Binding of the *Escherichia coli* cyclic AMP receptor protein to DNA fragments containing consensus nucleotide sequences

Kevin GASTON,* Annie KOLB† and Stephen BUSBY*‡

* School of Biochemistry, University of Birmingham, Birmingham B15 2TT, U.K., and † Département de Biologie Moléculaire, Institut Pasteur, 75724 Paris Cédex 15, France

Binding of the *Escherichia coli* CRP protein to DNA fragments carrying nucleotide sequences closely corresponding to the consensus is very tight with a dissociation time of over 2 h in our conditions. The concentration of cyclic AMP required for this binding is below the physiological range of intracellular cyclic AMP concentrations. Changes in nucleotide sequence at positions that are not well-conserved between different naturally-occurring CRP sites allow a more rapid dissociation of CRP–DNA complexes. There is an inverse correlation between the stability of CRP binding to sites *in vitro* and the repression by glucose of expression dependent on these sites *in vivo*: expression that is dependent on the tighter binding sites cannot be repressed by the inclusion of glucose in the growth medium.

INTRODUCTION

In *Escherichia coli*, increases in intracellular cyclic AMP levels are produced in response to glucose starvation, and are used to trigger the expression of gene products involved in the catabolism of alternative carbon substrates (reviewed by Ullmann & Danchin, 1983). Cyclic AMP binds to the cyclic AMP receptor protein (CRP), which undergoes a conformational change such that the CRP–cAMP complex binds to specific sites near catabolite-sensitive promoters. Binding of the complex can increase the rate of transcription initiation at these promoters (reviewed by de Crombrughe *et al.*, 1984). During growth of *E. coli* in media containing glucose, the intracellular level of cyclic AMP drops and CRP–cAMP-stimulated transcription is suppressed.

Comparison of a number of CRP–cAMP binding sites shows that the complex covers at least 22 base pairs and suggests the consensus sequence 5′ aaaTGTGAtctagaTCACAtt 3′, where the upper and lower case letters denote strongly and weakly conserved nucleotides respectively (Ebright *et al.*, 1984a; Berg & Von Hippel, 1988; Jayaraman *et al.*, 1989). In a recent study we introduced a CRP–cAMP binding site closely resembling this consensus sequence into the *E. coli* galactose (*gal*) promoter region, replacing the naturally occurring site (Gaston *et al.*, 1988). We found that CRP–cAMP bound to this site with an affinity that was at least 100-fold higher than to the original site. As expected, CRP–cAMP binding to the high-affinity site caused a stimulation of expression from the *galP1* promoter *in vivo* but surprisingly this stimulation was not suppressed by high concentrations of glucose in the growth medium. Here we show that the concentration of cyclic AMP required to trigger CRP binding to this site *in vitro* is very low. Moreover, the complex is exceptionally stable with a dissociation time of more than 2 h in our conditions, suggesting that CRP–cAMP is permanently anchored to this site *in vivo* in all growth conditions. However, point mutations at some of the

weakly conserved nucleotides in this tight-binding site cause large reductions in the stability of the complexes.

MATERIALS AND METHODS

The DNA fragments used in this work were purified from derivatives of plasmid pAA121 (Kelsall *et al.*, 1985) carrying the *galP1* promoter on a 107 bp *Bam*HI–*Hind*III fragment as illustrated in Fig. 1. Derivatives were grown in a Δ *crp* derivative of the strain M182 (Busby *et al.*, 1983) and plasmid DNA and fragments were purified using standard protocols (Maniatis *et al.*, 1982). By convention *gal* promoter sequences are numbered with the *galP1* transcription start as +1. The inserts used in this work all contained an *Sph*I site from bp –26 to –31, a *Bam*HI linker sequence upstream of –57 and a *Hind*III linker downstream of +45. The construction of these inserts, which are shown schematically in Fig. 1, was fully described in the Materials and methods section and Fig. 5 of Gaston *et al.* (1988). Synthetic *Bam*HI–*Sph*I fragments carrying derivatives of the CRP consensus binding sequence were bought from Alta Bioscience of the University of Birmingham and were cloned directly between the *Bam*HI–*Sph*I sites of the plasmid shown in Fig. 1. Using the numbering suggested by Lehming *et al.* (1988), different positions in any CRP site are numbered 1–11 leftward or rightward from the centre of symmetry (see Fig. 4). All nucleotide sequences were checked by the Maxam–Gilbert chemical method (Maniatis *et al.*, 1982). The 203 bp *lac* fragment was used as previously described by Kolb *et al.* (1983a).

