Entropic Nature of the Interaction between Promoter Bound CRP Mutants and RNA Polymerase†

Susan Krueger,§ Susan Gregurick,§ Ying Shi,‖ Shenglun Wang,‖ Brian D. Wladkowski,‡ and Frederick P. Schwarz*‖

National Institute of Science and Technology Center for Neutron Research, Gaithersburg, Maryland 20899, Department of Chemistry and Biochemistry, University of Maryland, Baltimore County, 1000 Hilltop Circle, Baltimore, Maryland 21250, Center for Advanced Research in Biotechnology/National Institute of Science and Technology, 9600 Gudelsky Drive, Rockville, Maryland 20850, and Department of Chemistry, Western Maryland College, 2 College Hill, Westminster, Maryland 21157

Received August 29, 2002; Revised Manuscript Received December 20, 2002

ABSTRACT: The interaction between CRP, T127L, S128A, and CRP* and RNA polymerase bound to a 104 bp synthetic promoter were determined by ITC at 298 K and ranges from a ∆G° = 1.4 ± 0.8 kJ mol⁻¹ (cAMP-ligated S128A) to 4.5 ± 0.3 kJ mol⁻¹ (cAMP-ligated double mutant CRP*) with endothermicities that range from 4 ± 3 kJ mol⁻¹ (cAMP-ligated CRP) to 47 ± 8 kJ mol⁻¹ (cGMP-ligated T127L). The interaction is, thus, entropically driven, exhibits enthalpy−entropy compensation, and increases the binding affinity of the RNA polymerase to the promoter by factors ranging from 1.7 ± 0.1 (cAMP-ligated S128A) to 6.1 ± 0.1 (cAMP-ligated CRP*). Although the binding affinities to the promoter alone, except for cAMP-ligated S128A, are the same as to a shorter 40 bp duplex containing the same CRP consensus binding site sequence (conDNA), the binding enthalpies of CRP/mutant to the promoter are lower by factors of 2−3x than the corresponding binding enthalpies to conDNA. Small angle neutron scattering measurements on the DNA−CRP/mutant complexes in D₂O/H₂O solutions exhibit an increase in the Rg of the CRP/mutant component from 22 to 27−31 Å that can be attributed to a conformational change in the N-terminal domain of CRP. The Rg = 27 Å for the bound conDNA can be attributed to a slight unwinding of the DNA in solution that would also enhance the activation of transcription. The Rg = 53 ± 3 Å for the bound promoter is attributed to bending of the promoter in solution that can be responsible for the lower CRP/mutant promoter binding endothermicities.

The transcription of enzymes that metabolize carbohydrates in Escherichia coli is enhanced by the binding of cAMP receptor protein (CRP)† to a site on the promoter that is centered at −61.5 bp upstream from the transcription start point and adjacent to the RNA polymerase binding site. This enhancement can be attributed to (i) an increase in the binding affinity of RNA polymerase to the promoter by CRP (1, 2), (ii) bending of the promoter by CRP resulting in more contacts between the RNA polymerase and the promoter (3), and/or (iii) an increase in the conversion rate of the RNA polymerase promoter complex from a closed form to the transscriptionally active open form by CRP (2). CRP is a 47 000 g mol⁻¹ dimer with the C-terminal domain of each subunit binding to the promoter site upon activation of the protein by binding of cAMP to the N-terminal domain. The cAMP-ligated CRP binding affinity to the promoter sequence can be altered by mutations at the CRP subunit interface such as the T127L mutation (T127L), the S128A mutation (S128A), double mutations consisting of T127L and S128A (CRP*), and by cGMP instead of cAMP ligation to the T127L and CRP* mutants (4). Although the binding affinities of the mutants to a 104 bp synthetic promoter containing the CRP consensus binding sequence (syncon promoter) range from 1.1 ± 0.1 × 10⁶ M⁻¹ (cAMP-ligated CRP*) to 6.6 ± 1.1 × 10⁶ M⁻¹ (cAMP-ligated CRP) (4), the binding affinities of the CRP mutants to the RNA polymerase−promoter complex vary only from 0.80 ± 0.20 to 2.2 ± 1.2 × 10⁶ M⁻¹ (5). This is in closer agreement with the observed variation from 11.7 ± 0.5 to 13 ± 1 in the enhancement of in vitro transcription by the CRP mutants (5). A weaker binding affinity of a CRP mutant to the promoter is apparently compensated by a stronger binding interaction between the CRP mutant and the bound RNA polymerase. Undoubtedly, CRP increases the binding affinity of RNA polymerase to the promoter by providing additional indirect interactions between the RNA polymerase and through the CRP to the promoter. However, the nature of the interaction between the CRP mutant and the bound RNA
polymerase still remains to be determined.

