Mode of Selectivity in Cyclic AMP Receptor Protein-Dependent Promoters in Escherichia coli †

ARTICLE *in* BIOCHEMISTRY · JANUARY 1996

Impact Factor: 3.02 · DOI: 10.1021/bi952187q · Source: PubMed

2 AUTHORS, INCLUDING:

James Ching Lee
University of Texas Medical Branch at Galves…

140 PUBLICATIONS **5,725** CITATIONS

SEE PROFILE

Mode of Selectivity in Cyclic AMP Receptor Protein-Dependent Promoters in *Escherichia coli*†

Erica A. Pyles and J. Ching Lee*

Department of Human Biological Chemistry and Genetics, The University of Texas Medical Branch, Galveston, Texas 77555-1055

Received September 13, 1995; Revised Manuscript Received November 20, 1995⊗

ABSTRACT: *Escherichia coli* cAMP receptor protein (CRP) controls more than 20 genes. There are significant differences in the promoter regions in these genes. Thus, an elucidation of the mechanism of CRP action requires knowledge about the mode of selectivity in these promoters. An earlier study [Heyduk, T., & Lee, J. C. (1990) *Proc. Natl. Acad. Sci. U.S.A. 81,* 1744−8] indicates that the CRP(cAMP)$_1$ conformer exhibits the highest affinity for the *lac* PI site in the *lac* operon. It is conceivable that the CRP conformer that binds with the highest affinity to these other sites may not be CRP(cAMP)$_1$. To investigate this possibility, the binding of CRP to nine CRP binding sites was studied as a function of cAMP concentration. The CRP binding sites employed in this investigation were chosen to represent the primary promoter sites from class I (*lac* site PI) and class II (sites PI of *gal* and *crp*) as well as secondary CRP binding sites (*crp* site PII and *cat* PII) to further understand the molecular mechanism of CRP in controlling the transcription of these bacterial genes. The affinity of CRP for three synthetic CRP binding sites was also examined to explore the contribution of the inverted repeat region and sequences surrounding the recognition motifs. The synthetic sequences are *gallac* which contains the *lac* recognition motifs in the background of *gal*, modified *cat* PII which contains an 8-base pair spacer between the recognition motifs rather than the 7-base pair sequence naturally found in *cat* PII, and a random sequence that has no known similarity to any CRP binding site found in nature. The apparent affinities of these sequences for CRP were quantitatively determined to be biphasic in their cAMP dependence. The CRP(cAMP)$_1$ conformer was found to have the highest affinity for all of the DNA sequences examined. No specific affinity was observed for these sequences with free CRP and CRP(cAMP)$_2$. The affinity of CRP for DNA was sequence-dependent and increased in the following order: random < *cat* site PII, modified *cat* site PII, *crp* sites PI and PII < *gal* site PI < *lac* site PI < *gallac*. These results indicate that the entire CRP binding site sequence and its natural variability provide information to CRP. These promoter sites which appear to have different mechanisms at the molecular level are transcriptionally controlled by the same CRP conformer, CRP(cAMP)$_1$. Thus, the regulation of transcription by CRP is more subtle than choosing different conformational forms of CRP. Using "physiological" concentrations of various components, a computer simulation study was conducted to illustrate the possible consequences of the thermodynamic parameters determined in this study. It is evident that the promoters of protein systems regulating the transport and metabolism of carbohydrates are responsive to low cAMP concentrations. However, the promoter for controlling the expression of CRP is highly regulated by the fluctuation of cAMP concentration.

The mechanisms of how the ubiquitous cAMP receptor protein, CRP,[1] controls transcription in *Escherichia coli* have received a great deal of attention from a variety of approaches such as structural, genetic, and solution studies [for reviews, see de Crombrugghe *et al.* (1984), Crothers and Steitz (1992), and Kolb *et al.* (1993)]. However, the molecular mechanisms by which CRP regulates transcription and selects the specific promoter from among the more than 20 CRP-dependent promoters are still unclear. Solution studies of CRP and DNA have indicated that the properties of both molecules contribute to the complexity of the molecular mechanisms of CRP action. CRP is allosterically controlled by cAMP, and the binding of cAMP to CRP is negatively cooperative in nature. The three solution conformers of CRP that have been detected *in vitro* are free CRP, CRP(cAMP)$_1$, and CRP(cAMP)$_2$, and they have been shown to exhibit different solution properties *in vitro* (Heyduk & Lee, 1989). In addition, three classes of CRP binding sites on DNA whose overall mechanisms may differ from one another at the molecular level have been identified (Ebright, 1993; Zhou *et al.*, 1994). The common feature shared by all three classes is that they require CRP to affect gene expression. The location of the CRP and RNA polymerase binding sites differentiates class I CRP binding sites from those in class II. Class I CRP binding sites are located upstream from the transcriptional start site at discrete distances and on the same face of the DNA helix as the RNA polymerase binding site, whereas class II CRP binding sites overlap the RNA

† This work was supported by the Robert A. Welch Foundation Grants H-0013 and H-1238 and NIH Grant GM-45579.

* Author to whom correspondence should be addressed.

⊗ Abstract published in *Advance ACS Abstracts,* January 15, 1996.

[1] Abbreviations: CRP, cAMP receptor protein; CPM, 7-(diethylamino)-3-(4′-maleimidylphenyl)-4-methylcoumarin; TEDKG buffer, 50 mM Tris, 1 mM ethylenediaminetetraacetic acid (EDTA), 1 mM dithiothreitol (DTT), 100 mM KCl, 10% (w/v) glycerol, and pH 7.80 at 25 °C; TE buffer, 50 mM Tris, 1 mM EDTA, and pH 7.80 at 25 °C; TEK(100) buffer, 50 mM Tris, 1 mM K_2EDTA, 100 mM KCl, and pH 7.80 at 25 °C.

0006-2960/96/0435-1162$12.00/0

polymerase binding site. Thus, activation and repression of gene expression is possible and is due to the interaction of CRP with DNA. Class II binding sites can be either activated or repressed by CRP−DNA complex formation, while interaction of CRP with class I binding sites can only induce transcriptional activation. The third class of CRP binding sites requires another regulatory protein, and the CRP binding sites are located at various distances from the transcriptional start site. Therefore, the mechanisms of transcriptional control by CRP for class III sites are more complex than those for sites in classes I and II.

Furthermore, it is common to find two CRP binding sites for promoters in class I and II. These sites are designated as CRP binding sites PI and PII. Binding of CRP to a secondary site usually results in transcription of a divergent promoter. However, there are cases where the opposite is true, such as in the *crp* gene, in which CRP represses the transcription of its own gene (Okamoto & Freundlich, 1986; Aiba, 1983). The sequences of CRP binding site PI usually contain the conserved TGTGA motif which is separated from a variable inverted repeat sequence by a spacer. The spacer sequence and the sequences outside the conserved and inverted repeat motifs vary considerably from one promoter site to another. In addition, the length of the spacer is also variable. The length most commonly found is 6 base pairs (bp); however, there are some that have 7 or 8 and possibly 9 base pairs between the recognition motifs (Barber & Zhurkin, 1990; Barber *et al.*, 1993). The sequences of CRP binding site PII frequently have more sequence variability than site PI sequences, and often the characteristic TGTGA motif is not as well-conserved.

These natural variations observed in the promoter sites may affect the interaction of CRP with DNA. It has been shown that nucleotide substitutions in the recognition motifs will affect the association and dissociation constants of CRP−DNA complexes. Changing the inverted repeat in the *lac* sequence from TCACT to TCACA increased the apparent binding affinity of CRP by 17-fold (Gartenberg & Crothers, 1988). It has been shown from filter binding assays that the Berg and von Hippel consensus binding site (ICAP) binds CRP ∼450 times greater than the natural *lac* promoter site binding to the protein (Ebright *et al.*, 1989). Furthermore, gel mobility shift assays have determined that the apparent affinity of CRP for the *gal* sequence is ∼5-fold lower than that of *lac* for CRP (Kolb *et al.*, 1983). Dissociation of CRP complexed to *lac*, *gal*, and ICAP were measured by addition of nonspecific DNA. It was found that the lifetimes for the *lac*−CRP, *gal*−CRP, and ICAP−CRP complexes were 8 min, 1 min, and >2 h, respectively (Gaston *et al.*, 1988). In an extensive study, Dalma-Weiszhausz *et al.* (1991) showed that the binding affinity of CRP to DNA depends on the sequence of the distal binding domain. There is a correlation between binding affinity and extent of bending. This evidence shows that the sequence variability is a contributing factor. However, it is not possible to directly deduce from these studies which CRP conformer(s) binds to these sequences.

Moreover, the mechanism of how CRP is capable of repressing its own gene by binding to its primary binding site, which results in the expression of a divergent promoter (Aiba, 1983; Okamoto & Freundlich, 1986; Hanamura & Aiba, 1992), is also not clearly understood. There is recent evidence to support the idea that the *crp* site PII also binds to CRP in the presence of cAMP and that CRP binding to site PII results in transcription of *crp* mRNA with concomitant reduction in the synthesis of the divergent promoter (Hanamura & Aiba, 1992). Synthesis of *crp* mRNA and the divergent mRNA were shown to differ in their cAMP and CRP requirements both *in vitro* and *in vivo*. These results imply that a different mechanism may be responsible for the activation and repression of the *crp* gene.