Gel binding assays using labelled fragments and purified CRP were performed exactly as before (Gaston *et al.*, 1988). In the experiment shown in Fig. 2 cyclic AMP was omitted from the gel and running buffer. DNA fragments carrying *galP1* with different CRP binding sequences were transferred to the chromosome of strain pop2094 such that the *lac* genes were controlled by *galP1* exactly as previously described (Gaston *et al.*, 1988).

Abbreviation used: CRP cAMP, cyclic AMP receptor protein cyclic AMP complex.

‡ To whom correspondence and reprint requests should be addressed.

Fig. 1. DNA used in this study

The Figure illustrates a derivative of pAA121 carrying the pBR322 origin of replication and β -lactamase genes (*ori* and *amp^R*) and the *galK* gene (hatched). The enlarged section shows the restriction enzyme cutting sites used in this work. A modified *galP1* promoter is carried between a *HaeIII*–*EcoRI*–*BamHI* linker upstream of -57 and a *HindIII* linker downstream of $+45$; distances are numbered with respect to the *galP1* transcription initiation site (marked by the wavy arrow). The *galP1* promoter used contains an *SphI* site introduced between -31 and -26 ; this allows the *gal* CRP–cAMP binding site to be replaced with synthetic *BamHI*–*SphI* fragments of any desired sequence.

Repression of this expression by glucose was also measured precisely as before (Fig. 6 of Gaston *et al.*, 1988).

RESULTS

CRP–cAMP binding to a consensus site

The CRP–cAMP complex binds to a 22 bp nucleotide sequence which, at the *galP1* promoter, is located from bp -52 to -31 . The starting point for this work was the plasmid illustrated in Fig. 1 which is a derivative of pAA121 into which a *BamHI*–*HindIII* fragment carrying the *galP1* promoter sequence was cloned. Gaston *et al.* (1988) described in detail how the internal *SphI* site could be exploited to facilitate the replacement of the wild-type *gal* CRP binding sequence with new sequences: essentially, synthetic oligonucleotides carrying *SphI* and *BamHI* cohesive ends are ligated between the *SphI* and *BamHI* sites of the plasmid shown in Fig. 1. In initial experiments we attempted to substitute the segment from -52 to -31 at the *galP1* promoter with a 22 bp sequence corresponding exactly to the consensus for CRP binding, 5' AAATGTGATCTAGATCACATT 3'. However our efforts to clone this sequence failed due

to the frequent appearance of spontaneous deletions (T. Palmer, unpublished work). In contrast, the sequence 5' AAATGTGATCTAGATCACACTG 3', which differs from the consensus only at two poorly conserved positions, could be maintained stably. Gaston *et al.* (1988) showed that, as expected, this site, which was labelled AA', bound CRP–cAMP exceptionally tightly. However, surprisingly, CRP–cAMP-dependent expression from *galP1* carrying this site was not repressed by the inclusion of glucose in the growth medium, in contrast to expression from the *gal* or *lac* promoters carrying naturally-occurring sites.

