

Interplay between Site-Specific Mutations and Cyclic Nucleotides in Modulating DNA Recognition by *Escherichia coli* Cyclic AMP Receptor Protein^{†,‡}

Jiayin Dai,[§] Shwu-Hwa Lin,[§] Carly Kemmis, Anita J. Chin, and J. Ching Lee*

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

Received January 8, 2004; Revised Manuscript Received March 31, 2004

ABSTRACT: Mutagenesis of various amino acids in *Escherichia coli* cyclic AMP receptor protein (CRP) has been shown to modulate protein compressibility and dynamics [Gekko et al. (2004) *Biochemistry* 43, 3844–3852]. Cooperativity of cAMP binding to CRP and the apparent DNA binding affinity are perturbed [Lin and Lee (2002) *Biochemistry* 41, 11857–11867]. The aim of this study is to explore the effects of mutation on the surface chemistry of CRP and to define the consequences of these changes in affecting specific DNA sequence recognition by CRP. Furthermore, the role of the interplay between mutation and specific identity of the bound cyclic nucleotide in this DNA recognition was explored. In the current study, effects of eight site-specific mutations (K52N, D53H, S62F, T127L, G141Q, L148R, H159L, and K52N/H159L) on DNA recognition of four sequences (Class I (site PI of *lac*), Class II (site PI of *gal*), and synthetic sequences that are hybrids of Classes I and II sites) modulated by three different cyclic nucleotides (cAMP, cCMP, and cGMP) were investigated. All mutations altered the surface chemistry of CRP as evidenced by the change in elution properties of these proteins from different matrixes. While T127L, S62F, K52N, and H159L exhibited unexpected behavior under combinations of specific experimental conditions, such as the identity of bound cyclic nucleotide and DNA sequence, in general, results showed that the affinities of CRP for DNA were sequence-dependent, increasing in the order of *lacgal26* < *gal26* < *lac26* < *gallac26* for all the mutants in the presence of 200 μ M cAMP. The apparent association constants significantly increased in the order of no cyclic nucleotide \approx cGMP < cCMP < cAMP for all the examined DNA sequences. Linear correlation between the ΔG for CRP–DNA complex formation and the cooperativity energy for cAMP binding was observed with *gallac26*, *gal26*, and *lacgal26*; however, the slope of this linear correlation is DNA sequence dependent. Structural information was presented to rationalize the interplay between CRP sequence and cyclic nucleotides in defining the recognition of DNA sequences.

Escherichia coli cyclic AMP receptor protein (CRP)¹ regulates the expression of more than 100 genes that are responsive to the fluctuation of the cellular concentration of cAMP during its cell cycle (1–4). Serving as an activator, cAMP induces conformational changes in the protein, which in turn confer the CRP–cAMP complex with a capability to recognize specific DNA sequences at or near the promoter. CRP is a dimer of two identical subunits, each consisting of

209 amino acids (5, 6). Each subunit is composed of two distinct functional domains that are connected by a hinge region (residues 134–138), as shown in Figure 1. The N-terminal domain is responsible for dimerization of CRP and for interaction with cAMP. The C-terminal domain contains the helix–turn–helix motif that is responsible for DNA recognition and binding (7). Binding of cyclic nucleotide to the N-terminal domain leads to an allosteric activation of CRP, which binds to specific DNA sequences with significantly higher affinity (1–4). Furthermore, the binding of cyclic nucleotide to their binding sites is characterized by positive cooperativity; thus, there is communication between binding sites (8–10). Hence, interdomain and intersubunit communications are essential features of normal functioning of CRP.

Each CRP subunit consists of a typical cyclic nucleotide binding motif, which consists of residues 18–50, 61–73, and 81–97 (7, 11). The network of communications is involved in interdomain and intersubunit interactions that seem to include residues 50–61 (loop 3), while residues 73–

[†] Supported by NIH Grant GM45579 and by the Robert A. Welch Foundation Grants H-0013 and H-1238.

[‡] Dedicated to the memory of Genevieve Ching-Wen Lee (1974–1994).

* Corresponding author. Tel: (409) 772–2281. Fax: (409) 772-4298. E-mail: jcllee@utmb.edu.

[§] Equal contribution from these authors.

¹ Abbreviations: cAMP, 3',5'-cyclic adenosine monophosphate; cGMP, 3',5'-cyclic guanosine monophosphate; cCMP, 3',5'-cyclic cytidine monophosphate; cNMP, 3',5'-cyclic nucleotide monophosphate; WT, wild type; CRP, cAMP receptor protein; CPM, *N*-[4-[7-(diethylamino)-4-methylcoumarin-3-yl]phenyl]maleimide; Tris, Tris(hydroxymethyl)-aminomethane; TEK(100), 50 mM Tris, 100 mM KCl, 1 mM EDTA at pH 7.8 and 25 °C; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis.

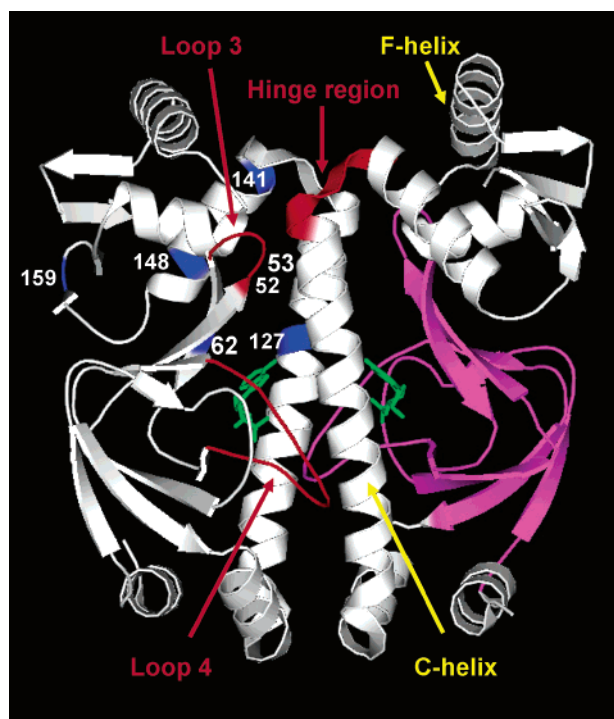


FIGURE 1: Structure of CRP dimer. Each subunit binds one cAMP (labeled in green). Interactions between the two subunits involve the hinge region and loop3 (labeled in red) and C-helix (yellow arrow). F-helix (yellow arrow) is the DNA recognition helix. Cyclic nucleotide binding motif is highlighted in magenta color (amino acids 18–97), especially the loop 4 (red color). Locations of mutation sites are highlighted in blue, except amino acids 52 and 53 in loop 3 (red). The structure was constructed using the PyMOL Molecular Graphics System (35) with input pdb file 1G6N from the Protein Data Bank.

81 (loop 4) are apparently involved in contacts with the coiled-coil C-helices (residues 111–134), which constitute a significant portion of the dimer interface (7, 11). Genetic data have identified some residues in this network as important in the normal mode of functioning in CRP (12–15). In a series of *in vitro* studies, Lin et al. showed that signals due to mutations involving K52N, D53H, S62F, T127L, G141Q, L148R, H159L, and K52N/H159L are mostly transmitted through the subunit interface (8, 16), although these mutation sites are distributed in various regions of the CRP molecule. For example, residues 52 and 53 are in loop 3; residues 62 and 127 are in the cyclic nucleotide binding motif, while the rest are in the DNA binding domain. The locations of these mutation sites are shown in Figure 1. These mutations affect neither the stoichiometry of cAMP binding nor the binding affinity of the first molecule of cAMP, except the S62F mutant, which shows a ~10-fold decrease in affinity (8). The cooperativity of cAMP binding is affected by mutation. It ranges from negative to positive cooperativity (8). With the exception of the T127L mutant, the ΔG for DNA–CRP complex formation increases linearly with increasing ΔG associated with the cooperativity of cAMP binding. This linear relationship implies that the protein molecule modulates the signal in the binding of cAMP (8). These results were obtained with only cAMP and *lac26* as the allosteric activator and targeted DNA binding site, respectively. At present, it is not known whether the modulation mechanisms for WT and mutant CRPs are similar. Does mutation change the specificity for cyclic

nucleotide binding and DNA recognition which is modulated by cyclic nucleotides?