These diverse observations that are detected in different CRP-dependent promoters lead to an obvious question on promoter selectivity by CRP. How does CRP select the correct promoter? There are at least three potential modes that CRP can adopt in exercising promoter selection: binding affinity of CRP to promoter is DNA sequence-dependent, the number of cAMP molecules bound to the homodimeric CRP may dictate the specific promoter to which the CRP binds, and the induced DNA bending differs among these promoters. Previous studies have shown that CRP$(cAMP)_1$ has the highest affinity for the *lac*−CRP binding site while CRP$(cAMP)_2$ and free CRP exhibit very weak or no affinity for sequence specific DNA, respectively (Heyduk & Lee, 1990). It is conceivable that different classes of promoter interact with different conformers of CRP.

The mode of promoter selectivity is a central issue in CRP function, and current knowledge does not provide adequate information to address this issue. Therefore, the cAMP dependence of CRP binding to *lac* site PI, *gal* site PI, *crp* sites PI and PII, and *cat* PII (catabolite aminotransferase promoter site; Le Grice *et al.*, 1982) was quantitatively determined (see Figure 1). In addition to determining the cAMP dependence of CRP binding to the binding sites of *lac*, *gal*, and *crp*, the binding of CRP to three synthetic sites was also investigated to further understand the effect of the DNA sequence on CRP−DNA interactions. These other CRP binding sites were chosen to quantitatively investigate the effect of the recognition motifs, spacer length, and sequences outside the recognition motifs on the affinity of CRP for DNA. In addition, a sequence generated by random numbers was employed to measure the affinity of CRP for a totally random, nonspecific sequence.

MATERIALS AND METHODS

Materials. Deprotected and desalted oligodeoxyribonucleotides were purchased from Genosys, or they were synthesized using β-cyanoethyl phosphoramidite chemistry on a Beckman Oligo1000. The fluorescent probe, 7-(diethylamino)-3-(4′-maleimidylphenyl)-4-methylcoumarin (CPM) was purchased from Molecular Probes. Other reagents and buffer materials were purchased from Beckman, Sigma, or Boehringer Mannheim.

Methods. Concentrations of solutions employed in this study were determined from the absorption spectrum using the following extinction coefficients: CRP monomer at 278 nm (20 400 M^{-1} cm^{-1}; Takahashi *et al.*, 1980) and cAMP at 259 nm (1.465×10^4 M^{-1} cm^{-1}; Windholz, 1976) in aqueous solution and CPM at 260 and 385 nm (2.3×10^4 and 3.3×10^4 M^{-1} cm^{-1}, respectively; Haughland, 1992) in methanol containing <1% w/v β-mercaptoethanol and <1% w/v sodium acetate. Absorption spectra were measured using a Hitachi U-2000 spectrophotometer.

CRP Preparation. CRP was isolated and purified from an overproducing strain (pPLcCRP1) of *E. coli* by the

CRP BINDING SITE		SEQUENCE		
LAC SITE PI	→	-80 -70 CAACGCAATTAA	-60 **TGTGA**GTTAGC**TCACT**	-50 CATTAGGCACCC
GAL SITE PI	←	-30 AGATGCGAAAAG	-40 **TGTGA**CATGGA**ATAAA**	-50 -60 TTAGTGGAATCG
GALLAC		AGATGCGAAAAG	**TGTGA**CATGGA**TCACT**	TTAGTGGAATCG
CRP SITE PI	←	+60 ACTGCACGGTAA	+50 +40 **TGTGA**CGTCC**TTGCA**	+30 TACATGCAGTAC
CRP SITE PII	→	-70 CGCGCAACGGAA	-60 GGCGACCTGGGTCATG	-50 CTGAAGCGAGAC
CAT PII		-140 CGAATAAATACC	-130 **TGTGA**CGGAAGA**TCACT**	-120 TCGCAGAATAAA
MODIFIED CAT PII		CGAATAAATACC	**TGTGA**CGGAAGAA**TCACT**	TCGCAGAATAAA
RANDOM		TTACGAACTCAG	TTCTGATACCAAGCAG	CCCAGTTGTGGG

FIGURE 1: Sequences of oligodeoxyribonucleotides employed in this study. The *gallac* sequence is a derivative of the primary *gal*−CRP binding site in which the inverted repeat sequence is replaced with the inverted repeat sequence of the *lac*−CRP binding site. The *cat* PII and modified *cat* PII differ from the other sequences in their spacer length between the two half-sites. *cat* PII and modified *cat* PII have spacer lengths of 7 and 8 nucleotides, respectively, rather than the more frequently observed 6-base pair spacer length. The numbers above the naturally occurring CRP binding sites indicate the distance in base pairs relative to the transcriptional start site. Arrows represent the direction of transcription relative to the CRP binding site sequences. The breaks in the sequences serve as guides so that the "essential" core of the CRP binding site can be more easily observed to show the recognition motifs (when present) and the spacer between the recognition motifs. The sequences are aligned by the highly conserved TGTGA motif (bold). The variable inverted repeat sequences are also in bold. All sequences are listed from the 5′-end to the 3′-end (left to right).

method of Heyduk and Lee (1989). CRP was stored in TEDKG buffer at −20 °C. Purified CRP was observed to migrate as one band on a sodium dodecyl sulfate−polyacrylamide gel. The integrity of the protein was judged by measuring the stoichiometry and affinity of the protein for the *lac*−CRP binding site in the presence of 200 μM cAMP, and similar results were obtained as previously described using the same solution conditions (Heyduk & Lee, 1990). CRP was routinely dialyzed against the desired buffer and carefully filtered (0.2 μm) before use.

DNA Preparation. The DNA sequences containing the CRP binding site were synthesized with the recognition sites centrally located in the strand, and each of the sequences flanking the binding site had the same number of nucleotides (see Figure 1). The modified *gal* sequence, *gallac*, examined is essentially the *gal*−CRP primary binding site with the inverted repeat ATAAA replaced with the inverted repeat sequence of the *lac*, TCACT. The modified *cat* PII sequence contains 8 base pairs in between the recognition motifs rather than 7 base pairs naturally found in the *cat* PII sequence. The other synthetic sequence was generated from random numbers with two restrictions. The generated sequence was 40 base pairs in length, and the number of A, C, T, and G in the sequence was equivalent. Approximately 20 oligomers were generated and were carefully searched for sequences that were similar to the known DNA recognition sites for CRP, both the highly conserved motif (TGTGA) and the more variable inverted repeats. Random sequences containing the conserved motif, inverted repeats, or partial recognition sequences of CRP binding sites were discarded. The one chosen for this study is shown in Figure 1. The rationale for employing such a sequence was to reduce the possibility of using a nonspecific sequence that may have the tendency to adopt various conformations favorable for the interaction of CRP with DNA.

The single-stranded oligodeoxyribonucleotides were routinely purified before synthetic modification and biophysical measurements were performed. The DNA was purified by denaturing polyacrylamide gel electrophoresis (12−15% acrylamide, 7.5 M urea). Solutions were desalted using a Waters Sep-Pak C_{18} column as previously described (Heyduk & Lee, 1990). Purified oligodeoxyribonucleotides were stored in water at −20 °C. The integrity of the purified oligomers was monitored by polyacrylamide gel electrophoresis, and each oligomer migrated as one band.

The fluorescent probe (CPM) was covalently attached to the 5′-end of DNA sequences and was located upstream or downstream relative to the transcriptional initiation start sites (Heyduk & Lee, 1990, 1992). The fluorescently labeled oligodeoxyribonucleotides were annealed with the appropriate complementary strand phosphorylated at the 5′-end. The integrity of the sample was monitored by nondenaturing polyacrylamide gel electrophoresis (12−15%). Typically, very little (≤5%) or no ss-DNA (single-stranded) was observed. Only nonfluorescently labeled ss-DNA was observed in samples containing ds-DNA (double-stranded). The ratio of probe to DNA was calculated from the absorption spectrum of the double-stranded, fluorescently labeled DNA, and it ranged from 0.4 to 1.0 mol of probe per mole of DNA. The samples were dialyzed against TE buffer and stored at −20 °C.

Fluorescence Measurements. The anisotropy of fluorescently labeled oligodeoxyribonucoleotides (see Figure 1 for sequences) binding to CRP as a function of cAMP concentration was measured using a SLM 8000C spectrofluorometer operated in the T-format with analog detection. The temperature of the solution in the cell was measured using a thermocouple and was constant at 25.0 ± 0.1 °C (Physitemp, Model BAT-12). The excitation wavelength was 390 nm, and an Oriel band-pass filter, 59816, was placed in the

excitation path to reduce second order scatter. Oriel bandpass filters 59850 (>25% transmittance between 470 and 570 nm) were used in the emission paths. The dimensions of the fluorescence cells used were 10 × 10 mm. TEK-(100) buffer (pH 7.80) at 25.0 °C was used.