Fig. 2 shows an experiment to determine the cyclic AMP concentration at which CRP binding to the AA' site is triggered. Radioactively labelled DNA fragments carrying either the AA' site at *galP1* (lanes a–i) or the *lac* site (lanes j–r) were incubated with purified CRP in the presence of different concentrations of cyclic AMP ranging from 0 to $5.0 \mu\text{M}$. Free and bound fragments were then resolved by polyacrylamide-gel electrophoresis. The result shows that CRP binding to both the *galAA'* or *lac* sites absolutely requires cyclic AMP but that the concentration of cyclic AMP required for binding to the *galAA'* site ($0.1 \mu\text{M}$) is ten times less than the concentration required for binding to the *lac* site ($1 \mu\text{M}$).

The tight binding of the CRP–cAMP complex to the *galAA'* site is clearly shown in Fig. 3(a). Here increasing concentrations of CRP were incubated with a mixture of fragments carrying the *galAA'* site and the *lac* site in the presence of a fixed concentration of cyclic AMP (which was also included in the gel and running buffer). As we previously noted (Gaston *et al.*, 1988), CRP binding to the *galAA'* site is an order of magnitude tighter than to the *lac* site, which is the tightest-binding naturally occurring site that has been previously characterized. However, experiments such as that shown in Fig. 3(a) do not give a true measure of the tightness of binding, as the amount of CRP required to shift the *galAA'* fragment in the assay is merely the amount required to titrate the fragment. As it is impossible to work at fragment concentrations of less than 0.2 nM , we cannot discriminate between different dissociation constants that are appreciably less than 10^{-10} M . In order to estimate the real strength of the *galAA'* site we have therefore measured the dissociation rate for CRP bound at the sequence. The half-time for dissociation in a particular set of conditions gives an empirical measure of binding strength and allows us to order the affinity of very tight-binding sites. Fig. 3(b) shows an experiment where equimolar amounts of fragments carrying the *galAA'* site and the *lac* CRP binding site were incubated with CRP–cAMP and the resulting complexes were then challenged with an excess of unlabelled DNA. Samples from the mixture were loaded on a gel at different times to separate free and bound *galAA'* and *lac* fragments. In the conditions of this experiment the half-time for dissociation from the *lac* fragment is around 8 min whereas very little free *galAA'* fragment appears even after 128 min, showing that the complex is very stable. Fig. 3(c) shows a similar experiment where the competition was performed at higher salt concentrations (150 mM-KCl); in these conditions rapid dissociation from the *lac* fragment occurs with a half-time of less than 2 min whereas after 140 min only 50% of the *galAA'* fragment has dissociated from the complex with CRP–cAMP.

Fig. 2. Dependence of CRP binding on cyclic AMP concentration

CRP (10 nM) was preincubated with 0.5 nM labelled DNA fragment in the presence of different concentrations of cyclic AMP as shown for each lane. The Figure shows an autoradiogram of a gel on which free (f) and bound (b) fragments were resolved. In lanes a–i a 203 bp *Eco*RI fragment carrying the wild-type *lac* promoter sequence was used, whilst in lanes j–r the DNA was a *Hae*III–*Hind*III fragment isolated from pAA121 (see Fig. 1) carrying the AA' site.

Fig. 3. Comparison of the *gal*AA' and *lac* CRP–cAMP binding sites

(a) Equimolar quantities of labelled fragments (0.2–0.5 nM) carrying the *lac* or *gal*AA' sites were incubated for 20 min at 22 °C in binding buffer [40 mM-Tris/HCl (pH 7.9)/10 mM-MgCl₂/100 mM-KCl/100 μg of bovine serum albumin/ml] with the concentrations of CRP shown in the presence of 0.2 mM-cyclic AMP. Free and bound fragments were then separated by gel electrophoresis, visualized by autoradiography and quantified by densitometer scanning. (b) Dissociation of CRP–cAMP–DNA complexes (10 nM-CRP, 0.2 mM-cyclic AMP and 0.5 nM-labelled DNA fragments in binding buffer) was initiated by the addition of excess unlabelled plasmid DNA (130 μg/ml) carrying the AA' site. Aliquots of the mixture were loaded onto a running polyacrylamide gel at the time intervals shown in the Figure in order to resolve the free and bound DNA. The sample in lane b was taken immediately prior to the addition of competitor. In the control lanes (a and i) labelled DNA was added after a mixture of CRP, cyclic AMP and excess unlabelled DNA had been preincubated. This shows that the CRP is irreversibly trapped by the unlabelled DNA. (c) The same experiment as in (b) carried out at higher salt concentration (150 mM-KCl). Under the conditions used here the half-life of the CRP–cAMP–*lac* DNA complex was under 5 min, whereas the half-life of the CRP–cAMP–*gal* DNA complex is about 120 min.