In a recent study, Gekko et al. reported a correlation between the effects of amino acid substitutions on the functional energetics and global structural properties in CRP (17). Mutations induce large changes in the partial specific volume and adiabatic compressibility from the corresponding values for the wild-type protein. These changes in global structural properties correlate with the rate of amide proton exchange. A linear correlation was established between adiabatic compressibility and the energetics of cooperativity of binding of cAMP molecules to the high affinity sites, regardless of the nature of cooperativity be it negative or positive. This linear correlation indicates that the nature and magnitude of cooperativity are a continuum. The protein is designed so that its functional properties can be modulated by not only changing the magnitude but also the nature of cooperativity. A similar linear correlation was established between compressibility and DNA binding affinity. Further linear correlations were established among the dynamics of CRP and functional energetics. Double mutation (K52N/H159L) at positions 52 and 159, whose α -carbons are separated by 34.6 Å, showed nonadditive effects on these physical parameter. These results demonstrate that a small alteration in the local structure due to amino acid substitution is dramatically magnified in the overall protein dynamics and that the structural flexibility plays an important role in modulating the allosteric behavior of CRP. The combination of results from the studies by Gekko et al. (17) and Lin et al. (8, 16) indicates that mutations affect the global structural properties of CRP and, as a consequence, affects the communications between cAMP binding sites in adjacent subunits, between cAMP and DNA binding sites.

On the basis of the current understanding, the CRP–DNA complexes assume different geometries as a function of the DNA sequence. The flanking sequences affect the energetics of DNA–CRP complex formation. Studying nine DNA sequences containing CRP binding sites, Lee and colleagues reported that the apparent affinity of CRP for DNA was sequence-dependent (18–22). The energetics of the CRP–DNA complex formation is contributed not only from the two conserved half-sites but also from their adjacent flanking sequences. The relative contributions of the half-sites and flanking sequences to the energetics of DNA recognition are operon specific (19). In addition, similar to other DNA binding proteins, CRP is able to induce DNA bending. *lac* bends symmetrically (19), whereas *gal* bends asymmetrically upon CRP binding (21, 22). Thus, DNA sequence encodes information regarding not only the energetics of CRP–DNA formation but also structural elements such as DNA bending geometry.

The aim of this study is to explore the effects of these CRP mutants on DNA sequence recognition, the interplay between mutation and specific cyclic nucleotide in this DNA recognition. Furthermore, we wish to explore the effects of mutation on the surface chemistry of CRP and to define the consequences of these changes in affecting specific DNA sequences recognition by CRP. Thus, in the current study, both DNA sequence and protein structure were altered to address this issue. The CRP binding sites employed in this investigation were chosen to represent the primary promoter sites of Class I (*lac* PI), Class II (*gal* PI), and the hybrids

Table 1: Sequence of CRP Binding Sites

	sequence of CRP binding sites
<i>gallac</i>	5'-AAAAGTGTGACATGGATCACTTTAGT-3'
<i>lac</i>	5'-ATTAATGTGAGTTAGCTCACTCATT-3'
<i>gal</i>	5'-AAAAGTGTGACATGGAATAAATTAGT-3'
<i>lacgal</i>	5'-ATTAATGTGAGTTAGCATAAACATTA-3'

sequence (*lacgal* and *gallac*). Sequences are listed in Table 1. The purified CRP mutants were tested for their ability to bind specific DNA sequences and the effects of cyclic nucleotides on modulating DNA sequence recognition. The binding of CRP to these sequences in the presence of various cyclic nucleotides was monitored by fluorescence anisotropy.

MATERIALS AND METHODS

Materials. Cyclic AMP, cyclic GMP, cyclic CMP, and Tris are products of Sigma (Saint Louis, MO). CPM, *N*-[4-[7-(diethylamino)-4-methylcoumarin-3-yl]phenyl]maleimide, was purchased from Molecular Probes. The 5'-end phosphorylated oligodeoxyribo-nucleotides were purchased from Fisher Scientific Genosys. All other reagents were of the highest grade commercially available.

All the experiments were conducted in TEK(100). The concentrations of protein, cyclic nucleotides, and fluorescence probe were determined spectrophotometrically with the following extinction coefficients: 40 800 M⁻¹ cm⁻¹ at 278 nm for CRP dimer (9), 14 650 M⁻¹ cm⁻¹ at 259 nm, 12 950 M⁻¹ cm⁻¹ at 254 nm, and 9100 M⁻¹ cm⁻¹ at 271 nm for cAMP, cGMP, and cCMP, respectively; and 33 000 M⁻¹ cm⁻¹ at 385 nm for CPM.

CRP and DNA Preparations. Wild-type and mutant CRPs were prepared from the well-established expression system in this laboratory as described (8, 16). CRPs were isolated and purified by sequential chromatography on Bio-Rex 70, hydroxyapatite (Bio-Rad), and phenyl sepharose (Amersham Pharmacia Biotech, Uppsala Sweden). The salt concentration of each fraction collected was determined by conductivity. All purified proteins are ~99% homogeneous as routinely judged by Coomassie Blue stained SDS-PAGE gels with a loading 50–60 µg/lane of protein. Protein mass was further confirmed with mass spectrometry. Before being used, CRPs were routinely dialyzed against the described buffer, and were then gently filtered through a membrane with a pore size of 0.22 µm (8).

Details for CPM labeled double-stranded DNA preparation were previously described (18, 20).

Determination of DNA Binding Affinity. Fluorescence anisotropy titration was performed with a SLM 8000C spectrofluorometer to quantitatively evaluate the CRP-DNA interaction. Details for sample preparation and instrumentation were previously described (16, 19). Experiments were performed in TEK(100), and data were fitted to the following equation by nonlinear least-squares with Marquardt-Levenberg algorithm provided by Sigmaplot (version 7.0) for Windows to determine the apparent association constant for CRP-DNA interaction, K_x :

$$A = A_D + \Delta A \times \frac{(K_x D_T + K_x P_T + 1 - \sqrt{(K_x D_T + K_x P_T + 1)^2 - 4K_x^2 D_T P_T})}{2K_x D_T} \quad (1)$$

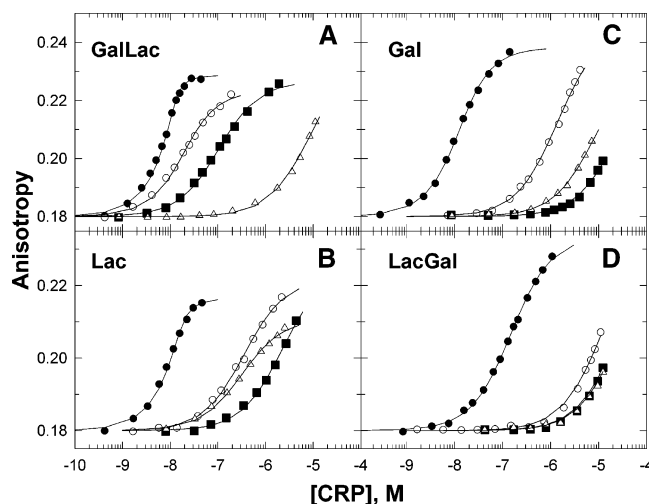


FIGURE 2: Fluorescence anisotropy isotherms of L148R with (A) *gallac*, (B) *lac*, (C) *gal*, and (D) *lacgal* as a function of various cyclic nucleotides: ●, cAMP; ○, cCMP; ■, cGMP; △, without cNMP. Experiments were carried out in TEK(100) at pH 7.8 and 25 °C with DNA concentration in the range from 12 to 20 nM. The solids represent the best fits of the observed anisotropies with eq 1, and the recovered parameters are listed in Table 2.

where A , A_D , and ΔA are the observed anisotropy, the anisotropy of free DNA, and the total change in anisotropy, respectively, D_T and P_T are the total molar concentrations of DNA and protein, and x is A, C, G, or N for cAMP, cCMP, cGMP, or without cNMP, respectively. Binding isotherms were carried out in the presence of 200 µM cyclic nucleotide for WT and all the CRP mutants, except for S62F. When experiments were conducted with S62F the concentration of cyclic nucleotides was set at 1 mM, since the affinity of cAMP with S62F is much weaker than that with WT CRP (8). On the basis of the binding affinities, at the chosen cAMP concentrations these proteins are saturated at the high affinity sites with only an insignificant amount of the low affinity site occupied. The design of experimental conditions is to avoid the inhibitory DNA binding effect upon the binding of cAMP to the low affinity site (8).