Binding Isotherms. The oligodeoxyribonucleotides (12 nM) and cAMP (from 0 to 29.4 mM) were titrated with CRP. The following reaction best describes the global binding of CRP to DNA:

$$PA_n + D \overset{K_{app}}{\rightleftharpoons} PA_nD \quad (1)$$

$$K_{app} = \frac{[PA_nD]}{[PA_n][D]} \quad (2)$$

where D, PA_n ($n = 0$, 1, or 2), and PA_nD are DNA and CRP−cAMP, and CRP−cAMP−DNA complexes, respectively. $P_T \equiv PA_n + PA_nD$. K_{app} is the apparent equilibrium constant. The affinity of CRP for DNA in the absence of cAMP is very weak. Thus, the contribution of CRP−DNA complex to the total CRP concentration was neglected.

The following equation results after substituting the total protein concentration, P_T, into eq 2 and accounting for all species with anisotropic motion:

$$r_{obs} = r_o + \Delta r[[K_{app}D_T + K_{app}P_T + 1 - [(K_{app}D_T + K_{app}P_T + 1)^2 - 4K_{app}{}^2D_TP_T]^{1/2}]/2K_{app}D_T] \quad (3)$$

r_{obs}, r_o, Δr, P_T, and D_T are the observed anisotropy, anisotropy of free DNA, total change in anisotropy, and total protein and DNA concentrations, respectively. The data were fit to eq 3 using the program SCIENTIST (MicroMath Scientific Software, Salt Lake City, UT). Dilution of the solutions during the titration affected the fitted parameters when the change of volume was >6%. Therefore, the effect of dilution was taken into consideration during data analysis when the volume change was >6%. The fitted parameters were obtained by nonlinear curve fitting. Errors are expressed in terms of 75% confidence intervals (support plane method).

Competition Experiments. CPM-labeled DNA for 26-base pair sequences of *lac*, *gal*, and *gallac* (labeling efficiency > 70%) and the respective unlabeled ds-DNA were titrated with CRP in the presence of 200 μM cAMP. The final mole ratio of CPM-labeled DNA to total DNA in the cell was 0.48. The data were fit to eq 3 as described above.

Stoichiometric Titrations. Stoichiometric binding of CRP to DNA was monitored by the change in anisotropy of the covalently attached probe on DNA as a function of CRP concentration. The cAMP concentration was 200 μM. The length of the DNA was 40 base pairs and the concentration of DNA used in the titrations was as follows: *gallac* (100 nM) and *gal* and *lac* (120 nM). Stoichiometric measurement of CRP complexed with *cat* PII, modified *cat* PII, and *crp* sites PI and PII was precluded due to excessive scattering of light caused by aggregated CRP−DNA complexes.

RESULTS

It has been previously demonstrated that covalent modification of DNA with a fluorescent probe is a useful and convenient tool for quantitative investigation of DNA−protein interactions (Heyduk & Lee, 1990). The stoichiometry and equilibrium binding of a protein to DNA can be measured under a variety of solution conditions by monitoring either the energy transfer from the protein to the fluorescent probe or the anisotropy of the probe. The latter measurement was shown to be more reproducible. Therefore, the change in the anisotropic motion of the fluorescent probe (CPM) covalently linked to several oligodeoxyribonucleotides was used as a signal to measure the binding of the protein to several CRP binding sites.

It is important to evaluate if covalent attachment of the probe to either end of the DNA sequences affects the stoichiometry and affinity of CRP for DNA. Results from control experiments show that the probe most likely does not interfere with CRP−DNA interactions since addition of unlabeled DNA to the respective CPM-labeled DNA sequences did not alter the apparent affinity of the DNA for the protein (data not shown). These results are consistent with previous competition studies that used several lengths of unlabeled and CPM-labeled *lac* sequences with CRP at various salt concentrations (Heyduk & Lee, 1990, 1992).

Effect of the DNA Sequence on the Spectroscopic Properties of the Fluorescent Probe

The anisotropies of free CPM-labeled DNA and the CPM-labeled DNA−CRP complex provide information about the local environment of the probe. A less mobile or immobile environment for the fluorophore results in higher values of anisotropy, while lower values are observed when the fluorophore has greater mobility.

The anisotropy values of free 26-base pair *lac*, *gallac*, and *crp* site PI were 0.1872 ± 0.0068 and 0.1865 ± 0.0048 when the label was attached close to TGTGA motifs and to inverted repeat motifs, respectively. The identity of these values indicates that the local environments of the probe within these DNA sequences are similar. However, the results on the *gal* sequence do not follow the same trend. The anisotropy of CPM attached downstream from the transcriptional start site (close to TGTGA motif) was 0.2005 ± 0.0074, while that of the upstream CPM labeled *gal* sequence was 0.186. In addition, the anisotropy for CPM-labeled sequences that were 40 base pairs long was 0.195 ± 0.002. These data imply that there may be some interaction between the probe and the downstream portion of the 26-base pair *gal* sequence. Nevertheless, if there is an interaction between the probe and the downstream end of the *gal* sequence, it is probably small because no difference in the apparent affinity of CRP for *gal* was observed in the competition of CPM−*gal* and unlabeled *gal* for CRP relative to the apparent affinities determined without additional unlabeled *gal*.

There is further indication that variation in the sequences can affect the local environment of the probe. The total change in anisotropy for CRP−DNA complex formation with a DNA length of 40 base pairs appeared to be sequence-independent. The total change in anisotropy for PI sites of *lac*, *gal*, and *crp* saturated with protein was 0.023 ± 0.002. In addition, the total change in anisotropy for the synthetic *gallac* saturated with protein was 0.027 ± 0.002. In contrast, the total change in anisotropy observed for the DNA whose length was 26 base pairs in the presence and absence of CRP was found to be DNA sequence-dependent (sites PI of *lac*,

gal, *crp*, and *gallac*). It was found that the *lac* site PI and *crp* site PI binding to CRP resulted in a 0.032 ± 0.002 total change in anisotropy. In contrast, the total change in anisotropy for *gallac* and *gal* site PI binding to CRP was 0.049 ± 0.004. This indicates that the local environment of the probe when CRP is complexed with CPM-labeled DNA is different for *gallac* and *gal* site PI relative to that of *lac* site PI and *crp* site PI. Another possible cause for the difference in the total change in anisotropy is aggregation of the DNA−CRP complex (Heyduk *et al.*, 1992). However, this difference in anisotropy cannot be explained by concomitant aggregation with specific binding since an increase in salt concentration (up to 150 mM) did not change the anisotropy values when the binding sites were saturated by CRP. Higher concentrations of salt only decreased the affinities of the DNA for CRP (data not shown). Thus, the greater change in anisotropy values observed for CRP bound to the *gal* and *gallac* sequences most likely reflects a particular property of the DNA sequence and is not artificially induced. This interpretation is further supported by the fact that the same apparent affinity of CRP for *gal* was obtained with the fluorescent probe attached to either end of the sequence. It is important to note that no useful information can be obtained from the total change in anisotropy found for CRP bound to *crp* site PII, *cat* PII, modified *cat* PII, and the random sequence since the contribution from CRP−DNA aggregation was significant to the anisotropy signal observed. It appears that specific binding of CRP to DNA is linked to DNA−CRP aggregation; however, this issue is beyond the scope of this study and will be addressed elsewhere.

Stoichiometry of CRP−DNA Complexes

The stoichiometries of the DNA−CRP complexes were measured for the CRP binding sites that were 40 base pairs in length, in the presence of 200 μM cAMP in TEK(100) buffer (pH 7.80) at 25.0 °C. The curves are shown in Figure 2. The stoichiometries of the *lac*−CRP, *gal*−CRP, and *gallac*−CRP complexes were 0.90 (0.70, 1.2), 0.97 (0.78, 1.2), and 0.90 (0.79, 1.0), respectively. The numbers in parentheses indicate a confidence interval of 75%. The stoichiometry of the *lac*−CRP complex determined in this study is consistent with previous results that were measured using a 32-base pair *lac* binding site and a lower salt concentration of 25 mM (Heyduk & Lee, 1990). The total change in anisotropy observed for the *gallac*−CRP complex is slightly greater in magnitude than that observed for the stoichiometric complexes of CRP−*lac* and CRP−*gal*. This difference may possibly reflect a larger contribution of scatter to the observed signal or a slightly different local environment for the probe located on the *gallac* sequence relative to those of *gal* and *lac*. The binding of 1 mol of CRP to 1 mol of DNA was also observed using the shorter 26-base pair DNA sequences (data not shown). The stoichiometries of CRP bound to random, *cat* PII, modified *cat* PII, and *crp* sites PI and PII could not be accurately measured due to DNA−protein aggregation.

Effect of the DNA Sequence on the CRP−DNA Equilibrium

The binding of CRP to these DNA sequences was monitored by the change in anisotropy as a function of CRP concentration at 25.0 °C. The measurements were performed in TEK(100) buffer (pH 7.80) in the presence of 200 μM cAMP.