Effects of point mutations in the consensus CRP binding sequence

To investigate the factors responsible for the great stability of complexes between CRP–cAMP and the *gal*AA' binding site we have made a number of mutations at positions throughout the AA' site. To do this synthetic oligonucleotides were cloned between the *Bam*HI and *Sph*I sites at *gal*P1 (see Fig. 1) such that the 22 bp CRP

binding sequence from –52 to –31 was altered. Because we were interested in changes that reduced the great stability of the complexes without destroying binding, the changes were made at positions where the bases are poorly conserved between different CRP sites.

Fig. 4 shows the 'top strand' of the 22 bp sequences of the *gal*AA' CRP binding site together with six derivatives, T4–T9, and, for reference, the consensus and the wild type *gal* and *lac* sequences. As the consensus sequence is

CRP consensus:		Left	1	11	Right	Half-time (min)	Expression during growth in glucose (% of maximum)
		a a a T G T G A t c t a g a T C A C A t t t	11	1			
AA':	A A A T G T G A T C T A G A T C A C A C T G					>120	85
T4:	T A A T G T G A T C T A G A T C A C A C T G					>120	90
T5:	A A A T G T G A G C T A G C T C A C A C T G					>120	N.D.
T6:	T A G T G T G A T C T A G A T C A C A C T A					>120	85
T7:	C C C T G T G A T C T A G A T C A C A G G G					2	21
T8:	G C A T G T G A T C T A G A T C A C A T G C					16	35
T9:	T A G T G T G A T C G C G A T C A C A C T A					16	35
gal site:	T A A T T T A T T C C A T G T C A C A C T T					1	17
lac site:	T A A T G T G A G T T A G C T C A C T C A T					8	N.D.

Fig. 4. Nucleotide sequence of CRP binding sites used in this study

Nucleotide sequences of the top strand of synthetic DNA carrying CRP-binding sites, AA' and T4–T9. Using the sites shown in Fig. 1 the nucleotide sequences were cloned from bp –52 to –31 at *galP1*. For reference the 22-base consensus sequence is shown on the top line, numbered leftward and rightward from the centre of symmetry. Nucleotide sequences of the CRP binding sites at the wild-type *gal* and *lac* promoters are shown on the lower two lines. The dissociation time for CRP–cAMP from each of the sites is shown next to each sequence: to make these measurements *Hae*III–*Hind*III fragments containing each sequence were complexed to CRP–cAMP then challenged with an excess of DNA precisely as in Fig. 3(b). The effect of glucose on expression *in vivo* from *galP1* carrying each of the sites is shown in the final column of the figure. Data are presented as the % of expression measured after growth in medium containing glucose (100 % is taken as expression in the absence of glucose).

symmetric we have numbered the bases divergently from the centre of symmetry using the convention of Lehming *et al.* (1988). DNA fragments carrying each of the derivatives were titrated with CRP as in Fig. 3(a) and the dissociation time was measured in a competition experiment as in Fig. 3(b). Each of the derivatives, T4–T9, bound CRP tightly giving results identical to that in Fig. 3(a) (not shown). However, the dissociation times measured in challenge experiments such as in Fig. 3(b) vary widely from one site to another (Fig. 4). The changes at positions 3, 9 and 11 in derivatives T4, T5 and T6 have little or no effect on the stability of the CRP–cAMP–DNA complexes. In contrast, the changes at position 1 and 10 in derivatives T9 and T8 respectively reduce the half-time of the complexes to around 16 min. In derivative T7 the cumulative changes at positions 9, 10 and 11 destabilize the complexes further and the dissociation time is reduced to 2 min.