The bending of the promoter by CRP to increase interactions between the RNA polymerase and the promoter was inferred from early X-ray crystal structures of the cAMP-ligated CRP—short DNA complexes that showed the bound DNA to be bent by 90° (6). In addition to bending of the promoter, conformational changes in the CRP may also occur upon binding to the promoter to facilitate interactions with the adjacent bound RNA polymerase. This is based on the premise that the interaction between CRP and bound RNA polymerase is analogous to the interactions between transcription factors and an RNA polymerase in the preinitiation transcription complex in eukaryotic transcription. Proteolysis probes (7), circular dichroism measurements (7–9), fluorescence measurements (8), and NMR measurements (9) show that these transcription factors undergo local and global folding processes in their DNA-binding domains upon binding to the promoter. These conformational changes, such as in the transcription factor Ets-1 (7), can provide additional contact points with other proteins in preinitiation complex to activate transcription in eukaryotes. Protein footprinting analysis of the cAMP-ligated CRP complex in solution do indeed imply that CRP undergoes an additional conformational change in the N-terminal domains upon binding to ≥30 bp DNA duplexes containing the CRP consensus binding-site sequence (10). However, since this conformational difference in the N-terminal domain is not observed in a comparison between the crystal structures of the DNA-bound and unbound cAMP-ligated CRP complexes (6, 11–13), it may be minimized by crystal-packing forces in the crystallization of the protein complexes. Furthermore, the 30 bp DNA in the crystal structure consisted of two overlapping DNA fragments in contrast to the >30 bp DNA duplexes used in the solution studies (4, 10). It is, thus, necessary to determine if bending of the promoter analogous to the bent DNA in the X-ray crystal structure of cAMP-ligated CRP—DNA complexes as well as conformational changes in the CRP upon promoter binding does occur in solution to provide a more complete understanding of how the activation of transcription is enhanced by CRP.

In this investigation, the enthalpy and entropy changes that determine the binding affinities of the CRP mutants to the syncon promoter (Figure 1) were determined by isothermal titration calorimetry (ITC) measurements and compared to the corresponding changes that determine the binding affinities of the CRP mutants to an RNA polymerase—promoter complex. The promoter consists of a 104 bp section from −82 to 22 of the lacUV5 promoter that contains the RNA polymerase binding site and was modified to a consensus CRP binding sequence centered at −61 bp (syncon promoter). Differences between these thermodynamic parameters are then attributed to the interaction between the bound CRP mutant and the RNA polymerase. Since this interaction between the bound CRP mutant and the RNA polymerase may be facilitated by a conformational change in CRP upon binding to the promoter, small angle neutron scattering (SANS) measurements were employed to directly determine if any conformational change occurs in CRP upon binding to the promoter site, as mimicked by a shorter 40 bp CRP consensus duplex (conDNA) and a 40 bp duplex that contained the CRP binding sequence from the lacUV5 promoter (lacDNA) (Figure 1). An increase in the radius of gyration of the CRP mutant from 22 to 30 Å in the promoter site-bound state is indeed observed in SANS measurements on the CRP/mutant—conDNA complex in solution. The SANS data are then compared to simulated SANS data from a distorted energy-minimized X-ray crystal structure of the protein component of the cAMP-ligated CRP—DNA complex to determine the nature of this conformational change. The DNA component was modeled as a bent rigid rod in these calculations and exhibited partial unwinding, characteristic of the active open form of the RNA polymerase—promoter transcription complex. Furthermore, SANS data on the 104 syncon promoter bound to the CRP/mutants show that it is in the bent conformation, which would facilitate contacts between the promoter and the RNA polymerase. Although there is general agreement between the CRP mutant binding affinities to the 104 bp syncon promoter and the shorter 40 bp conDNA duplex, the binding enthalpy change to the shorter sequence is larger than that to the longer sequence. This difference is explored in terms of the long-range electrostatic forces contributing to the binding enthalpy with the longer promoter.