RESULTS

Binding Isotherms of L148R CRP. The DNA binding sites employed in the current investigation, as listed in Table 1, represent the primary promoter sites from Class I (*lac* site PI), Class II (site PI of *gal*), and synthetic sequences that are hybrids of Classes I and II sites (*lacgal* and *gallac*). Figure 2 shows representative fluorescence anisotropy titrations of L148R binding to four DNA sequences in the presence of various cyclic nucleotides.² L148R can bind to all four oligomers, *lac*, *gallac*, *gal*, and *lacgal* in the presence of cAMP, cCMP, cGMP, and in the absence of cyclic nucleotide. Each of the individual titration isotherms is well separated, suggesting that the affinity for the same DNA sequence is modulated by binding of specific cyclic nucle-

² L148R CRP is a difficult system to study because it is easily contaminated by trace amounts of nucleic acid during protein purification, rendering the protein to aggregate during the DNA binding assay, and consequently the protein would gradually lose DNA binding affinity during storage. Nevertheless, it is still possible to obtain high-quality titration isotherms, as long as the L148R sample for experiments is freshly purified.

Table 2: The Apparent Associations for the DNA–CRP Complex Formation

	$K_{app} (\times 10^6 M^{-1})$ with cAMP				$K_{app} (\times 10^6 M^{-1})$ with cCMP			
	<i>gallac</i>	<i>lac</i>	<i>gal</i>	<i>lacgal</i>	<i>gallac</i>	<i>lac</i>	<i>gal</i>	<i>lacgal</i>
wt	160 ± 24	49 ± 6.4	12 ± 0.8	0.95 ± 0.02	0.46 ± 0.05	0.07 ± 0.01	0.07 ± 0.01	0.07 ± 0.01
K52N	132 ± 8	6.7 ± 0.4	1.5 ± 0.1	<0.17	0.23 ± 0.01	0.60 ± 0.01	0.18 ± 0.02	ND
D53H	970 ± 170	731 ± 122	216 ± 11	14.9 ± 0.8	83.9 ± 5.9	3.3 ± 0.1	0.60 ± 0.02	<0.18
S62F	103 ± 18	1.9 ± 0.1	0.45 ± 0.02	<0.07	51.1 ± 6.1	1.1 ± 0.1	0.20 ± 0.02	ND
T127L	5.3 ± 0.3	0.12 ± 0.01	<0.05	<0.05	0.29 ± 0.05	<0.11	<0.05	ND
G141Q	264 ± 37	72.1 ± 12	12.1 ± 0.6	1.81 ± 0.13	80.6 ± 2.1	8.04 ± 0.55	1.87 ± 0.07	<0.80
L148R	1911 ± 744	809 ± 277	141.3 ± 10.8	7.22 ± 0.27	79.5 ± 8.8	2.91 ± 0.22	0.72 ± 0.03	<0.08
H159L	215.4 ± 22.5	38.9 ± 4.3	9.4 ± 0.4	0.62 ± 0.02	0.10 ± 0.01	ND ^a	ND	ND
K52N/H159L	103.4 ± 11	3.2 ± 0.2	1.1 ± 0.1	<0.19	<0.06	<0.05	ND	ND

	$K_{app} (\times 10^6 M^{-1})$ with cGMP				$K_{app} (\times 10^6 M^{-1})$ without cNMP			
	<i>gallac</i>	<i>lac</i>	<i>gal</i>	<i>lacgal</i>	<i>gallac</i>	<i>lac</i>	<i>gal</i>	<i>lacgal</i>
wt	<0.05	<0.07	<0.07	<0.08	ND	ND	ND	ND
K52N	<0.15	ND	ND	ND	ND	ND	ND	ND
D53H	10.7 ± 0.6	0.48 ± 0.03	<0.13	<0.1	<0.08	2.33 ± 0.03	0.69 ± 0.07	ND
S62F	ND	ND	ND	ND	ND	ND	ND	ND
T127L	2.4 ± 0.1	<0.07	<0.05	ND	ND	ND	ND	ND
G141Q	26.1 ± 1.0	1.4 ± 0.1	0.43 ± 0.08	<0.08	0.42 ± 0.01	<0.10	<0.08	<0.09
L148R	10.2 ± 0.7	0.45 ± 0.05	ND	ND	<0.12	3.26 ± 0.26	<0.12	ND
H159L	ND	ND	ND	ND	ND	ND	ND	ND
K52N/H159L	ND	ND	ND	ND	ND	ND	ND	ND

^a ND, the binding affinity of CRP with DNA is too weak to be determined.

otides. The apparent association constants increase significantly in the order of cGMP < cCMP < cAMP for all of the examined DNA sequences. It is worthy to note that the affinity of L148R with *lac* or *gal* in the absence of cNMP is stronger than that in the presence of cGMP. This unique phenomenon is only observed in L148R and D53H. The affinity of L148R for DNA is sequence dependent and increases in the order of *lacgal*, *gal*, *lac*, and *gallac* at the same concentration of various cyclic nucleotides, as illustrated by shifting binding isotherms to the lower concentration of L148R. This observation is in good agreement with that of WT CRP (19, 21).

The observed data from the binding study are well described with eq 1, fitting results shown as solid lines in Figure 2. The recovered parameters are summarized in Table 2. Some of the binding isotherms, such as *lacgal* with cCMP or cGMP, lack plateau values at a high CRP concentration due to their weak binding affinity. When an isothermal titration failed to reach a saturation level, the fitting was carried out with a constraint of ΔA , a recovered value from the fit of the data observed in the presence of cAMP, which allowed us to estimate the maximum limits for DNA binding affinity.

