl sulfate (Roth, 1970) and transformed with plasmid pRS550 Δ−287 containing the aldA–lacZ fusion, had no significant effect on aldA expression.
**DISCUSSION**

It is well established that the ALDH encoded by the aldA gene oxidizes different hydroxyaldehydes and serves several metabolic pathways (Baldomà & Aguilar, 1987).

The wide spectrum of catalytic activity and metabolic function is not uncommon to most ALDHs and is probably the cause of the requirement for multiple regulatory mechanisms. These include: (i) at least two regulatory mechanisms for two classes of inducer
molecules sharing a common regulator protein; (ii) a cAMP-dependent and a cAMP-independent catabolite repression; and (iii) an aerobic control by the ArcAB regulatory system. It is not clear whether IHF regulates \textit{aldA} expression.

Two inducing mechanisms are proposed on the basis of the following observations. Full induction by rhamnose, hence by aldehydes such as lactaldehyde, required sequences of the promoter up to $-170$, whereas full induction by casein acid hydrolysate, hence by 2-oxoglutarate, required only sequences up to $-93$. This is in accordance with two DNA-regulator interactions that could be mediated by one or two regulator proteins. Previous characterization of induction of \textit{aldA} by different inducers suggested a single regulator model for which the two kinds of inducer would compete (Quintilla et al., 1991).

Our results reveal the constitutive presence of regulator protein in the extracts and its ability to bind the proximal promoter when the inducer is complexed either \textit{in vivo} or \textit{in vitro}. The results are consistent with a higher affinity of the regulator-inducer complex for DNA when the inducer is 2-oxoglutarate instead of glycolaldehyde.

Band shift assays confirmed that the tsp proximal operator overlaps with the RNA polymerase binding site, a situation common for repressor controls but not for activator controls like the one found in the \textit{aldA} gene. In these and other cases, RNA polymerase binding for transcription initiation is synergistic with the binding of other factors such as regulator proteins or Crp (Adhya et al., 1993). It has been proposed that these factors may contribute to the formation of an RNA polymerase open complex in promoters with weak $-10$ and $-35$ sequences, as exemplified by the galactose operon promoter (Chan et al., 1990), or with abnormal distances between the two boxes (Hidalgo & Demple, 1994; Summers, 1992). In this context, it has also been proposed by Kumar et al. (1994) that formation of the RNA polymerase complex, in the absence of the $-35$ box, might be facilitated by the extended $-10$ hexamer found in these weak promoters.

The interaction with the distal sequences required for full induction by aldehydes (growth on rhamnose) is not understood. It is likely that interactions with the distal regulatory site, centred at $-134$, could play a role, especially taking into account the requirement of sequences upstream of position $-119$ for full induction. This could be mediated either by interactions with another protein regulator factor or by induced conformational changes of the promoter DNA. In this context, bending of DNA by Crp interaction, as described for other systems such as the gal promoter (Lavigne et al., 1992), may influence the mechanism of transcription initiation. Furthermore, curved sequences could also account for uncharacterized regulatory effects (decrease in induction) caused by the presence of sequences upstream of position $-193$ (Fig. 3). The cooperation of this distal upstream regulatory site for induction would be competed for by the cAMP-independent repressor formed in the presence of glycerol (Fig. 3). Absolute requirement of the proximal Crp site, centred at $-60$, for induction by either of the two types of inducer molecule would explain total repression by glucose under all conditions.

Catabolite repression of \textit{aldA} expression has been well-documented in studies of enzyme levels under different physiological conditions. Here the repression at the transcriptional level is confirmed by the use of \textit{lacZ} fusions to the \textit{aldA} promoter. Our promoter deletion analysis, mobility shift assays and DNase protection assays seem to rule out the upstream control site centred at $-134$ as a Crp site and confirm the function of the Crp site centred at $-60$. At this site, the DNase-I-protected sequences of each strand displayed an overlap of 28 nucleotides, which included the Crp consensus described by de Crombrugghe et al. (1984), with nine of the ten nucleotides of the palindrome identical and a common six nucleotide spacer. Affinity of Crp for this site was one order of magnitude lower than other examples described in the literature (Barber et al., 1993). However, several Crp-dependent promoters have a $K_d$ value similar to the one described here for \textit{aldA} (Kolb et al., 1993).

The inability to restore full expression in the presence of cAMP in the fusion experiments with the complete promoter is in accordance with the presence of a cAMP-independent catabolite repression, as described by Ullmann et al. (1976) and revised by Kolb et al. (1993).

Another level of control of \textit{aldA} expression is suggested by the absolute oxygen requirement for ALDH activity. No enzyme is detected by catalytic activity or immuno-detection under anaerobic conditions (Baldoma & Aguilar, 1988). Arc regulatory protein seems to supply such a control, as indicated by the anaerobic increase in $\beta$-galactosidase activity in \textit{aldA} promoter–\textit{lacZ} fusion strains defective for Arc function and hence unable to repress under anaerobiosis.

Three IHF sites were found in the \textit{aldA} promoter which, according to DNA mobility shift experiments, bind pure IHF regulatory protein. No significant effect of IHF on \textit{aldA} expression was detected under our conditions, although the possibility of its widely accepted role as stabilizer of DNA conformation remains open (Nash & Granston, 1991). The large number of IHF-binding sites spread over the \textit{E. coli} genome, the lack of many strong phenotypes associated with gene expression in IHF mutants (Friedman, 1988) and the ability of IHF to bend and compact DNA (Schmid, 1990) are all consistent with this hypothesis.

ACKNOWLEDGEMENTS

We thank S. Adhya and H. A. Nash for purified proteins, M. Aldia and E. C. C. Lin for bacterial strains and R. Rycroft for editorial assistance. This work was supported by grant PB94-0829 from the Direccion General de Investigacion Cientifica y Tecnica, Madrid, Spain, and by the help of the ‘Comissionat
per Universitats i recerca de la Generalitat de Catalunya’. A.L. was the recipient of a predoctoral fellowship (FPI) from the Ministerio de Educación y Ciencia of Spain.

REFERENCES


Received 28 October 1996; revised 3 February 1997; accepted 7 February 1997.