**EXPERIMENTAL PROCEDURES**

**Materials.** The CRP and mutants were expressed, purified, and assayed according to the procedures described previously (4). The concentration of the CRP/mutants was determined

\[
\text{conDNA} = 5'\text{GCAACGCAATAAATGTGATCTGACATCAATTTAGCACC}3' \\
3'\text{GGTGGCTTTATTTACATGTCTGTAATAACATTCCGGG}5',
\]

\[
\text{lacDNA} = 5'\text{GCAACGCAATTATGTTACGTCCTACCTACATAGGACC}3' \\
3'\text{GGTGGCTTTATTTACATGTCTGTAATAACATTCCGGG}5',
\]

**Figure 1:** Sequences for the conDNA, lacDNA, and syncon promoter. In the syncon promoter, the CRP consensus binding site is in bold letters, and P1 is the transcription start point.
from UV measurements at 280 nm using an extinction coefficient of $3.5 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$ (14). The RNA polymerase holoenzyme was prepared and purified according to the procedure described by Lowe et al. (15), and an extinction coefficient of $2.77 \times 10^2 \text{ cm}^{-1} \text{ M}^{-1}$ at 280 nm was used to determine the concentration of the RNA polymerase (16). The complementary 40 bp conDNA and lacDNA strands were obtained commercially at a purity level >90% as assayed by capillary electrophoresis and an analytical ion exchange column. Gel electrophoresis of the conDNA and lacDNA strands revealed essentially one intense band at 12,000 g, the molecular mass of the strand. The DNA duplexes were prepared by annealing complementary single DNA strands as described previously (4), and their concentrations were determined from UV absorption measurements using an extinction coefficient of $5.33 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$ (17).

The 104 bp syncon promoter was from -82 to +22 of the lacUV5 promoter, so that it contained the CRP and RNA polymerase binding site sequences shown in Figure 1. Each complementary strand of the 104 bp promoter was synthesized on the $\mu$mol level on a DNA synthesizer and purified by gel electrophoresis. The complementary sequences were annealed by heating equal amounts of each strand in 10 mM Tris/HCl buffer containing 1 mM MgCl$_2$ and 0.5 M NaCl at pH = 7.4 up to 95 °C followed by slow cooling to room temperature. The concentration of the 104 bp syncon promoter was determined from OD measurements at 260 nm using an extinction coefficient of $1.3 \times 10^4 \text{ cm}^{-1} \text{ M}^{-1}$ based on an OD of 1 for a 50 ng $\mu$L$^{-1}$ solution (18) and a molecular mass of 65,000 g mol$^{-1}$.

The phosphate buffer solution was 50 mM K$_2$HPO$_4$ and contained 0.2 mM DTT, 0.2 mM EDTA, and 0.15 mM KCl at pH = 7.0. For the cAMP-ligated and cGMP-ligated CRP/mutant titrations, the buffer contained 1 mM of the cyclic nucleotide monophosphate to ensure that the CRP/mutant was all complexed with the ligand. The potassium phosphate salts, NaCl, KCl, Tris, MgCl$_2$, sodium salts of cAMP and cGMP, mercaptoethanol, polyacrylamide, bromophenol blue, and urea were reagent grade from Sigma Chemical Co. The DTT was ultrapure brand from Gibcobrl. The sodium salt of EDTA was from Serva Co. The HCl and glycerol were reagent grade from Mallinckrodt.

**ITC Measurements.** The binding affinity, enthalpy, and entropy of the cNMP-ligated CRP/mutants to the 104 bp syncon promoter and the promoter–RNA polymerase complex were determined from ITC measurements, using a Microcal, Inc. VP Titration Calorimeter as described previously (4, 5). The sample vessel contained either the RNA polymerase, the RNA polymerase–104 bp syncon promoter complex, or the promoter alone in the phosphate buffer, while the reference vessel contained just the buffer solution. First, 2–4 $\mu$L aliquots of the 0.03–0.1 mM promoter solution were titrated 3–4 min apart into the 1–3 $\mu$L RNA polymerase sample solution until the binding was saturated as evident by the lack of a heat exchange signal. After adding cNMP to the promoter–RNA polymerase solution, 10 $\mu$L aliquots of a 0.03–0.06 mM cNMP-ligated CRP/mutant solution were titrated into the promoter–RNA polymerase complex solution. In a separate titration, the cNMP-ligated CRP/mutant solution was titrated into the sample vessel containing just a 1–3 $\mu$L promoter solution. For each of the titrations, the additions were continued for 2–3 times past saturation so that a heat of dilution of the titrant could be determined from these additional peak areas. For the promoter into RNA polymerase titrations, these extra additions amounted to about a 7% excess of the promoter to RNA polymerase concentration. The heats of dilution obtained from titrating the ligand just into the buffer solution were then subtracted from the heats obtained during the titration prior to analysis of the data.

A nonlinear, least-squares minimization software program from Microcal, Inc., Origin 5.0 (19), was used to fit the incremental heat of the $i$th titration ($\Delta Q(i)$) of the total heat, $Q_i$, to the total titrant concentration, $X$, according to the following equations:

$$Q_i = nC_i \Delta H_i \cdot V(1 + X/nC_i + 1/nK_i C_i) - [(1 + X/nC_i + 1/nK_i C_i)^2 - 4X/nC_i]^{1/2}/2 \quad (1a)$$

$$\Delta Q(i) = Q(i) + dV_{li}/2V \left\{ Q(i) + Q(i - 1) \right\} - Q(i - 1) \quad (1b)$$

where $C_i$ is the total RNA polymerase–promoter complex or promoter concentration in the sample vessel, $V$ is the volume of the sample vessel, and $n$ is the stoichiometry of the binding reaction, to yield values of $K_i$, $\Delta H_i$, and $n$.