Effect of DNA Sequence on DNA Binding by CRP. Figure 3 summarizes the results on the energetics of CRP–DNA formation of CRP mutants as a function of DNA sequence. The effects of the presence of cAMP, cCMP, cGMP, and in the absence of cNMP are shown in Figure 3, panels A, B, C, and D, respectively. Only in the presence of cAMP would measurable DNA binding affinities be observed for all nine CRPs and four DNA sequences. The association constants are listed in Table 2. The binding affinity of CRPs with DNA was sequence dependent, decreasing in the order of *gallac* > *lac* > *gal* > *lacgal*. Some CRP mutants (L148R, D53H, and G141Q) exhibit higher DNA binding affinity than WT CRP for each of the examined DNA sequences. In contrast,

other CRP mutants (K52N, T127L, H159L, and K52N/H159L) exhibit lower DNA binding affinity than WT CRP. The effect of CRP mutation on the binding affinity for *gallac* is not obvious, i.e., the difference in the DNA binding affinity with the examined CRPs for *gallac* lies in a relatively narrow range, compared with that for *lac*, *gal*, and *lacgal*. In the presence of cCMP (Figure 3B), the DNA binding affinity is much weaker than that in the presence of cAMP, even though the affinity is measurable. The pattern of binding energetics as a function of the examined DNA sequences and CRP mutants is basically the same as that in the presence of cAMP, i.e., K52N, T127L, H159L, and K52N/H159L CRP mutants exhibit higher DNA binding affinity than WT CRP whereas the other CRP mutants, except S62F, display lower DNA binding affinity than WT CRP. S62F exhibits significant DNA binding affinity in the presence of cCMP. D53H, G141Q, and L148R CRP exhibit very similar binding affinity for *gallac* in the presence of cCMP. In the presence of cGMP (Figure 3C), the binding affinities of DNAs with CRPs are too weak to be determined precisely among 50% of the binding isotherms. D53H and G141Q display measurable interactions with all four oligomers. Unexpectedly, T127L displays a significant free energy change for the formation of CRP–*gallac* complex. The DNA–CRP formation was also studied in the absence of nucleotide (Figure 3D), since CRP's capability of activating transcription in the absence of cAMP is one of the in vivo characteristics of some CRP mutants (references in ref 15). As expected, only these CRP mutants (L148R, D53H, and G141Q) retain detectable DNA binding affinity in the absence of cyclic nucleotides. The effect of DNA sequence on G141Q–DNA formation in the absence of cyclic nucleotides is the same as that in the presence of cyclic nucleotides, i.e., *gallac* > *lac* > *gal* > *lacgal*. However, the affinities of D53H and L148R with *lac* or *gal* are stronger than those with *gallac* in the absence of cyclic nucleotide, even though the difference is small.

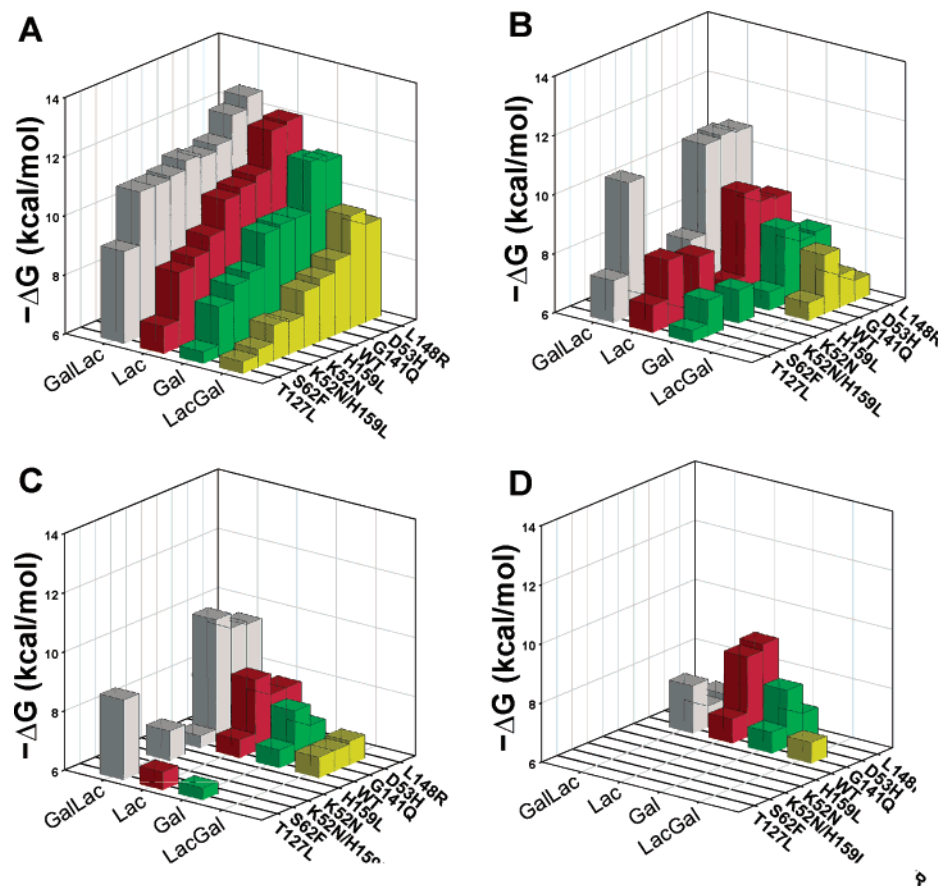


FIGURE 3: Changes in free energy of CRP–DNA complex formation as a function of cNMP. (A) cAMP, (B) cCMP, (C) cGMP, and (D) without cNMP. 200 μ M cyclic nucleotides were used for all the investigated CRP mutants except S62F. With S62F, 1 mM cyclic nucleotide was present in the experiment due to weak cyclic nucleotide affinity. The apparent association constants for the CRP–DNA interaction are summarized in Table 2.

Effect of Cyclic Nucleotides on DNA Binding by CRP.

Figure 4 shows the energetics of complex formation for CRPs in the presence of (A) *gallac*, (B) *lac*, (C) *gal*, and (D) *lacgal* as a function of cyclic nucleotides. For all the four examined DNA sequences, the DNA binding affinity with CRPs decreases in the order of cAMP > cCMP > cGMP \approx no cNMP with some exceptions. The first exception is the binding affinity of *lac* and *gal* with L148R and D53H in the absence of cyclic nucleotide. The binding affinity is stronger than that in the presence of cGMP. Second, the binding affinity of *gallac* with T127L in the presence of cGMP is stronger than that in the presence of cCMP. Quantitatively, the binding affinity of DNA with S62F in the presence of cCMP is only slightly weaker than that in the presence of cAMP, suggesting that S62F has a reduced ability to distinguish among the various cyclic nucleotides.

Effect of Mutation on Protein Surface Properties. During the purification procedures, it became obvious that mutations affect the interaction of the mutant proteins with the chromatographic matrixes as evidenced by a change in the concentration of salt to elute the protein. Figure 5A shows the concentration of KCl at which the peak of mutant CRP was eluted from the BioRex column. Mutations alter the adsorption of these CRP mutants to the column matrix. In most cases, it took a higher KCl concentration to elute the protein. There is no apparent correlation between the salt concentration and the nature of side chains involved in the substitution, e.g., a substitution of H by L in residue 159

leads to a requirement of a lower concentration of KCl for elution; however, a T-to-L mutation in residue 127 has the opposite effect. Figure 5B shows the results of the phenyl sepharose column. Again there is a change in surface properties of CRP as evidenced by the change in salt concentration for elution, without a correlation with the nature of side chains. These results reflect the apparent global nature of the perturbations conferred on the protein by site-specific mutations.

DISCUSSION

A series of studies were conducted to define the effects of specific-site mutations on the ability of CRP to differentiate the various cyclic nucleotides in activating CRP in recognition of the *lac26*-specific DNA sequence. Evidence suggests that cGMP and cCMP are able to replace cAMP to activate some CRP mutants, including D53H, G141Q, L148R, T127L, and S62F, resulting in a measurable binding affinity with the *lac26*-specific DNA sequence (8, 16). Nevertheless, cAMP is the strongest activator for DNA recognition of this particular sequence, namely, PI promoter of the *lac* operon (16). However, CRP is able to recognize more than 100 promoters with different DNA sequences (15, 23, 24), and information regarding both the energetics of CRP–DNA formation and DNA bending geometry is encoded within each of the DNA sequences (18, 19, 21, 22). Although none of the DNA sequences investigated show geometry of wrapping around CRP, the extent of DNA