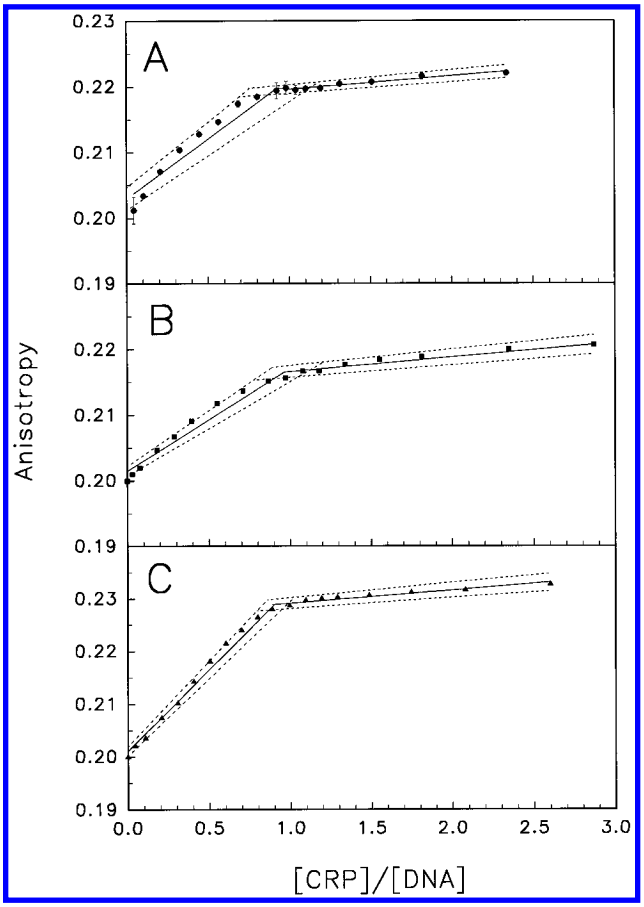

FIGURE 2: Stoichiometric titration curves of CRP binding to DNA sequences in the presence of 200 μM cAMP in TEK(100) buffer (pH 7.80) at 25.0 °C. The symbols correspond to the following DNA sequences: (A) *lac* site PI (●), (B) *gal* site PI (■), and (C) *gallac* (▲). The solid lines represent the best fits, and the dashed lines represent the 75% confidence limits.

A. Twenty-Six Base Pairs. Typical binding isotherms of CRP binding to *lac*, *gallac* and *gal*, *crp* sites PI and PII, and the random sequence are shown in Figure 3. The anisotropy values of free CPM-labeled DNA were normalized to 0.1800 so that the curves could be more easily compared. Normalizing the anisotropy affects neither the overall shape of the curves nor the affinities of CRP for the DNA sequences. The data were fit to eq 3, and the solid curves represent the best fit. The goodness of fit can be observed from the random distribution of the residuals about 0. The apparent equilibrium constants for CRP binding to these sequences are listed in Table 1. The isotherms of CRP binding to *cat* PII and modified *cat* PII are similar to the *crp* site PII−CRP binding isotherm and are therefore not shown in Figure 3 to simplify the graphs. The apparent equilibrium constant of *lac* binding to CRP is similar to the value previously determined under the same experimental conditions (Heyduk & Lee, 1990). The apparent affinity of CRP for *lac* was found to be 4-fold less than that of *gallac*, indicating that the affinity of CRP for its binding sites on DNA is not solely determined by the recognition motifs (conserved and inverted repeat motifs; see Figure 3A,B). The difference in the apparent affinities of CRP for *gal* and *lac* was only 4-fold, with CRP having a greater affinity for the *lac* binding site than the *gal* binding site. This is in good

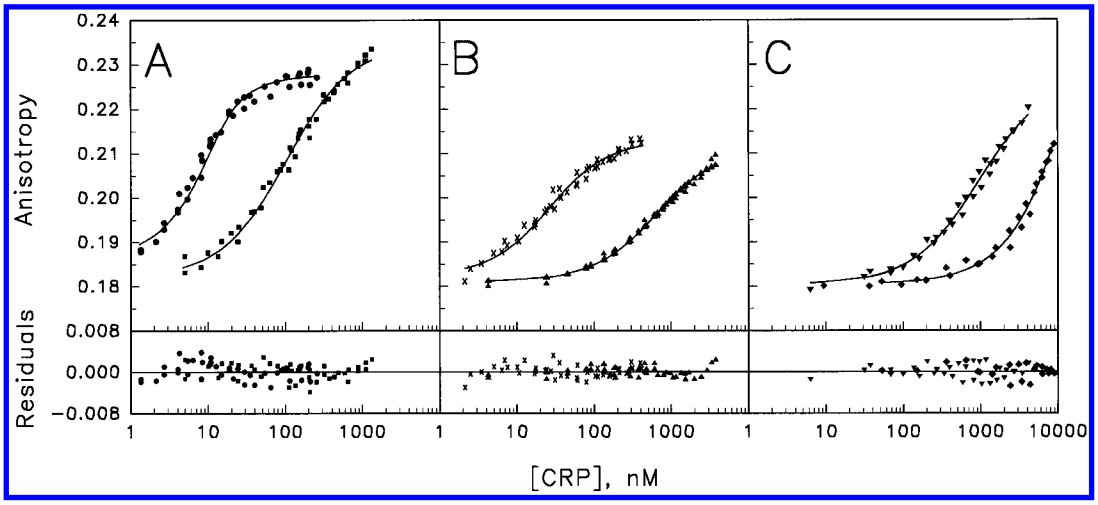

FIGURE 3: Typical binding isotherms of CRP binding to DNA sequences in the presence of 200 μM cAMP in TEK(100) buffer (pH 7.80) at 25.0 °C. The anisotropy of CPM-labeled DNA free in solution was normalized to 0.1800. The symbols correspond to the following DNA sequences: (A) *gallac* (●) and *gal* site PI (■), (B) *lac* site PI (×) and *crp* site PI (▲), and (C) *crp* site PII (▼) and random sequence (◆). The length of the CRP binding sites was 26 base pairs. Data were fit to eq 3 (in Materials and Methods), and at least two binding curves were fit simultaneously for each DNA sequence. The goodness of fit is shown by the distribution of the residuals about 0.

Table 1: Microscopic Equilibrium Association Constants for CRP−cAMP and CRP−cAMP−DNA Complexes[a]

DNA	$K_{app} \times 10^{-6}$, M^{-1}[b,c]	$K_{app} \times 10^{-6}$, M^{-1}[b,d]	$K_1 \times 10^{-4}$, M^{-1}[d,e]	$K_2 \times 10^{-3}$, M^{-1}[d,e]	$K_3 \times 10^{-6}$, M^{-1}[d,e]
lac site PI	200 (160, 250)[f,g]	41 (35, 49)[g]	5.5 (−1.7, 11)	1.1 (−.037, 2.7)	48 (23,72)
gal site PI	56 (35, 88)	9.5 (8.5, 11)	4.5 (−1.4, 10)	0.85 (−0.15, 1.9)	13 (7.3, 19)
gallac	330 (230, 470)	160 (130, 200)	4.7 (−2.0, 11)	0.89 (−0.28, 2.0)	250 (130, 350)
crp site PI	22 (15, 28)	1.7 (1.4, 2.1)	4.6 (2.7, 8.9)	1.0 (0.05, 2.0)	2.4 (1.6, 3.2)
crp site PII[h]	2.7 (2.0, 3.4)	1.0 (0.84, 1.2)	*i*	*i*	*i*
cat PII[h]	3.8 (2.4, 5.1)	1.1 (0.95, 1.3)	*i*	*i*	*i*
modified cat PII[h]	4.6 (3.5, 5.7)	2.5 (1.8, 3.4)	*i*	*i*	*i*
random[h,j]	0.48 (0.42, 0.54)	0.045 (0.017, 0.073)	*i*	*i*	*i*

[a] Measured in TEK(100) buffer (pH 7.80) at 25.0 °C. [b] Apparent equilibrium constants were calculated from eq 3 in Materials and Methods. The concentration of cAMP was 200 μM. [c] The length of DNA was 40 base pairs with the exception of *cat* PII and modified cat PII. In these cases, the length was 41 and 42 base pairs, respectively, to accommodate the greater spacer between the recognition sites. [d] The length of DNA was 26 base pairs with the exception of *cat* PII and modified *cat* PII. In these cases the length was 27 and 28 base pairs, respectively, to accommodate the greater spacer between the recognition sites. [e] The parameters K_1, K_2, and K_3 were obtained from fitting the apparent equilibrium constants as a function of cAMP concentration to eq 5 in Results with $K_4 = 0$. [f] Errors in parentheses are expressed in terms of 75% confidence intervals. [g] The apparent equilibrium constant is in good agreement with previously published results (Heyduk & Lee, 1990, 1992). [h] The apparent affinity of this binding site for the protein was found to display a similar biphasic cAMP dependence relative to that observed for the other sequences examined (see Results). [i] Not determined. [j] The random sequence was treated as a single binding site and not as multiple phosphate sites (52 and 80 phosphate sites for the 26 bp and 40 bp sequences, respectively) in the calculation of the apparent binding affinities.

agreement with apparent affinity values previously determined by the gel shift method (Gaston *et al.*, 1988). The apparent affinity of the *lac*−CRP complex was ∼25−40-fold greater than those determined for CRP binding to *crp* sites PI and PII (See Figure 3B,C). CRP has the same apparent affinity for *crp* sites PI and PII in the presence of 200 μM cAMP even though the latter binding site does not contain the recognition half site sequence frequently found in many CRP binding sites. However, there is a slight difference in the shape of the two binding isotherms (see Figure 3B,C). The binding of CRP to the *crp* site PII appears to have a higher degree of nonspecific character relative to the isotherm of CRP−*crp* site PI. The shape of the curve resembles the binding isotherm of CRP binding to the random sequence, namely, the lack of a plateau value at high CRP concentrations. Yet, the binding isotherm is shifted to the left relative to the binding isotherm for the random sequence, implying that some specific interaction exists between *crp* site PII and CRP.