**SANS Measurements.** The concentrations of the SANS solutions ranged from 0.01 to 0.09 mM of cAMP-ligated CRP, cAMP-ligated T127L, or CRP* and were complexed with the conDNA, lacDNA, or syncon promoter in the phosphate buffer. To confirm that the complexes had the 1:1 stoichiometry for DNA binding, the complexes were formed in an ITC by titrating the DNA into the protein solution until the binding reached saturation. For the cAMP-ligated complexes, the buffer contained 1 mM cAMP so that the protein in the SANS solution was saturated with bound cAMP. For the contrast variation technique, the solutions were dialyzed in the appropriate D$_2$O/H$_2$O buffer.

SANS measurements were performed on the CHRSN 30 m SANS instrument at the NIST Center for Neutron Research in Gaithersburg, MD (20). The neutron wavelength, $\lambda$, was 5 Å, with a wavelength spread, $\Delta \lambda/\lambda$, of 0.15. The neutron scattering intensity was corrected for background scattering from the buffers and incoherent scattering from hydrogen in the samples. Data were placed on an absolute scale by normalizing the scattering intensity to the incident beam flux. Finally, the data were radially averaged to produce scattering intensity, $I(Q)$, versus $Q$ curves, where $Q = 4\pi \sin(\theta)/\lambda$, and $2\theta$ is the scattering angle. The Guinier approximation, $I(Q) = I(0) \exp(-Q^2 R_g^2/2)$, was used on the low-$Q$ portions of the data to obtain initial values for the radius of gyration, $R_g$, and the forward scattering intensity, $I(0)$, of the samples. This analysis is valid in the region $QR_g \sim 1$. In the cases where higher $Q$ data were available, the GNOM program (21) was used to determine the distance distribution function, $P(r)$, the radius of gyration, and the forward scattering intensity. The advantage of this approach is that all of the data are used, rather than a limited data set.

---

2 Certain commercial materials, instruments, and equipment are identified in this manuscript to specify the experimental procedure as completely as possible. In no case does such identification imply a recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that the materials, instruments, or equipment identified is necessarily the best available for the purpose.
Entropic Interaction between CRP and RNA Polymerase

at small $Q$ values, leading to more accurate determinations of $R_g$ and $I(0)$ that are less influenced by possible aggregation effects.

The scattering intensities from the CRP–DNA complexes were decomposed into the scattering from their components, $I_{CRP}(Q)$ (or $I_{CRP*}(Q)$) and $I_{DNA}(Q)$ using the following equation:

$$I(Q) = \Delta \rho_{CRP}^2 I_{CRP}(Q) + \Delta \rho_{CRP} \Delta \rho_{DNA} I_{CRP+DNA}(Q) + \Delta \rho_{DNA}^2 I_{DNA}(Q)$$

(2)

where $\Delta \rho = (\rho - \rho_0)$ is the contrast, or the difference between the scattering length density of the molecule ($\rho$) and the solvent ($\rho_0$). The cross-term, $I_{CRP+DNA}(Q)$, represents the interference function between the CRP and the DNA components. The known quantities in eq 2 are $\Delta \rho_{CRP}$ and $\Delta \rho_{DNA}$, and the unknowns are $I_{CRP}(Q)$, $I_{DNA}(Q)$, and $I_{CRP+DNA}(Q)$. Since measurements were made at three different contrasts, or $D_2O/H_2O$ buffer conditions, there is sufficient information to solve for the three unknown component intensities from the set of simultaneous equations for $I(Q)$ at each contrast.

Monte Carlo Simulations. Monte Carlo simulations of the SANS intensity from the CRP component of the complex were performed by representing each amino acid as a homogeneous sphere and calculating the size and neutron scattering strength of each sphere from the known volume and chemical composition of the amino acid. Once a sphere has been drawn around each amino acid, centered about the coordinates of its $\alpha$-carbon atom, a box is drawn around the entire molecule with a region of bound water up to 5 Å thick. Points are generated at random within the box. The distance distribution function, $P(r)$, for the total volume is calculated in real space by making a histogram representing the distances between all possible pairs of points, weighted according to the product of the neutron scattering strengths at each point (22). A distorted energy-minimized X-ray crystal structure of a cAMP-ligated CRP–DNA complex was used for simulating the SANS intensities. The starting structure for the minimizations was the protein portion of the Protein Data Bank file 1BER by Parkinson et al. (12) with a distortion in the structure that consisted of relaxing the torsional angle between the C and the B $\alpha$-helicis. The addition of cAMP to the structure as described by Passner and Steitz (13) made very little difference in the simulations. Minimization of the starting structure employed the AMBER force field parameters and used the method of steepest descent.