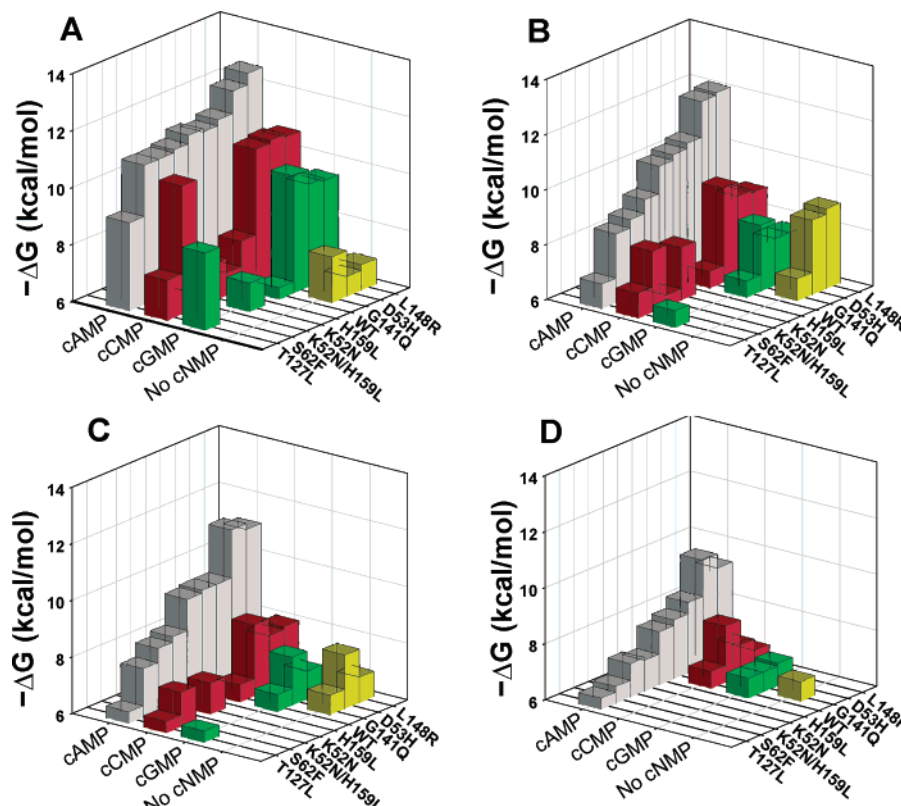


FIGURE 4: Changes in free energy of CRP–DNA complex formation as a function of the DNA sequence (A) *gallac*, (B) *lac*, (C) *gal*, and (D) *lacgal*.

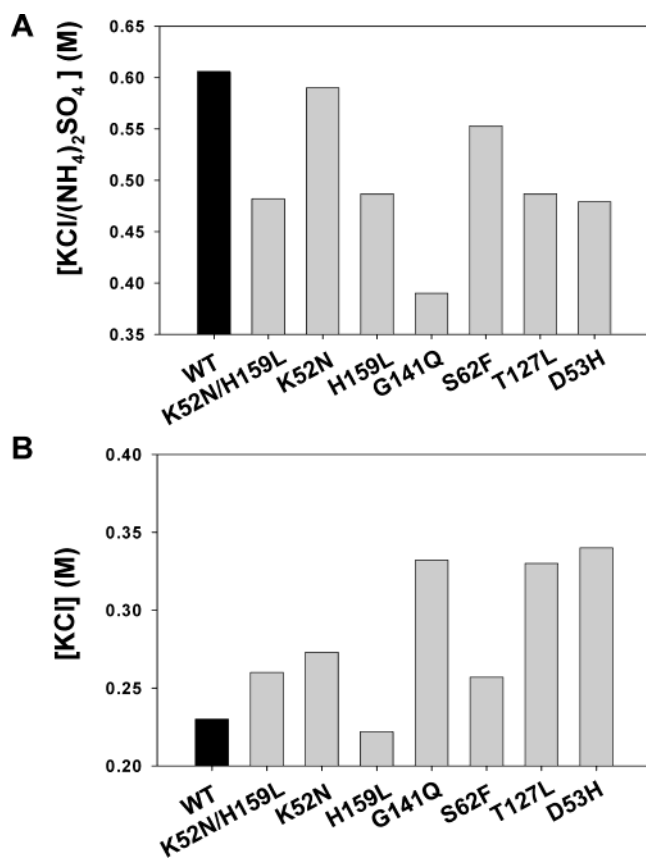


FIGURE 5: Relation between mutations and concentrations of salt for elution. (A) Data for the BioRex column and (B) data for the hydrophobic column.

bending is dependent on the flanking sequence (21). It is conceivable that the distribution of surface charges on CRP

might, in conjunction with the tendency of the DNA sequence in bending, play a role in affecting the chemistry of DNA recognition. In addition, it is conceivable that there are mutation-induced changes in the surface properties of CRP and in conjunction with the cyclic nucleotide bound, be it cAMP or otherwise, may exhibit an alteration in their interactions with DNA sequences because these DNA molecules can assume different geometries upon formation of DNA–CRP complexes. To investigate the interplay between protein sequence and cyclic nucleotides in defining the recognition of DNA sequences, this investigation was initiated.

In an earlier study, a linear correlation was observed between ΔG for subunit assembly of various mutants and the difference in energetics of DNA binding in the presence of cAMP and cCMP, $\Delta\Delta G^3$ (16). These results imply that the binding of cyclic nucleotides to CRP modulates DNA recognition via subunit–subunit and domain–domain communications. In this study, in addition to the *lac* P1 sequence, three DNA sequences were included to explore the effect of the geometry of the DNA–CRP complex. Using the *lac* sequence as the target binding site, there is a linear relation between $\Delta\Delta G$ and the energetics of subunit assembly, with the exception of the data points for S62F and T127L, as shown in Figure 6. The data on the left of the figure indicate that the DNA binding affinities in the presence of cAMP or cCMP are similar and these mutants have low energetics for subunit assembly. However, the data on the right of the figure indicate that these mutants exhibit significantly different

³ The choice of using the difference in binding energetics in the presence of cAMP and cCMP is due to the availability of more reliable data when cCMP replaces cAMP. The DNA binding affinity in the presence of cGMP is usually too weak for reliable quantification.

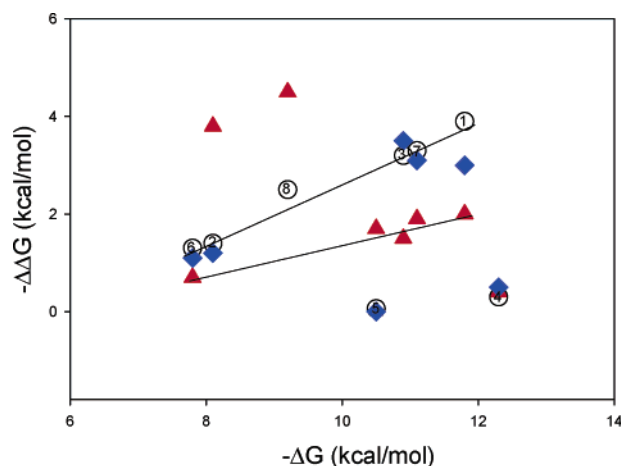


FIGURE 6: Relation between $-\Delta\Delta G$ and $-\Delta G$ of subunit assembly. $\Delta\Delta G$ is the difference in DNA binding in the presence of cAMP and cCMP. The symbols and DNA sequence are (○) *lac*; (◆) *gal*; and (▲) *gallac*. Mutants and symbols are 1, WT; 2, K52N; 3, D53H; 4, S62F; 5, T127L; 6, G141Q; 7, L148R; 8, H159L.

affinity for DNA depending on the specific cyclic nucleotide bound and these mutants have higher energetics for subunit assembly. Such a linear correlation implies that these mutations can modulate the communication between subunits and such interaction(s) is intimately linked to the ability of CRP in differentiating between cAMP and cCMP as allosteric activators. According to the data, S62F and T127L show exceptionally low ability to differentiate cAMP and cCMP. In regards to the energetics of subunit assembly associated with these mutants, they are expected to exhibit a greater ability to differentiate between these two cyclic nucleotides in recognizing specific DNA sequence. The same trend is observed in using *gal* as the target binding site. However, the observation is significantly different when *gallac* is the target site. While the behavior of S62F is still exceptional, the behavior of T127L is part of the linear relation following the trend exhibited by the rest of the mutants and WT CRP, with the exception of K52N and H159L. These two latter two mutants exhibit exceptionally high values for $-\Delta\Delta G$, i.e., greater ability to differentiate between cAMP and cCMP than expected.