It was also found that the affinity of CRP for the two sequences that contain a larger spacer between the recognition half-sites of TGTGA and TCACT, *cat* PII and modified *cat* PII, are very weak. The apparent equilibrium constants for CRP binding to *cat* site PII and to modified *cat* site PII are about 40- and 15-fold weaker than that for *lac* site PI binding to CRP; although, the recognition half-sites are the same as those of the *lac* (see Table 1 and Figure 1). This implies that the alignment of the major grooves of DNA relative to the DNA binding domain of the protein is an important factor in determining the degree of specificity. The results found for CRP binding to the modified *cat* PII and *cat* PII sequences are consistent with previously determined apparent affinities for synthetic sequences containing 8-base pair spacers under different solution conditions using the gel shift assay (Barber *et al.*, 1993). There is no specific interaction between the random, CPM-labeled sequence for CRP as indicated by the lack of saturation in the binding isotherm. The affinity was ∼20-fold lower than the weakest CRP−DNA interactions examined in this study.

B. Forty Base Pairs. Previous results have shown that the 26-base pair DNA is just short of the thermodynamically defined binding site (Heyduk & Lee, 1992). Thus, the interaction of the analogous 40-base pair DNA sequence with CRP was examined. It should be noted that the affinities of CRP for the analogous DNA sequences that are 40 base pairs in length were determined to be greater than the affinity of

CRP for the respective 26-base pair sequences under the same solution conditions (data not shown). The apparent equilibrium constants for CRP−DNA complexes with a DNA length of 40 base pairs in the presence of 200 μM cAMP are also listed in Table 1 for comparative purposes. The dependence of the length on the affinity of CRP for the *lac* binding site is consistent with previous binding measurements (Heyduk & Lee, 1990). The apparent equilibrium values determined for *lac* and *gallac* binding to CRP are relevant since a 2-fold increase in the concentration of DNA employed in this study relative to the apparent dissociation constants for the complexes is not sufficient to compromise the accuracy of measuring equilibrium constants (Weber, 1992). The relative strength of CRP binding to the 40-base pair sites with respect to the analogous 26-base pairs sites binding to CRP was found to differ depending on the DNA sequence. This demonstrates the contribution of the flanking ends of the CRP binding site to the overall binding energy of the DNA−CRP complex and that the extent of the contribution is sequence-dependent (Dalma-Weiszhausz *et al.*, 1991).

Dependence of CRP−DNA Interactions on cAMP Concentration

The apparent equilibria of CRP−DNA complexes in the presence of 200 μM cAMP indicate that CRP−DNA interactions are sequence dependent. The apparent affinity of CRP for DNA decreases in the following order: *gallac* > *lac* > *gal* > *crp* site PI ~ *crp* site PII ~ *cat* PII ~ modified *cat* II > random sequence. One cannot deduce from these data if these sequences preferentially bind to a specific CRP conformer. The apparent equilibrium constant is a composite of several equilibria and reflects the affinities of various CRP conformers for the DNA sequences. A more detailed description of DNA binding to CRP (compared to eq 1 in Materials and Methods) includes CRP, both cAMP-liganded CRP forms, $CRP(cAMP)_1$, and $CRP(cAMP)_2$, and the possible interactions of these conformers with DNA. The complete reaction scheme is shown below. P, A, PA, PA_2,

$$\begin{array}{ccccccccc} & & & \overset{K_1}{\rightleftharpoons} & & & & \overset{K_2}{\rightleftharpoons} & \\ P & + & A & & PA & + & A & & PA_2 \\ + & & & & + & & & & + \\ D & & & & D & & & & D \\ \Updownarrow K_6 & & & & \Updownarrow K_3 & & & & \Updownarrow K_4 \\ & & & \underset{K_7}{\rightleftharpoons} & & & & \underset{K_5}{\rightleftharpoons} & \\ PD & + & A & & PAD & + & A & & PA_2D \end{array}$$

D, PD, PAD, and PA_2D are CRP, cAMP, $CRP(cAMP)_1$, $CRP(cAMP)_2$, DNA, CRP-DNA, $CRP(cAMP)_1$−DNA, and $CRP(cAMP)_2$−DNA, respectively. K_1 and K_2 are the microscopic association constants for binding of the first and second molecule of cAMP to CRP, respectively. K_3, K_4, and K_6 are the association constants for the formation of $CRP(cAMP)_1$−DNA, $CRP(cAMP)_2$−DNA, and CRP−DNA complexes, respectively. K_5 and K_7 can be calculated from the quotients $K_5 = K_2K_4/K_3$ and $K_7 = K_1K_3/K_6$, respectively. Therefore, the cAMP dependence of CRP binding to the DNA sequences was examined to dissect out the affinities of the CRP conformers with the DNA sequences employed in this study. The dependence of the apparent equilibrium constants for CRP−DNA complexes as a function of cAMP concentration is depicted in Figure 4 (the length of the DNA was 26 base pairs). It is clear from the data that all sequences examined here show a similar biphasic response to the cAMP concentration. The affinity of DNA for CRP was observed to decrease at very low and high cAMP concentrations (approximately <50 μM and >1 mM, respectively), while the strongest affinity of CRP for the DNA sequences was observed with a cAMP concentration in the range of 0.1−1 mM (see Figure 4). In the absence of cAMP, DNA−CRP complex can only be observed at very high concentrations of CRP yielding apparent affinities similar to that of the nonspecific DNA sequence. This biphasic response was found to be independent of the sequence and DNA length (data not shown). The dependence of the apparent equilibrium constants for CRP−DNA complexes as a function of cAMP concentration was fit, in accordance with the thermodynamic reaction scheme depicted above, to the following:

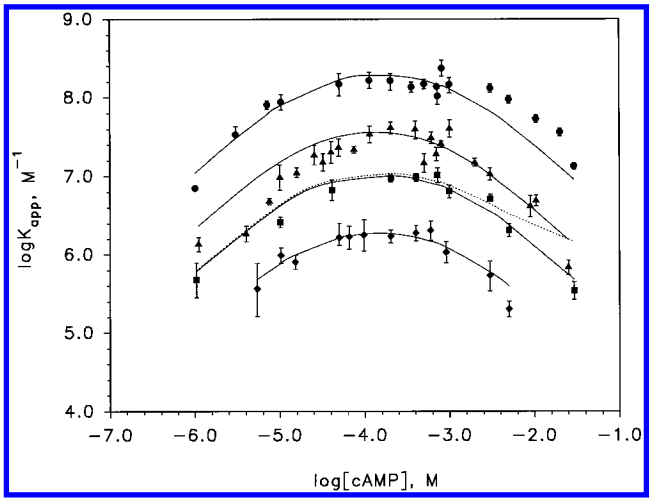

FIGURE 4: cAMP dependence of CRP binding to DNA sequences in TEK(100) buffer (pH 7.80) at 25.0 °C. The symbols correspond to the following DNA sequences: *gallac* (●), *lac* site PI (▲), *gal* site PI (■), and *crp* site PI (◆). The length of the DNA was 26 base pairs. The error bars represent the 75% confidence intervals of the K_{app} values at each cAMP concentration. The data were fit to eq 5 (in Results), and the best fits are shown by the solid curves and with $K_4 = 0$. The dashed curve represents the best fit of the cAMP dependence for the affinity of CRP for *gal* site I where $K_4 \neq 0$ during the data analysis for comparison purposes.

$$K_{app} = \frac{[PD] + [PAD] + [PA_2D]}{([P] + [PA] + [PA_2])[D]} \qquad (4)$$

The concentration of PD approaches 0 since no specific interaction was observed between DNA and CRP in the absence of cAMP. Taking this information into consideration along with expression of the concentration of species in terms of microscopic equilibrium constants, eq 4 becomes

$$K_{app} = \frac{K_1K_3[A] + K_1K_2K_4[A]^2}{1 + K_1[A] + K_1K_2[A]^2} \qquad (5)$$

The values for the microscopic association constants for cAMP bound to CRP, K_1 and K_2, obtained from fitting the data in this study to eq 5, are in good agreement with the values obtained from direct equilibrium measurements of cAMP binding to CRP (see Table 1; Heyduk & Lee, 1989). The equilibrium constants K_3 and K_4 were obtained from

fitting the data to eq 5 (Table 1). The best fits of the data were obtained when K_4 was equal to 0. When K_4 is allowed to vary during the data analysis, then the fitted curves show nonrandom distribution of residuals and tend to shift to considerably higher values at high cAMP concentrations (in the millimolar range) than the experimental data; although, the data points at low cAMP concentrations continue to fit well. The difference in the fitted curves when $K_4 = 0$ and $K_4 \neq 0$ is shown for the cAMP dependence of CRP binding to the *gal* binding site for comparison purposes (see Figure 4). Similar values for K_1, K_2, and K_3 are obtained with K_4 equal to 0 or allowed to vary during the analysis, and the errors in the parameters are smaller when K_4 is equal to 0. Thus, when K_4 is set to 0, the fitted curve is a better representation of the experimental data. The decrease in affinity of CRP for the DNA sequences at high concentrations of cAMP (in the millimolar range) cannot be attributed to the increase in phosphate concentration (Heyduk & Lee, 1990). The affinity of CRP binding to *lac* (26 base pairs in length) in the presence of 200 μM cAMP in 50 mM phosphate, 1 mM EDTA, and 101 mM K^+ (pH 7.80) at 25.0 °C [$K_{app} = 5.1$ (4.4, 5.9) $\times 10^7$ M^{-1}] was found to be the same as the apparent affinity of the CRP−*lac* complex measured in the presence of 200 μM cAMP in TEK(100) buffer (pH 7.80) at 25.0 °C (see Table 1).