The DNA structure was represented as a bent 10 Å diameter rod of uniform neutron scattering length density. The DNA portion of the Parkinson et al. (12) crystal structure was used as a guide to determine the position and structure of the bends in the DNA. The total length of the rod is dependent on the number of base pairs in the DNA, where each base pair is represented by a segment 3.4 Å in length. Thus, for the 40 bp lacDNA and conDNA the total rod length was 136 Å, and for the 104 bp syncon promoter the total rod length was 354 Å.

RESULTS

ITC Measurements. The results of ITC titrations of cAMP-ligated CRP solution into the 104 bp syncon promoter solution without bound RNA polymerase are shown in Figure 2 and with bound RNA polymerase in Figure 3. The results of these binding reactions are summarized in Table 1 for cAMP-ligated CRP, T127L, S128A, and CRP* and cGMP-ligated T127L. In the absence of bound RNA polymerase, the binding affinities to the syncon promoter range from $1.6 \pm 0.2 \times 10^6$ (cAMP-ligated CRP) to $8.4 \pm 0.5 \times 10^6$ M$^{-1}$ (cAMP-ligated S128A) and are consistently slightly higher but still, with the exception of the cAMP-ligated S128A mutant, within experimental error of the corresponding binding affinities to the 40 bp conDNA duplex containing just the 26 bp consensus sequence. It is not clear as to why the cAMP-ligated S128A binding affinity to the syncon promoter is 4–6× larger than to the conDNA duplex. The ligated CRP/mutant binding affinities to the promoter with RNA polymerase bound to the promoter are all higher, ranging from $9.8 \pm 0.1 \times 10^6$ (cAMP-ligated CRP*) to 16.1 $\pm 4.6 \times 10^6$ M$^{-1}$ (cAMP-ligated CRP), as observed earlier (5) and are attributed to additional interactions between the CRP and the RNA polymerase (5). Differences in the binding free energy, enthalpy, and entropy between the ligated CRP/mutant binding to the promoter with and without bound RNA polymerase are presented in Table 2. These differences are attributed to the interaction thermodynamic parameters between the bound RNA polymerase and the ligated CRP/mutant.
polymerase to the CRP to yield the following net reaction for binding of the RNA polymerase to the promoter.

$$\text{polymerase} + \text{CRP} \rightarrow \text{RNA polymerase}$$

To determine how much the interaction between CRP/mutant and the RNA polymerase bound to the promoter is compensated by a strong interaction on the longer promoter can account for this difference through additional exothermic contributions to the binding enthalpy. ITC titrations were performed on CRP and several mutants binding to the promoter at a higher ionic strength of 0.5 M KCl that would shield these long-range electrostatic interactions. The results at this higher ionic strength are compared to similar binding results to the conDNA at 0.5 M KCl from ref 4 in Table 3. Comparison of the results in Tables 1 and 3 show that the conDNA and the syncon promoter binding free energy changes are all lower at the higher ionic strength. Although values for $\Delta G^o$ are in agreement for binding of conDNA and the syncon promoter at the higher ionic strength, the conDNA and syncon promoter binding enthalpies are the same only for cAMP-ligated CRP but different for cAMP-ligated T127L and for cAMP-ligated S128A. Since the differences between the conDNA and promoter binding enthalpies are maintained at the high salt concentration for two of the three complexes investigated, it is unlikely that long-range electrostatic interactions on the promoter can account for the differences in the cNMP-ligated CRP/mutant binding enthalpies to conDNA and to the syncon promoter.