Since the results for T127L and S62F do not follow the trend exhibited by the other mutants, it is useful to explore the rationale for these unexpected behavior. A closer examination of the data shows that a replacement of L for T in residue 127 leads to an unexpectedly higher affinity for *lacgal* sequence in the presence of cGMP. The OH group of Thr127 located at the subunit interface (C-helix) forms an H-bond with the adenine 6-amino group of cAMP. Thus, it can be expected that T127L may affect cyclic nucleotide binding, or selectivity of cyclic nucleotide to modulate specific DNA recognition. T127L mutation does not affect the binding constant of cAMP, but reduces the cooperativity between the cAMP binding sites, and suppresses the binding affinity with specific DNA sequences (8). These observations from this laboratory are consistent with that reported by Gorshkova et al., who showed that a T127L mutation decreases the cAMP binding cooperativity without disturbing cAMP binding (25). They also observed that cGMP-liganded T127L maintains about 35% of transcriptional activation of the cAMP-liganded CRP, and the repression of transcription

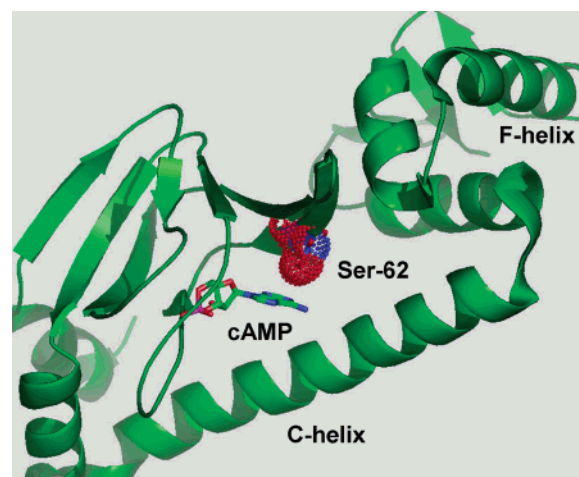


FIGURE 7: Structural presentation showing the spatial relation between serine 62 and bound cAMP.

by T127L was dependent on cGMP rather than cAMP (25, 26). Thus, the authors concluded that a T127L mutation altered the specificity of cyclic nucleotide selectivity by CRP (25). If the interpretation were correct, T127L would be expected to change the order of DNA binding affinity in the presence of different cyclic nucleotides, namely, cGMP should be a more effective ligand to activate CRP for specific DNA sequence recognition. The affinity of T127L CRP with *lacgal* in the presence of different cyclic nucleotides is too weak to resolve this issue (Table 2). The binding affinity of T127L CRP with *gallac*, *lac*, or *gal* in the presence of cGMP is stronger than that in the presence of cCMP. Nevertheless, cAMP is still the strongest activator for the binding of *gallac* DNA sequence. Therefore, a T127L mutation exhibits the ability to switch cyclic nucleotide selectivity, although in the present case it does not alter the selectivity in favor of cGMP over cAMP.

While a replacement of L for T in residue 127 leads to an unexpectedly higher affinity for *lacgal* sequence in the presence of cGMP, the S62F mutant shows a similar behavior but in the presence of cCMP. Although there is no obvious explanation of this unexpected result, one might begin by exploring the interaction between the side chain of residue 62 and the adenine ring of cAMP, as shown in Figure 7. There is most likely van der Waals interaction between the side chain and the adenine ring. It is most likely that a replacement of the serine side chain with a much larger phenyl ring as in phenylalanine an unfavorable steric interaction may ensue. Thus, it may not be surprising that a smaller ring structure, as represented by a cytosine, could fit better as a steric partner for the phenyl group within that space. Although sterically this pair of chemicals may fit better in the spatial confine, specific interactions have been altered leading to a change in the surface chemistry of the protein as shown by the changes in elution condition demonstrated by Figure 5. One might speculate on the reason for the differences in the behavior of T127L and S62F toward the binding of these proteins to the *gallac* sequence. The bending of the *gallac* sequence upon complex formation with CRP is symmetric and at a more severe angle than *lac* (21). Such a conclusion is based on the fluorescence energy transfer data which showed a higher efficiency of transfer between the fluorescence probe, attached to the end of DNA, and

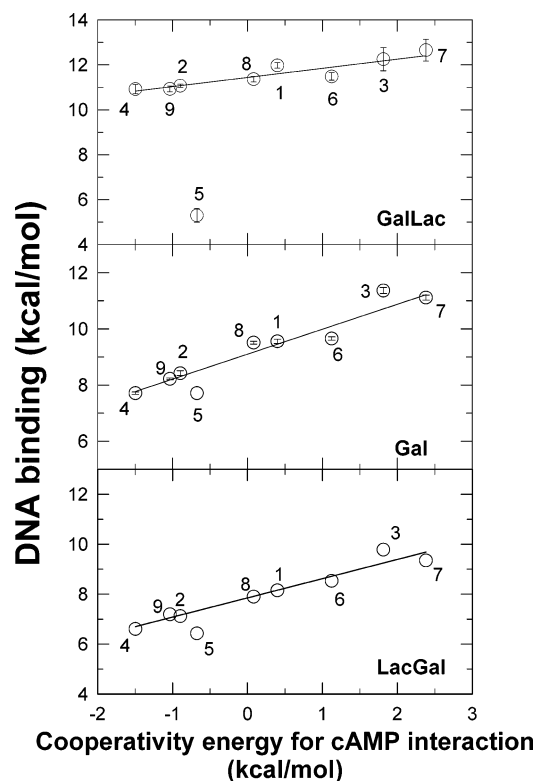


FIGURE 8: Linear correlation between the energetics of ligand binding cooperativity and that of DNA–CRP formation. Mutants and symbols are the same as in Figure 6, and data point 9 is K52N/H159L. The DNA sequence for which the data were determined are *gallac*, *gal*, and *lacgal* shown in the upper, middle, and lower panels, respectively. The energetics of ligand binding cooperativity is adopted from the previous cAMP–CRP binding study (8). Solid lines represent the results of linear regression fit to the observed data with slopes of 0.406 ± 0.133 , 0.888 ± 0.069 , and 0.770 ± 0.075 for *gallac*, *gal*, and *lacgal*, respectively. The value of data point #5, T127L, was excluded during the fitting routine for *gallac* because of its significant deviation from other observations.

tryptophan residues in the N-terminal domain of CRP when the *gallac* sequence is employed (21, 22). This greater bending of DNA observed in the *gallac*–CRP complex may bring parts of the DNA molecule closer to the protein surface. Thus, a change in the appropriate surface area of CRP induced by mutation may lead to the observed results.

The unexpectedly stronger affinity of K52N and H159L mutants for *gallac* might be the result of the same mechanism presented for the S62F mutant, namely, closer interaction between the bent DNA and these sites, as shown by relative location of these residues to the bent DNA which interacts with the F-helices (see Figure 1). These residues are located in the surface of two exposed loops, and they have been reported to be recognized by RNA polymerase to activate transcription (27). Furthermore, the study by Gekko et al. (17) showed that the effects of mutations on these two residues are not additive, i.e., there is communication between these two residues.