To determine whether the mutations in the series T4–T9 affect the function of the CRP binding sites *in vivo*, DNA fragments carrying the *galP1* promoter with the different sites were transferred to the chromosome of the strain pop2094 such that *lacZ* expression was dependent on *galP1*. β -Galactosidase levels were then determined in cultures grown on fructose as a carbon source: in each case 1000 ± 100 units of β -galactosidase activity was found (full details of the transfers, growth conditions and assays are presented in Gaston *et al.*, 1988). The assays were repeated with 0.1 % glucose included in the growth media: in the final column of Fig. 4 the measured activity is expressed as a percentage of the maximal activity in the absence of glucose. The data show that glucose has little effect on expression from *galP1* carrying the AA' CRP site or the T4 and T6 derivatives, whereas significant reductions are found with the T7, T8 and T9 sites. Fig. 4 shows that there is an

inverse correlation between the repression of *galP1* activity by glucose *in vivo* and the stability of complexes between CRP–cAMP and the different sites *in vitro*: the most tightly binding sites are not susceptible to repression by glucose.

DISCUSSION

CRP is a dimeric protein which, when complexed to cyclic AMP, binds to target sites on DNA covering at least 22 bp. The consensus sequence for this target, 5' aaatGTGATctagaTCACAtt 3' consists of an inverted repeat of an 11 bp element. All the available evidence from genetic, biochemical and model-building studies suggest that the amino acid side chains of the protein make strong contacts with the bases at positions 4–8 (by convention we number the sequence leftwards and rightwards from the centre of symmetry: see Fig. 4); these are the positions that are strongly conserved between the 26 different sites that have been collated (Ebright *et al.*, 1984b; Weber & Steitz, 1984; de Crombrughe *et al.*, 1984; Gent *et al.*, 1987; Berg & Von Hippel, 1988; Jayaraman *et al.*, 1989). All the synthetic sites we have examined here contain the consensus sequence at these positions and all the fragments containing these sites bind tightly to CRP–cAMP. However, our results clearly show that the stability of these complexes is dependent on the precise nucleotide sequence at the poorly conserved positions (positions 1–3 and 9–11).

The affinity of the DNA fragments carrying the AA' site for CRP–cAMP is so high that it is difficult to quantify. In our conditions the dissociation time is over 2 h even at salt concentrations of 0.15 M. Interestingly, *in vivo* expression that depends on this site is not repressed by glucose (Fig. 4 and Gaston *et al.*, 1988). We found that the cyclic AMP level required to trigger CRP

binding to this site is $0.1 \mu\text{M}$, confirming the observation of Kolb *et al.* (1983a) that tighter binding sites require lower concentrations of cyclic AMP for binding. Since the intracellular concentration of cyclic AMP never falls below $1 \mu\text{M}$ (even with cells grown on glucose) (Joseph *et al.*, 1982; Ullmann & Danchin, 1983), this explains why transcription dependent on the AA' site cannot be turned off and why the consensus sequence is useless as a switch and is never found in a wild-type gene.