**SANS Measurements.** SANS measurements on the cAMP-ligated CRP−conDNA complex in 0% D_2O buffer solution over a concentration range of 0.01−0.04 mM are presented in Figure 4, in the form of Guinier, or ln(I) versus $Q^2$, plots. Consistent values of $Rg(\text{Complex})$ were extracted from the slopes of the fitted curves in Figure 3, $28.8 \pm 1.7$ Å at 0.04 mM, $26.6 \pm 0.6$ Å at 0.02 mM, and $26.4 \pm 0.8$ Å at 0.01 mM, and thus, show that aggregation effects are absent in these solutions. Averages of $Rg(\text{Complex})$ values over all concentrations yield a value of $27.3 \pm 0.8$ Å for the CRP complex and $28.7 \pm 0.6$ Å for the CRP* complex as shown in Table 4. These values are larger than the $Rg$ values of $21.6 \pm 0.2$ Å for CRP alone and $22.2 \pm 0.2$ Å for CRP* alone (22). Similar increases in $Rg(\text{Complex})$ that are independent of concentration were also observed for the cAMP-ligated CRP−lacDNA complex and are shown in Table 4. A similar value for $Rg(\text{Complex})$ was also observed for cAMP-ligated T127L−conDNA complex. The increase in $Rg(\text{Complex})$ could result from addition of the DNA component and/or from an increase in $Rg(\text{CRP})$ upon DNA binding. Table 4 also shows the values of $Rg$ for the cAMP-ligated CRP/mutant−syncon promoter complexes. The $Rg$ is considerably larger for the syncon promoter complexes, indicating that the addition of the DNA is clearly making some contribution to the increased $Rg$ in this case, most likely because of the addition of the much larger DNA component.

To differentiate the neutron scattering intensity of the DNA component from that of the CRP, a contrast variation series
of SANS measurements were performed on cAMP-ligated CRP−conDNA, CRP*−conDNA, cAMP-ligated CRP−syncon promoter, and cAMP-ligated CRP*−syncon promoter complexes in 0, 15, and 70% D$_2$O/H$_2$O buffer solutions. The results of the SANS measurements from the cAMP-ligated CRP-mutant−conDNA complexes in these D$_2$O/H$_2$O buffer mixtures are presented in Figure 5. As can be seen in Figure 5, the maximum intensities are at 0% D$_2$O buffer where the intensity is a sum of the intensity contributions from the protein and the DNA components of the complex, while at 70% D$_2$O, the intensities are at a minimum since the contribution from the DNA component is essentially zero. As shown in Table 4, the $R_g$(Complex) values were the same for cAMP-ligated CRP−conDNA and CRP*−conDNA complexes in all the D$_2$O/H$_2$O buffer mixtures, even though the scattering from the DNA component is negligible in 70% D$_2$O buffer and maximum in 0% D$_2$O buffer. This indicates that a conformational change of the CRP component, rather than the addition of the DNA component, is primarily responsible for the larger-than-predicted $R_g$(Complex).

![Figure 4: Low angle portions of the SANS intensities, on a ln(I) vs $Q^2$ scale, from the cAMP-ligated CRP−DNA complex in 0% D$_2$O buffer, measured at concentrations of 2 (○), 1 (△), and 0.5 mg ml$^{-1}$ (∆). Radii of gyration, extracted from the slopes of the fitted curves (−) using a Guinier analysis, were 28.8 ± 1.7, 26.6 ± 0.6, and 26.4 ± 0.8 Å, respectively.](image-url)

This is further substantiated by decomposition of the SANS contrast variation intensities from the complexes into the $I_{CRP}(Q)$ and $I_{DNA}(Q)$ components of eq 2 as shown in Figure 6a for the cAMP-ligated CRP−conDNA complex and in Figure 6b for the CRP*−conDNA complex. Radii of gyration for the CRP component of the complex were then determined from the distance distribution functions, $P(r)$. The results for the components of the complexes are summarized in Table 5 along with those values expected from the CRP and DNA components in the energy-minimized X-ray crystal structure by Parkinson et al. (12). The $R_g$(CRP) component in both complexes is ~28.5 Å, more than 6 Å larger than the 21.6 ± 0.2 Å observed for the cAMP-ligated CRP complex alone (22) and the 22.6 Å value predicted from the energy-minimized X-ray crystal structure by Parkinson et al. (12). It should be emphasized that the DNA duplex in the crystal structure consists of two staggered 17/14 base sequences annealed together (12) and not the continuously linked conDNA duplex employed in this investigation. This may account for the larger $R_g$(DNA) in the X-ray crystal structure. The $R_g$ previously reported (4) for the CRP component of the cAMP-ligated CRP−lacDNA was smaller,
experiments have been performed on three CRP influenced by aggregation effects. Now that additional
mined from data obtained at twice the concentration and was eliminated by verifying that
complexes at lower concentrations, and aggregation effects have been eliminated by verifying that \( R_g \) value of the CRP component of the complex undergoes an increase upon binding to the DNA. This conclusion is further substantiated by the results from the contrast variation series performed on the cAMP-ligated CRP/mutant complexes shown, in Table 4, a larger \( R_g \) in 0 and 15% \( D_2O \), where the DNA component makes a strong contribution to the scattering intensity. However, the \( R_g \) value in 70% \( D_2O \) is approximately the same as the \( R_g \) for the CRP–conDNA complexes under the same conditions. Upon decomposition of the data into CRP and DNA components, the \( R_g \) value of the CRP and CRP* components were found to be \( \sim 28 \) Å, as shown in Table 5.