Another linear correlation was observed between the energetics of cooperativity of cAMP interaction and *lac26* binding affinity, namely, the mutant with greater favorable energetics of cooperativity exhibits higher affinity for specific DNA sequence (8). The effect of DNA sequence on this relation was explored in this study. Figure 8 shows the results of this correlation for *gallac*, *gal*, and *lacgal*. As was the

case in *lac*, linear correlations are observed in all cases, although the slopes are DNA sequence dependent. The slopes for *gallac* and *lacgal* are less than those of *lac* and *gal*. The cases in which the slopes are shallower indicate that mutations have a lesser effect in altering the ability of CRP in recognizing these specific DNA sequences. It is interesting to note that the DNA sequences in these cases are the synthetic hybrid ones while mutations have a much greater effect on CRP in recognizing the natural *lac* and *gal* sequences. In all these linear relationships for all four DNA sequences, T127L CRP always behaves differently and does not fit into this linear correlation. The greatest deviation from the correlation is observed in the case of *gallac*, as shown in Figure 8, upper panel. Thus, a T-to-L mutation in residue 127 yields a mutant that behaves significantly different and thus becomes a worthwhile subject for more in-depth investigation.

For most of the mutants investigated in this study, there are several aspects to support the conclusion that mutation does not change the mechanism of recognition of ligands by CRP. First, except T127L and S62F, mutations do not change CRP's specificity for DNA recognition modulated by cyclic nucleotides, which decreases in the order of cAMP > cCMP > cGMP, as shown in Figure 3. Although cGMP and cCMP are able to replace cAMP to induce activation in some CRP mutants, cAMP remains the strongest activator for DNA recognition. In the presence of cAMP, mutations maintain the relative affinity for different DNA sequences in the order of *gallac* > *lac* > *gal* > *lacgal*. Second, biphasic dependence of DNA–CRP formation on cAMP concentration is a basic mode of the reaction between WT CRP and DNA (18, 20). This biphasic behavior is a direct consequence of stronger affinity with a specific DNA sequence when CRP is loaded with cAMP to the high affinity binding sites, and weaker affinity when the low affinity cAMP binding sites are also filled. All the mutants presently studied do not affect this characteristic biphasic behavior, as demonstrated by the bell-shaped isotherm of the interaction of CRPs with DNA as a function of cAMP (16).

A study of these mutants leads to an identification of structural elements involved in signal transmission. The observed data of T127L CRP, which deviated from the linear correlation between the energetics of cooperativity of ligand-binding and that of DNA binding (Figure 8), suggest that T127L might not share the same mechanism of subunit–subunit or domain–domain communication in CRP. It has also been reported that specific mutation at Thr127 leads to a cAMP-independent phenotype (28). Residue 62 is located at the end of loop 3, which links between β -sheets 4 and 5 in the cyclic nucleotide-binding β -roll domain. Recently, Chen and Lee have shown that loop 3 is involved in intersubunit communication (29). Thus, the mutation of S62 may affect not only its direct interaction with cyclic nucleotide but also influence the ability of loop 3 to interact with other structural elements involved in signal transmission such as the hinge region.

We previously demonstrated that loading of the first cAMP to the high affinity cAMP binding sites significantly increases the affinity of the second cAMP to CRP, i.e., positive cooperativity. Mutations of residues 62 and 127 lead to changes in the extent of cooperativity. This finding suggests that the two subunits of CRP communicate to each other

through cAMP binding. Several lines of evidence support that the transmission of information may be through the interfacial domain–domain and subunit–subunit interactions (30–33). In the absence of cAMP, mutant G141Q is susceptible to proteolysis at the subunit interface. Such susceptibility is an established property of WT CRP only in the presence of cAMP. These observations imply that the G141Q mutant exhibits a subunit alignment similar to that of the activated form of CRP (32). In addition, chemical modification of the C-terminal DNA binding domain of the G141Q mutant shows that the DNA binding domain responds quantitatively to the binding of cyclic nucleotide to the N-terminal domain (32). This result indicates that an additional process leads to communication between domains. Furthermore, correlation between a weak affinity for subunit assembly and relaxation of cyclic nucleotide selectivity in mutants G141Q and S128A/G141Q suggests that intersubunit interaction is important for cyclic nucleotide discrimination in CRP (30). Recently, this correlation has been confirmed with other CRP mutants (K52N, D53H, S62F, T127L, L148R, H159L, and K52N/H159L), even though the sites of mutations are located in different structural elements in CRP (16).

The mechanism that is responsible for the alteration in signal transmission elicited by site-specific mutation probably involves protein dynamics. We previously reported that large positive entropy changes are the major driving force for the cooperativity binding of the two cAMP molecules to CRP, implying that the protein dynamics might play a key role in subunit–subunit communication (8). Dong et al. reported that relative protein dynamics of CRP is affected by cAMP binding based on a hydrogen–deuterium exchange study (34). Upon cAMP binding, the conformation of the DNA-binding domain changes to a more flexible structure, whereas the larger cAMP-binding domain shifts to a more rigid or compact structure. Unlike cAMP, the binding of cGMP to CRP exhibits no cooperativity between the two high affinity cyclic nucleotide sites. Interestingly, the interaction between cGMP and CRP is enthalpy driven without significant increases in entropy change. This observation is consistent with the hydrogen–deuterium exchange data, which indicate that cGMP is unable to induce a similar relaxation effect as cAMP on the DNA binding domain. Thus, cAMP-induced protein dynamics change might play a key role in subunit–subunit and domain–domain communications in CRP. Why is cGMP unable to induce similar change in protein dynamics, even though cGMP binds to CRP with the same affinity as cAMP? Recently, protein compressibility has provided new insights into the relationship between protein dynamics and protein function (17). Cavities, which are easily compressed by pressure in a macromolecule, offer protein the space for internal motions or the flexibility to respond to thermal or mechanical forces. Thus, protein compressibility is able to reflect the protein flexibility or volume fluctuation. CRP mutants significantly influence DNA binding affinity without changing the protein secondary structure. In addition, the currently favored model for the allosteric activation of CRP by cAMP involves subtle changes in the CRP structure, such as a rolling change in the relative position of the C-helices leading to a rigid body motion of the domain with respect to each other (15). This rolling event of the C-helices inside CRP likely influences amino acid repacking, thus

changing the protein cavity. cAMP induces protein conformational change through protein dynamics. CRP mutants (D53H, L148R, and G141Q) are able to recognize specific DNA sequences in the presence of cGMP, cCMP, and in the absence of cyclic nucleotide, as shown in Figure 3. Thus, it is safe to conclude that mutations in CRP are able to change protein dynamics. In other words, mutation changes protein compressibility, which in turn confers CRP with a capability to recognize specific DNA sequences without protein conformational changes induced by cAMP. Experiments to test this hypothesis are currently in progress.

In summary, this study provides evidence for the interplay between protein sequence and cyclic nucleotide in conferring the ability of macromolecular recognition. Qualitatively, most of the CRP mutants studied maintain the selectivity of cyclic nucleotide to modulate DNA recognition. The linear correlation between the energetics of cooperativity of cAMP binding and that of DNA binding suggests that mutation does not switch the communication mechanism between the cAMP binding domain and DNA binding domain (Figure 8). Since protein dynamics is involved in the long-range communication in CRP, perturbation of protein microstate distributions most likely contributes to the molecular events, such as cAMP binding, signal transmission, and DNA recognition.

ACKNOWLEDGMENT

The assistance of Drs. Jiang Qin and Bi-Hung Peng in the preparation of the manuscript is appreciated.