It is important to note that the cAMP concentration dependencies of the association constants of DNA−CRP complexes are similar to one another even though the affinities of the DNA sequences for CRP are different. It was also found that the apparent equilibrium constants determined for these CRP binding sites at 200 μM cAMP are similar to the values of K_3 and are therefore a good estimation of the affinity of the DNA sequences for CRP-$(cAMP)_1$. This estimation will hold true for all DNA sequences that have a value of K_3 that is much greater than the value of K_4. Due to the very low affinity of *crp* site PII, *cat* PII, and modified *cat* PII for CRP at 200 μM cAMP, the more rigorous cAMP concentration dependence measurements were not performed. However, the affinity was measured at very low and very high cAMP concentrations for *crp* site PII binding to the protein. These measurements resulted in $K_{app} = 5.4$ (4.4, 6.3) $\times 10^5$ M^{-1} and $K_{app} = 3.1$ (−2.0, 8.3) $\times 10^4$ M^{-1} at 20 μM and 10 mM cAMP, respectively. Similar results were obtained for CRP binding to *cat* PII and modified *cat* PII relative to those of CRP binding to *crp* site PII at 10.0 mM cAMP, yielding $K_{app} = 6.9$ (5.1, 8.7) $\times 10^4$ M^{-1} and $K_{app} = 7.6$ (5.6, 9.9) $\times 10^4$ M^{-1} for *cat* PII and modified *cat* PII, respectively. Furthermore, these CRP binding sites did not bind to CRP in the absence of cAMP. This indicates that the $CRP(cAMP)_1$ conformer has the strongest affinity for all CRP binding sites investigated in this study. Given these results, it is valid to assume that the K_{app} measured at 200 μM cAMP is a fairly good estimate of K_3, and therefore, this value can be used to make a direct comparison of the apparent equilibrium constants determined for CRP binding to *crp* site PII, *cat* PII, modified *cat* PII, and the random sequence with the equilibrium constants of CRP binding to the other DNA sequences examined in this study.

DISCUSSION

There is substantial variation in the natural DNA sequences employed in this thermodynamic study of CRP−DNA interactions. Nevertheless, results of this study indicate that CRP employs the same motif in selecting promoters. The $CRP(cAMP)_1$ species exhibits the highest affinity for specific DNA sequences. Furthermore, the DNA sequence outside of the two half sites apparently plays a role in defining the energetics of CRP−DNA complex formation.

Table 2: Change in Free Energy of CRP−DNA Complexes[a]

DNA	-ΔG, kcal/mol^{-1},[b] 40 base pairs[c]	26 base pairs
lac site PI	11.3 (11.2, 11.5)	10.5 (10.3, 10.5)
gal site PI	10.6 (10.3, 10.9)	9.72 (9.38, 9.94)
gallac	11.6 (11.4, 11.8)	11.5 (11.1, 11.7)
crp site PI	10.0 (9.80, 10.2)	8.72 (8.48, 8.89)
crp site PII	8.79 (8.61, 8.92)	8.20 (8.09, 8.31)[c]
cat site PII	8.99 (8.72, 9.16)	8.25 (8.17, 8.35)[c]
modified *cat* PII	9.10 (8.72, 9.23)	8.74 (8.55, 8.92)[c]
random	7.76 (7.68, 7.83)	6.34 (5.78, 6.64)[c]

[a] Determined from equilibrium association constants measured in TEK(100) buffer (pH 7.80) at 25.0 °C. [b] Calculated from $\Delta G = -RT \ln K_3$ at 25.0 °C. [c] ΔG was calculated using the apparent equilibrium constants.

CRP can be considered a molecular switch for turning on and off the transcription of many bacterial genes. It has been previously proposed that CRP may be regulating transcription by using different conformational states that are available to CRP (Heyduk & Lee, 1990). It is clear from the data presented here that $CRP(cAMP)_1$ is the high-affinity form responsible for the transcriptional regulation of the CRP binding sites examined. These results imply that the mechanism of transcriptional regulation by CRP may be much more subtle than choosing different cAMP-liganded states of CRP. Heyduk and Lee (1989) proposed that one of the rationales for $CRP(cAMP)_1$ to be the active species is structural so that the asymmetric CRP−ligand complex can form a complex with the asymmetric DNA sequence that provides directionality. Certainly the results from this study are consistent with this hypothesis. Apparently, the feature of negative cooperativity in cAMP binding is to ensure the presence of the $CRP(cAMP)_1$ species and the absence of the $CRP(cAMP)_2$ species under physiological concentrations of cAMP.

The difference in energetics of DNA binding to CRP is modest. It is clear from these results that there is a minimal requirement of ~8 kcal/mol for a favorable specific interaction between CRP and a binding site (see Table 2). The interaction of the random sequence with CRP was found to be weaker by ~1.0−2 kcal/mol than that of the weakest CRP−DNA complexes, including *crp* site PII, which does not contain the recognition motifs found in most CRP binding sites. This implies that there is information in site-specific DNA sequences for CRP other than the consensus sequence. The total difference in the change of free energies observed for specific CRP−DNA interactions is approximately 3 kcal/mol. These observations are true for the protein interacting with CRP binding sites that are 40 and 26-base pairs long. This implies that the 26-base pair sequences contain sufficient information necessary for CRP recognition. The sequences outside the 26-base pair region only serve to slightly strengthen the binding energetics. The modest difference in the change of free energy observed between site-specific DNA sequences bound to CRP is consistent with the DNA sequence-dependent studies that were measured on other

DNA−protein systems such as *lac* (Mossing & Record, 1986), λ (Sarai & Takeda, 1989), and *cro* (Takeda *et al.*, 1989) operators and the *Eco*RI site (Lesser *et al.*, 1990).

The regulation of transcription involves many factors, and a simple examination of the linkages involved in cAMP binding to CRP and followed by CRP(cAMP)$_1$ binding to DNA provides a rationale for these small differences in the affinities of CRP for different CRP binding sites. The ratio of DNA bound to CRP per total DNA concentration can be described by the following equation:

$$\frac{[\text{CRP}-\text{DNA complex}]}{[\text{DNA}]_{\text{total promoter site}}} = \frac{[\text{PAD}] + [\text{PA}_2\text{D}]}{[\text{D}]_o + [\text{PAD}] + [\text{PA}_2\text{D}]} \quad (6)$$

where

$$D_T = [D]_o + [\text{PAD}] + [\text{PA}_2\text{D}] \quad (7)$$

and

$$[\text{P}]_T = [\text{P}]_o + [\text{PA}] + [\text{PA}_2] + [\text{PAD}] + [\text{PA}_2\text{D}] \quad (8)$$

The concentration terms in the equations were subsequently expressed in the form of equilibrium constants. Then, substituting the expression for $[\text{D}]_o$ into eq 8 and solving for the free protein concentration, one can obtain the expression for $[\text{P}]_o$:

$$[\text{P}]_o = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad (9)$$

where

$$a = (K_1K_3[\text{A}] + K_1K_2K_4[\text{A}]^2)(1 + K_1[\text{A}] + K_1K_2[\text{A}]^2) \quad (10)$$

$$b = 1 + K_1[\text{A}][(1 + K_2[\text{A}]) + (K_3 + K_2K_4[\text{A}])([\text{D}]_T - [\text{P}]_T)] \quad (11)$$

$$c = -[\text{P}]_T \quad (12)$$

Now eq 6 can be expressed in the following manner:

$$\frac{[\text{CRP}-\text{DNA complex}]}{[\text{DNA}]_{\text{total promoter site}}} = \frac{K_1K_3[\text{A}][\text{P}]_o + K_1K_2K_4[\text{A}]^2[\text{P}]_o}{1 + K_1K_3[\text{A}][\text{P}]_o + K_1K_2K_4[\text{A}]^2[\text{P}]_o} \quad (13)$$