**Model of the CRP Component.** To identify the most likely source of the increase in \( R_g \) (CRP), Monte Carlo simulations of the scattering neutron intensities from an energy-minimized distorted conformation of the CRP–DNA complex were generated and compared to the SANS data. A distorted conformation of the X-ray crystal structure of the cAMP-ligated CRP–30 bp DNA complex (12) was employed for the starting conformation in the energy-minimization procedure. This distortion was based on the consistent hypersensitivity to several proteases observed in the B-domain to the protease. Consistent with this notion, the distortion of the CRP conformation consisted of increasing the torsional angle between the C and Bα-helices. The resulting energy-minimized distorted structure is shown in the lower structure in Figure 7. The upper structure in Figure 7 is the energy-minimized closed form of the cAMP-ligated CRP structure without the bound DNA (7). Ramachandran plots comparing the protein component of the minimized X-ray crystal structure (12), which gives a model SANS intensity similar to the cAMP-ligated CRP structure without the bound DNA (7) to the minimized model structure, are presented in Figure 8. The peptide backbones of the CRP

### Table 4: Radius of Gyration of the DNA-Bound CRP Complex in \( D_2O/H_2O \) Buffer Mixtures

<table>
<thead>
<tr>
<th>Complex</th>
<th>DNA</th>
<th>Complex concentration in mM (no. of det.)</th>
<th>Radius of Gyration (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>cAMP-ligated CRP</td>
<td>conDNA</td>
<td>0.01±0.04 (5)</td>
<td>27.3±0.8</td>
</tr>
<tr>
<td>in 0% ( D_2O )</td>
<td>conDNA</td>
<td>0.04</td>
<td>28.8±1.7</td>
</tr>
<tr>
<td>in 15% ( D_2O )</td>
<td>conDNA</td>
<td>0.04</td>
<td>29.9±0.08</td>
</tr>
<tr>
<td>in 70% ( D_2O )</td>
<td>conDNA</td>
<td>0.04</td>
<td>30.3±1.2</td>
</tr>
<tr>
<td>cAMP-ligated CRP*</td>
<td>lacDNA</td>
<td>0.05–0.90 (2)</td>
<td>31±2</td>
</tr>
<tr>
<td>in 0% ( D_2O )</td>
<td>conDNA</td>
<td>0.04–0.06 (3)</td>
<td>28.7±0.6</td>
</tr>
<tr>
<td>in 15% ( D_2O )</td>
<td>conDNA</td>
<td>0.06</td>
<td>28.9±1.0</td>
</tr>
<tr>
<td>in 70% ( D_2O )</td>
<td>conDNA</td>
<td>0.04</td>
<td>28.0±1.4</td>
</tr>
<tr>
<td>cAMP-ligated T127L</td>
<td>conDNA</td>
<td>0.02(1)</td>
<td>28.8±0.6</td>
</tr>
<tr>
<td>cAMP-ligated CRP</td>
<td>syncon promoter</td>
<td>0.04</td>
<td>55±2</td>
</tr>
<tr>
<td>in 15% ( D_2O )</td>
<td>syncon promoter</td>
<td>0.04</td>
<td>56±2</td>
</tr>
<tr>
<td>in 70% ( D_2O )</td>
<td>syncon promoter</td>
<td>0.03</td>
<td>30±3</td>
</tr>
<tr>
<td>cAMP-ligated CRP*</td>
<td>syncon promoter</td>
<td>0.04</td>
<td>53±1</td>
</tr>
<tr>
<td>in 15% ( D_2O )</td>
<td>syncon promoter</td>
<td>0.03</td>
<td>53±1</td>
</tr>
<tr>
<td>in 70% ( D_2O )</td>
<td>syncon promoter</td>
<td>0.04</td>
<td>28±2</td>
</tr>
</tbody>
</table>

\( ^* \) Exp. \( R_g \) values were obtained at each concentration using the Guinier approximation, as stated in the text, at only the low \( Q \) data points from \( l(Q) \) in the region where \( QR_g \sim 1 \). Those from x-ray were calculated directly from the protein component of the minimized Parkinson et al. X-ray crystal structure (12) shown in Figure 7 and those from model were calculated from the minimized distorted structure shown in Figure 7 and described in the text. Except for the T127L + cAMP sample, the stated \( R_g \) value is the average \( R_g \) from all of the separate determinations, and the uncertainty is the standard deviation of that average divided by \( n^{1/2} \), where \( n \) is the number of determinations. The uncertainty for the T127L + cAMP sample was obtained from the statistical uncertainty in the slope of the Guinier approximation equation.