REFERENCES

1. Busby, S., and Ebright, R. H. (1999) Transcription Activation by Catabolite Activator Protein (CAP), *J. Mol. Biol.* 293, 199–213.
2. Harman, J. G. (2001) Allosteric regulation of the cAMP receptor protein, *Biochim. Biophys. Acta* 1547, 1–17.
3. de Crombrughe, B., Busby, S., and Buc, H. (1984) Cyclic AMP Receptor Protein: Role in Transcription Activation, *Science* 224, 831–838.
4. Kolb, A., Busby, S., Buc, H., Garges, S., and Adhya, S. (1993) Transcriptional regulation by cAMP and its receptor protein, *Annu. Rev. Biochem.* 62, 749–796.
5. Aiba, H., Fujimoto, S., and Ozaki, N. (1982) Molecular cloning and nucleotide sequencing of the gene for *E. coli* cAMP receptor protein, *Nucleic Acids Res.* 10, 1345–1361.
6. Cossart, P., and Gicquel-Sanzey, B. (1982) Cloning and sequence of the *crp* gene of *Escherichia coli* K 12. *Nucleic Acids Res.* 10, 1363–1378.
7. Weber, I. T., and Steitz, T. A. (1987) Structure of a Complex of Catabolite Gene Activator Protein and Cyclic AMP Refined at 2.5 Å Resolution, *J. Mol. Biol.* 198, 311–326.
8. Lin, S.-H., and Lee, J. C. (2002) Communications between the High-Affinity Cyclic Nucleotide Binding Sites in *E. coli* Cyclic AMP Receptor Protein: Effect of Single Site Mutations, *Biochemistry* 41, 11857–11867.
9. Takahashi, M., Blazy, B., and Baudras, A. (1980) An Equilibrium Study of the Cooperative Binding of Adenosine Cyclic 3',5'-Monophosphate and Guanosine Cyclic 3',5'-Monophosphate to the Adenosine Cyclic 3',5'-Monophosphate Receptor Protein from *Escherichia coli*, *Biochemistry* 19, 5124–5130.
10. Takahashi, M., Blazy, B., Baudras, A., and Hillen, W. (1989) Ligand-modulated Binding of a Gene Regulatory Protein to DNA. Quantitative Analysis of Cyclic-AMP Induced Binding of CRP from *Escherichia coli* to Non-specific and Specific DNA Targets, *J. Mol. Biol.* 207, 783–796.
11. McKay, D. B., Weber, I. T., and Steitz, T. A. (1982) Structure of catabolite gene activator protein at 2.9-Å resolution. Incorporation of amino acid sequence and interactions with cyclic AMP, *J. Biol. Chem.* 257, 9518–9524.

12. Sanders, R. and McGeoch, D. (1973) A Mutant Transcription Factor That is Activated by 3':5'-Cyclic Guanosine Monophosphate, *Proc. Natl. Acad. Sci. U.S.A.* 70, 1017–1021.
13. Dessein, A., Schwartz, M., and Ullmann, A. (1978) Catabolite Repression in *Escherichia coli* Mutants Lacking Cyclic AMP*, *Mol. Gen. Genet.* 162, 83–87.
14. Garges, S., and Adhya, S. (1985) Sites of Allosteric Shift in the Structure of the Cyclic AMP Receptor Protein, *Cell* 41, 745–751.
15. Passner, J. M., Schultz, S. C., and Steitz, T. A. (2000) Modeling the cAMP-induced Allosteric Transition Using the Crystal Structure of CAP-cAMP at 2.1 Å Resolution, *J. Mol. Biol.* 304, 847–859.
16. Lin, S.-H., Kovac, L., Chin, A. J., Chin, C. C. Q., and Lee, J. C. (2002) Ability of *E. coli* Cyclic AMP Receptor Protein To Differentiate Cyclic Nucleotides: Effects of Single Site Mutations, *Biochemistry* 41, 2946–2955.
17. Gekko, K., Obu, N., Li, J., and Lee, J. C. (2004) A Linear Correlation Between the Energetics of Allosteric Communication and Protein Flexibility in *E. coli* Cyclic AMP Receptor Protein Revealed by Mutation-Induced Changes in Compressibility and Amide Hydrogen/Deuterium Exchange, *Biochemistry* 43, 3844–3852.
18. Pyles, E. A., and Lee, J. C. (1996) Mode of Selectivity in Cyclic AMP Receptor Protein-Dependent Promoters in *Escherichia coli*, *Biochemistry* 35, 1162–1172.
19. Pyles, E. A., Chin, A. J., and Lee, J. C. (1998) *Escherichia coli* cAMP Receptor Protein-DNA Complexes: I. Energetic Contributions of Half Sites and Flanking Sequences in DNA Recognition, *Biochemistry* 37, 5194–5200.
20. Heyduk, T., and Lee, J. C. (1990) Application of fluorescence energy transfer and polarization to monitor *Escherichia coli* cAMP receptor protein and *lac* promoter interaction, *Proc. Natl. Acad. Sci. U.S.A.* 87, 1744–1748.
21. Pyles, E. A., and Lee, J. C. (1998) *Escherichia coli* cAMP Receptor Protein-DNA Complexes: II. Structural Asymmetry of DNA Bending, *Biochemistry* 37, 5201–5210.
22. Lin, S.-H., and Lee, J. C. (2003) Determinants of DNA Bending in the DNA-cyclic AMP Receptor Complexes in *E. coli*, *Biochemistry* 42, 4809–4818.
23. Ebright, R. H. (1993) Transcription activation at Class I CAP-dependent promoters, *Mol. Microbiol.* 8, 797–802.
24. Zhou, Y., Merkel, T. J., and Ebright, R. H. (1994) Characterization of the Activating Region of *Escherichia coli* Catabolite Gene Activator Protein (CAP) II. Role at Class I and Class II CAP-dependent Promoters, *J. Mol. Biol.* 243, 603–610.
25. Gorshkova, I. I., Moore, J. L., McKenney, K. H., and Schwarz, F. P. (1995) Thermodynamics of Cyclic Nucleotide Binding to the cAMP Receptor Protein and Its T127L Mutant, *J. Biol. Chem.* 270, 21679–21683.
26. Moore, J. L. (1993) Thesis/Dissertation, The Mechanism of Activation of the *Escherichia coli* Cyclic AMP Receptor Protein by Cyclic AMP. The George Washington University, Washington, DC.
27. Williams, R., Bell, A., Sims, G., and Busby, S. (1991) The Role of Two Surface Exposed Loops in Transcription Activation by the *Escherichia coli* CRP and FNR Proteins, *Nucleic Acids Res.* 19, 6705–6712.
28. Harman, J. G., McKenney, K. H., and Peterkofsky, A. (1986) Structure-Function Analysis of Three cAMP-independent Forms of the cAMP Receptor Protein, *J. Biol. Chem.* 261, 16332–16339.
29. Chen, R., and Lee, J. C. (2003) Functional Roles of Loops 3 and 4 in the Cyclic Nucleotide Binding Domain of Cyclic AMP Receptor Protein from *Escherichia coli*, *J. Biol. Chem.* 278, 13235–13243.
30. Cheng, X. D., and Lee, J. C. (1998) Interactive and Dominant Effects of Residues 128 and 141 on Cyclic Nucleotide and DNA Bindings in *E. coli* cAMP Receptor Protein, *J. Biol. Chem.* 273, 705–712.
31. Cheng, X. D., Kovac, L., and Lee, J. C. (1995) Probing the Mechanism of CRP Activation by Site-Directed Mutagenesis: The Role of Serine 128 in the Allosteric Pathway of cAMP Receptor Protein Activation, *Biochemistry* 34, 10816–10826.
32. Cheng, X. D., and Lee, J. C. (1998) Differential Perturbation of Intersubunit and Interdomain Communications by Glycine 141 Mutation in *E. coli* CRP, *Biochemistry* 37, 51–60.
33. Shi, Y., Wang, S., Krueger, S., and Schwarz, F. P. (1999) Effect of mutations at the monomer-monomer interface of cAMP receptor protein on specific DNA binding, *J. Biol. Chem.* 274, 6946–6956.
34. Dong, A., Malecki, J., Lee, L., Carpenter, J. F., and Lee, J. C. (2002) Ligand-Induced Conformational and Structural Dynamics Changes in *Escherichia coli* Cyclic AMP Receptor Protein, *Biochemistry* 41, 6660–6667.
35. DeLano, W. L. (2002) The PyMOL Molecular Graphics System, DeLano Scientific, San Carlos, CA.

BI0499359