$[\text{D}]_T$ and $[\text{P}]_T$ are estimations of the total promoter site concentration (2×10^{-9} M; Takahashi *et al.*, 1989) and total cellular concentration of CRP (2×10^{-6} M; Anderson *et al.*, 1971; Berg & von Hippel, 1988). The curves were simulated using eq 13 and the values of 4.5×10^4 M^{-1} and 1.0×10^3 M^{-1} for K_1 and K_2, respectively. These values were taken from direct binding measurements (Heyduk & Lee, 1989). The ratio of $[\text{CRP}-\text{DNA complex}]/[\text{DNA}]_{\text{total promoter site}}$ as a function of cAMP concentration is shown in Figure 5. The simulations of the fractional saturation of CRP binding sites by CRP depicted in Figure 5A use the experimental values for K_3 determined in this study (see Table 1) and $K_4 = 0$. It is clear that the cAMP concentration necessary for obtaining fractional saturation of DNA by CRP is within the physiological range (~1−2 μM). The extent of fractional saturation of the binding site by CRP is also dependent on the affinity of the DNA for CRP(cAMP)$_1$. The simulated binding curves show that the high-affinity CRP binding sites are very responsive to the concentration of cAMP and therefore require very low concentrations of cAMP for initiation of transcriptional processes. For example, the *lac* promoter begins to respond to cAMP concentration at values as low as 3×10^{-7} M and approximately 80% of the promoter site is occupied at 10^{-6} M cAMP, as shown in curve 2 of Figure 5A. However, the high affinity sites are quite insensitive to fluctuations in

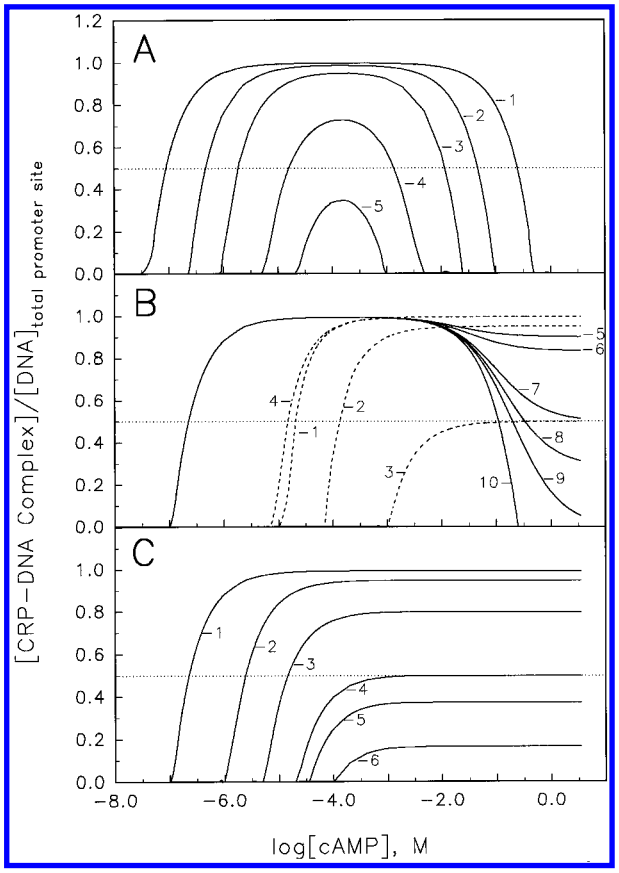

FIGURE 5: Simulations of the fraction of DNA bound to CRP with respect to the total promoter site concentration as a function of cAMP concentration. The curves were simulated using eq 13 (in Discussion). The values of K_1 and K_2 used in the calculation are from Heyduk and Lee (1989). The values of 2×10^{-9} M (Takahashi *et al.*, 1989) and 2×10^{-6} M (Anderson *et al.*, 1971; Berg & von Hippel, 1988) were used as estimates for the cellular concentrations of total promoter site concentrations and CRP, respectively. The dotted line in each graph is drawn to indicate half-saturation of the CRP binding sites on DNA by CRP. Curves are represented by the number in parentheses. (A) $K_4 = 0$ M^{-1}, and experimentally determined values of K_3 were used (see Table 1 for values): *gallac* (1), *lac* site PI (2), *gal* site PI (3), *crp* site PI (4), and *crp* site PII (5). (B) Dashed curves represent simulations where $K_3 < K_4$. The equilibrium values used are as follows: $K_3 = 0$ M^{-1} and $K_4 = 1 \times 10^8$ M^{-1} (1), $K_3 = 0$ M^{-1} and $K_4 = 1 \times 10^7$ M^{-1} (2), $K_3 = 0$ M^{-1} and $K_4 = 1 \times 10^6$ M^{-1} (3), and $K_3 = 1 \times 10^6$ M^{-1} and $K_4 = 1 \times 10^8$ M^{-1} (4). Solid curves represent simulations where $K_3 > K_4$ and $K_4 \neq 0$. $K_3 = 1 \times 10^8$ M^{-1} and was held constant for these simulations, and the value of K_4 was varied. The values of K_4 used are as follows: 5×10^6 M^{-1} (5), 3×10^6 M^{-1} (6), 1×10^6 M^{-1} (7), and 7×10^5 M^{-1} (8), 5×10^5 M^{-1} (9), 1×10^5 M^{-1} (10). (C) Simulations of the cAMP dependence on the fractional saturation of CRP binding sites when the affinities for DNA binding to CRP(cAMP)$_1$ and CRP(cAMP)$_2$ are the same, $K_3 = K_4$. The values of the equilibrium association constants used are as follows: 1×10^8 M^{-1} (1), 1×10^7 M^{-1} (2), 2.5×10^6 M^{-1} (3), 1×10^6 M^{-1} (4), 8×10^5 M^{-1} (5), and 6×10^5 M^{-1} (6).

cAMP concentration greater than 10^{-6} M, since the plateau region of fully saturated DNA by CRP spans a large cAMP concentration range. In contrast, CRP binding sites that have a low affinity for CRP, such as the *crp* promoter site, require higher concentrations of cAMP to form protein−DNA complex. Furthermore, low-affinity CRP binding sites are much more sensitive to fluctuations in cAMP concentration, as shown by curves 4 and 5 in Figure 5A. For example, curve 4 indicates that a small plateau region is observed and a 10-fold increase in cAMP concentration from the highest degree of saturation reduces the fraction of DNA sites occupied by CRP to zero. Thus, a strong affinity of CRP for DNA promoter site leads to a formation of protein−DNA complex at low cAMP concentrations, but the interaction of a low-affinity DNA binding site for CRP results in a biphasic behavior in the occupancy of promoter site. Do these simulated behaviors make biological sense? Transcriptional activation of the *lac* and *gal* genes results in the transport and metabolism of carbohydrates. The bacterium needs to respond to changes in energy source rapidly, i.e. at small changes in cAMP concentration. Thus, a very low concentration of cAMP is needed to activate the transcriptional process, and therefore, the cell is able to respond to small fluctuations in the level of carbohydrates. In contrast, regulation of transcription of *crp* would benefit from both a repression and a derepression process, as shown by the biphasic behavior shown in curves 4 and 5 in Figure 5A. As the initial demand for rapid action is required, an increase in the cellular cAMP concentration leads to formation of more cAMP−CRP complex and an increased transcriptional activity of genes. Subsequently, an increase in cAMP concentration results in repression of the *crp* gene by binding CRP−cAMP to the *crp* PI site. A repression of the *crp* gene would keep the concentration of CRP from increasing. This would lead to a faster increase in the concentration of cAMP−CRP complex available to activate transcriptional processes with any further increase in cAMP concentration. If further activation of operons is required, as reflected by a further increase in cAMP concentration, then formation of the CRP(cAMP)$_2$ complex may occur. This weakens the CRP−DNA complex, and a derepression of the *crp* gene is observed. A weak affinity for the *crp* operon by cAMP−CRP provides such a biphasic regulating mechanism.

Previously published work suggests that the bindings of CRP to *crp* sites PI and PII differ in their cAMP and CRP dependencies (Hanamura & Aiba, 1992), implying that different mechanisms of control may be effective. These observed differences can be easily explained from linkage theory without involving different mechanisms. The affinity of CRP for *crp* site PII is ∼2-fold less than that for *crp* site PI, and therefore, higher concentrations of cAMP will be needed to obtain fractional saturation of the binding site (see curve 5 in Figure 5A). A comparison between curves 4 and 5 demonstrates the effect of a simple small difference of 2-fold in affinity. No additional change in mechanism needs to be involved. The simulations show fractional saturation of the *crp* binding sites at higher cAMP concentrations than the cAMP concentrations used in the *in vitro* and *in vivo* studies. The experiments performed *in vitro* and *in vivo* monitored the binding of CRP to these two sites simultaneously, and thus, more than two reactions were proceeding during the measurement. The linked reactions can shift the fractional saturation of the CRP−DNA complexes to lower concentrations of cAMP as indicated in this simulation study by linking the cAMP and DNA binding reactions. The work by Aiba and co-workers (1983, 1991) nicely illustrates the strength of linked reactions in a multiple reaction system. It is clear from this study that the differences observed in the cAMP and CRP concentrations necessary to bind the two sites are not necessarily due to different cAMP-liganded states of CRP.