The AA' site differs from the consensus at positions 9R and 11R. The results with site T4, T5 and T6 show that other mutations at positions 9 and 11 changes at position 3 have no measurable effect on the stability of the complexes (Fig. 4). In contrast, changes at positions 1 and 10 (sites T7, T8 and T9) cause significant decreases in complex stability and, *in vivo*, allow glucose repression from dependent promoters. CRP-cAMP binding at a number of sites has been shown by several authors to be sensitive to changes in the flanking sequences (positions 9, 10 and 11) (Kolb *et al.*, 1983b; Yu & Reznikoff, 1984; Gartenberg & Crothers, 1988). However, to our knowledge, clear effects of changes at the centre of the sites (positions 1, 2 and 3) have not previously been demonstrated. The most likely explanation for the effects of mutations reported here is that they alter the bendability of the DNA. It is well known that, when CRP-cAMP binds to a site, the flanking sequences wrap round the CRP, stabilizing the binding and leading to an extended footprint and a bend in the DNA that may exceed 90° (Kolb *et al.*, 1983a; Wu & Crothers, 1984; Liu-Johnson *et al.*, 1986; Warwicker *et al.*, 1987; Thompson & Landy, 1988). Alterations in the flanking sequences such as in sites T7 and T8 can make bending more difficult and reduce the stability of the complexes (Gartenberg & Crothers, 1988). We propose that the changes at positions 1 in site T9 also somehow reduce the 'bendability' or 'twistability' of the CRP binding site, 'resisting' the CRP-induced conformational changes in the DNA and causing reduced stability in binding. Interestingly, similar changes at the centre of the 434 operator significantly affect the affinity of the operator for the 434 protein (Koudelka *et al.*, 1988).

We are grateful to Mrs. C. Webster for expert technical assistance. This work was funded by SERC with a grant no. GRD 53494 and a studentship to K.G.

REFERENCES

- Berg, O. & Von Hippel, P. (1988) *J. Mol. Biol.* **200**, 709–723
- Busby, S., Kotlarz, D. & Buc, H. (1983) *J. Mol. Biol.* **167**, 259–274
- de Crombrughe, B., Busby, S. & Buc, H. (1984) *Science* **224**, 831–838
- Ebright, R., Cossart, P., Gicquel-Sanzey, B. & Beckwith, J. (1984a) *Nature (London)* **311**, 232–235
- Ebright, R., Cossart, P., Gicquel-Sanzey, B. & Beckwith, J. (1984b) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 7274–7278
- Gartenberg, M. & Crothers, D. (1988) *Nature (London)* **333**, 824–829
- Gaston, K., Chan, B., Kolb, A., Fox, J. & Busby, S. (1988) *Biochem. J.* **253**, 809–818
- Gent, M., Gronenborn, A., Davies, R. & Clore, G. (1987) *Biochem. J.* **242**, 645–653
- Jayaraman, P.-S., Cole, J. & Busby, S. (1989) *Nucleic Acids. Res.* **17**, 135–146
- Joseph, E., Bernsley, C., Guiso, N. & Ullmann, A. (1982) *Mol. Gen. Genet.* **185**, 262–268
- Kelsall, A., Evans, C. & Busby, S. (1985) *FEBS Lett.* **180**, 155–159
- Kolb, A., Spassky, A., Chapon, C., Blazy, B. & Buc, H. (1983a) *Nucleic Acids Res.* **11**, 7833–7851
- Kolb, A., Busby, S., Herbert, M., Kotlarz, D. & Buc, H. (1983b) *EMBO J.* **2**, 217–222
- Koudelka, G., Harbury, P., Harrison, S. & Ptashne, M. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4633–4637
- Lehming, N., Sartorius, J., Oehler, S., von Wilcken-Bergmann, B. & Müller-Hill, B. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 7947–7951
- Liu-Johnson, H.-N., Gartenberg, M. & Crothers, D. (1986) *Cell* **47**, 995–1005
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Thompson, J. F. & Landy, A. (1988) *Nucleic Acids Res.* **16**, 9687–9705
- Ullmann, A. & Danchin, A. (1983) *Adv. Cyclic Nucleotide Res.* **15**, 1–46
- Warwicker, J., Engelman, B. & Steitz, T. (1987) *Proteins* **2**, 283–289
- Weber, I. & Steitz, T. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 3973–3977
- Wu, H.-M. & Crothers, D. (1984) *Nature (London)* **308**, 509–513
- Yu, X.-M. & Reznikoff, W. (1984) *Nucleic Acids. Res.* **12**, 5449–5464

Received 28 February 1989/12 April 1989; accepted 13 April 1989