The results of this study establish the potential significance of the species of CRP(cAMP)$_1$ in regulation of transcription in CRP-dependent promoters; however, this conclusion may not totally rule out any potential role for the CRP(cAMP)$_2$ species. To further explore this possibility, additional simulations were conducted. Other systems may invoke other possible scenarios for the transcription factor CRP. Let us first examine the possibility of promoter sites controlled solely by the CRP(cAMP)$_2$ conformer. The dashed curves 1−3 in Figure 5B show the degree of saturation of a promoter site by CRP(cAMP)$_2$ as a function of cAMP concentration. The K_4 values range from 1×10^8 to 1×10^6 M^{-1}. The cAMP concentration necessary to reach half-saturation of the CRP binding site by CRP(cAMP)$_2$ is shifted to higher concentrations relative to those in Figure 5A ($K_3 > 0$ and $K_4 = 0$), and the extent of saturation is dependent on the affinity of CRP(cAMP)$_2$ for the CRP binding sites. Systems that have weak affinity for the CRP(cAMP)$_2$ complex will result in <100% saturation of the binding site. Systems of this sort are finely controlled by a very narrow range of cAMP concentrations, and fluctuations at high concentrations of cAMP are ineffective at perturbing the degree of saturation of the promoter site. Transcription is either on or off.

Figure 5C illustrates the possibility where transcriptional control is governed by both CRP(cAMP)$_1$ and CRP(cAMP)$_2$ and that the affinities of these two conformations for the CRP binding sites on DNA are equal to one another. It is apparent that, as the affinities of these conformers weaken for the binding site, the highest level of fractional saturation of the binding site is decreased and the cAMP concentration required to attain fractional saturation of the binding site shifts to higher cAMP concentrations. These systems are relatively insensitive to fluctuations in the cAMP concentration (less than or equal to micromolar) when it is greater than the concentration needed to have CRP occupy the binding site. However, transcriptional control (on or off) by CRP is more sensitive to small fluctuations in the cAMP concentration ($\leq \mu$M) when both conformers bind equally well to the promoter site. The only method to determine if any one of these simulated linkages is truly found in nature is to quantitatively determine the affinity of CRP for the CRP binding site as a function of cAMP concentration. Then one is able to dissect out which conformer(s) of CRP is responsible for the transcriptional control of that particular promoter site.

Curve 4 in Figure 5B illustrates a case where $K_3 < K_4$ and $K_3 \neq 0$ ($K_3 = 1 \times 10^6$ M^{-1} and $K_4 = 1 \times 10^8$ M^{-1}). A CRP binding site that has a weak affinity for CRP(cAMP)$_1$ and a strong affinity for CRP(cAMP)$_2$ will result in cAMP dependence on the saturation of the binding site that is dominated by CRP(cAMP)$_2$. There is only a slight shift to lower cAMP concentrations to obtain total saturation of the binding site relative to the case where only CRP(cAMP)$_2$

binds to the CRP binding site with high affinity (compare curves 1 and 4). The solid curves in Figure 5B illustrate the possibility of a high affinity for CRP(cAMP)$_1$ but weaker affinities for CRP(cAMP)$_2$ (where $K_4 \neq 0$). The first portions of the curves (curves 5−10) are similar to one another and also similar to those in Figure 5A (see curve 1 for direct comparison). A decrease in the affinity of the binding site for CRP(cAMP)$_2$ results in the binding site on DNA not being fully saturated at molar concentrations of cAMP. It is only when the affinity of CRP(cAMP)$_2$ is very weak, 1×10^5 M^{-1}, that the CRP(cAMP)$_2$ conformer probably does not have the capability of controlling transcriptional processes. Curve 10 illustrates this point and is similar to the fractional saturation of the CRP binding sites examined in this study (see Figure 5A). So binding sites that have a very weak affinity for CRP(cAMP)$_2$ are not able to control the fractional saturation of the binding site. Curves 5−9 show the possibility that CRP(cAMP)$_2$ and CRP(cAMP)$_1$ are capable of controlling transcriptional processes. The simulations show that CRP(cAMP)$_2$ would be an active conformer at very high cAMP concentrations (in molar range) that are physiologically impractical; however, these simulations are not just a trivial exercise. It is important to remember that only two of the reactions in a multiple reaction pathway are being linked in this simulation. There could be other reactions that could shift these curves so that the CRP(cAMP)$_2$ conformer would be prevalent at physiological concentrations of cAMP. The CRP(cAMP)$_2$ conformer would predominate at physiological conditions and would be capable of regulating the fraction of DNA sites occupied by CRP if the CRP(cAMP)$_2$−DNA complex had a greater affinity for another ligand, i.e. RNA polymerase, relative to the affinity of the CRP(cAMP)$_1$−DNA complex for that ligand.

The linkage of CRP−DNA complexes with other transcriptional processes can be explored further as alluded to above, and the reactions can be linked to other interactions such as the CRP−DNA complex interacting with RNA polymerase. It is evident from Figure 5A that saturation of CRP binding sites examined in this study occurs within the physiological range of cAMP concentrations. Discrimination between various CRP binding sites still occurs due to their differences in affinity for CRP(cAMP)$_1$. Coupling the equilibria for these reactions clearly shows us how small changes in the equilibrium constant in one step of the overall transcriptional mechanism can be transformed into larger differences in the pathway. Linked reactions give the cell the ability to efficiently fine tune regulation pathways.

ACKNOWLEDGMENT

We thank Dr. Petr Herman for writing the software program used to generate random numbers. We dedicate this paper to the memory of Jeffries Wyman (1901−1995), who introduced to the field of biothermodynamics the concept of linked functions.

REFERENCES

Aiba, H. (1983) *Cell 32*, 141−9.
Anderson, W. B., Schneider, A. B., Emmer, M., Perlman, R. L., & Pastan, I. (1971) *J. Biol. Chem. 246*, 5929−37.
Barber, A. M., & Zhurkin, V. B. (1990) *J. Biomol. Struct. Dyn. 8*, 213−32.
Barber, A. M., Zhurkin, V. B., & Adhya, S. (1993) *Gene 130*, 1−8.
Berg, O. G., & von Hippel, P. H. (1988) *J. Mol. Biol. 200*, 709−23.
Crothers, D. M., & Steitz, T. A. (1992) in *Transcriptional Regulation*, pp 501−32, Cold Spring Harbor Press, Cold Spring Harbor, NY.
Dalma-Weiszhauz, D. D., Gartenberg, M. R., & Crothers, D. M. (1991) *Nucleic Acids Res. 19*, 611−6.
de Crombrugghe, B., Busby, S., & Buc, H. (1994) *Science 224*, 831−8.
Ebright, R. H. (1993) *Mol. Microbiol. 8*, 797−802.
Ebright, R. H., Ebright, Y. W., & Gunasekera, A. (1989) *Nucleic Acids Res. 17*, 10295−305.
Gartenberg, M. R., & Crothers, D. M. (1988) *Nature 333*, 824−9.
Gaston, K., Chan, B., Kolb, A., Fox, J., & Busby, S. (1988) *Biochem. J. 253*, 809−18.
Hanamura, A., & Aiba, H. (1991) *Nucleic Acids Res. 19*, 4413−9.
Hanamura, A., & Aiba, H. (1992) *Mol. Microbiol. 6*, 2489−97.
Haughland, R. P. (1992) in *Molecular Probes Handbook of Fluorescent Probes and Research Chemicals* (Larison, K. D., Ed.) 5th ed., Molecular Probes, Inc., Eugene, OR.
Heyduk, E., Heyduk, T., & Lee, J. C. (1992) *Biochemistry 31*, 3682−8.
Heyduk, T., & Lee, J. C. (1989) *Biochemistry 28*, 6914−24.
Heyduk, T., & Lee, J. C. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 1744−8.
Heyduk, T., Lee, J. C., Ebright, Y. W., Blatter, E. E., Zhou, Y., & Ebright, R. H. (1993) *Nature 364*, 548−9.
Kolb, A., Busby, S., Herbert, M., & Kotlarz, D. (1983) *EMBO J. 2*, 217−22.
Kolb, A., Busby, S., Buc, H., Garges, S., & Adhya, S. (1993) *Annu. Rev. Biochem. 62*, 749−95.
Le Grice, S. F., Matzura, H., Marcoli, R., Iida, S., & Bickle, T. A. (1982) *J. Bacteriol. 150*, 312−8.
Lesser, D. R., Kurpiewski, M. R., & Jen-Jacobson, L. (1990) *Science 250*, 776−86.
Okamoto, K., & Feundlich, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 5000−4.
Sarai, A., & Takeda, Y. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 6513−7.
Takahashi, M., Blazy, B., Baudras, A., & Hillen, W. (1989) *J. Mol. Biol. 207*, 783−96.
Takeda, Y., Sarai, A., & Rivera, V. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 439−43.
Weber, G. (1992) in *Protein Interactions*, pp 14−8, Chapman and Hall, New York.
Windholz, M., Ed. (1976) in *The Merck Index*, 9th ed., Merck & Co., Inc., Rahway, NJ.
Zhou, Y., Merkel, T. J., & Ebright, R. H. (1994) *J. Mol. Biol. 243*, 603−10.

BI952187Q