**Figure 5:** Contrast variation SANS intensities from cAMP-ligated CRP–DNA complex (a) and the CRP*–DNA complex (b) in 0 (○), 15 (○), and 70% \( D_2O \) (*) buffers.

**Figure 6:** Decomposition of SANS contrast variation intensities from the cAMP-ligated CRP–DNA complex (a) and the CRP*–DNA complex (b) into the intensities from the individual components, \( I_{cAMPR}(Q) \) (●) and \( I_{DNA}(Q) \) (○), as described in eq 2.

\( 23 \pm 1 \) Å, but was determined without including the data at the lowest \( Q \) values to minimize what, at the time, were thought to be aggregation effects. This value is also smaller than the previous value (4) of \( 32 \pm 1 \) Å for \( R_g(DNA) \) of the cAMP-ligated CRP–lacDNA complex, which was determined from data obtained at twice the concentration and was influenced by aggregation effects. Now that additional experiments have been performed on three CRP–DNA complexes at lower concentrations, and aggregation effects have been eliminated by verifying that \( R_g \) (Complex) does not change with concentration, it can be concluded that the \( R_g \) value of the CRP component of the complex undergoes an increase upon binding to the DNA. This conclusion is further substantiated by the results from the contrast variation series performed on the cAMP-ligated CRP/mutant complexes shown, in Table 4, a larger \( R_g \) in 0 and 15% \( D_2O \), where the DNA component makes a strong contribution to the scattering intensity. However, the \( R_g \) value in 70% \( D_2O \) is approximately the same as the \( R_g \) for the CRP–conDNA complexes under the same conditions. Upon decomposition of the data into CRP and DNA components, the \( R_g \) value of the CRP and CRP* components were found to be \( \sim 28 \) Å, as shown in Table 5.
From the model were calculated from the protein component of the minimized distorted structure shown in Figure 7 and described in the text. X-ray were calculated directly from the protein portion of the minimized Parkinson et al. (12) cAMP-ligated CRP without bound DNA and the P(r) function of a 1:1 mixture of the same bent rods with cross-sectional radii of 10 Å and the peak at \( r \approx 20 \) Å in the experimental data correspond remarkably well to those obtained from the simulations.

Model of the DNA Component. The bent rod uniform scattering length density model of the DNA is shown in the lower structure of Figure 7. Since the DNA binding site of the protein portion of the distorted energy-minimized structure is not significantly different from that in the undistorted energy-minimized structure of Parkinson (12), the latter structure was used to help position the bent rod DNA model in the distorted structure in Figure 7. The distance distribution function, P(r), was calculated for the bent rod DNA model and compared to that obtained from the DNA portion of the data from the cAMP-ligated CRP–conDNA complex. As shown in Figure 10, there is a maximum at \( r \approx 10 \) Å in the model P(r) function, but the maximum in P(r) obtained from the experimental data occurs at \( r \approx 25 \) Å. This shift in the P(r) maximum can be attributed to an increase in the cross-sectional radius of the DNA duplex in solution. To determine if this can account for the shift in the P(r) maximum, the P(r) function of a 1:1 mixture of the same bent rods with cross-sectional radii of 10 and 20 Å was simulated and is also shown in Figure 10. Although Rg(DNA) = 27 Å from this scattering model is still larger than the Rg = 25 Å value obtained from the data, the distance distribution of this mixed bent rod model fits that of the experimental DNA component data fairly well, and this larger cross-sectional area would be consistent with a slight localized unwinding of about 29° for the promoter upon CRP binding, as reported by Douc-Rasy et al. (23) employing an electrophoretic procedure.

SANS measurements on the cAMP-ligated CRP–105 bp syncon promoter complex in D2O/H2O buffer mixtures exhibited a Rg of 53 ± 3 Å for the promoter, as shown in Table 5. Using bend rod models for simulating the neutron scattering from the104 bp syncon promoter, it was found that the Rg for the 104 bp DNA duplex bent as shown in

Table 5: Radius of Gyration of the CRP and DNA Components of the Complexes

<table>
<thead>
<tr>
<th>CRP complex</th>
<th>DNA</th>
<th>complex concentrated (mM)</th>
<th>( P(r) )</th>
<th>X-ray</th>
<th>model</th>
<th>( P(r) )</th>
<th>rod model</th>
</tr>
</thead>
<tbody>
<tr>
<td>cAMP-ligated CRP</td>
<td>conDNA</td>
<td>0.04</td>
<td>28.2 ± 0.4</td>
<td>22.9</td>
<td>28.2</td>
<td>26 ± 1</td>
<td>29.1</td>
</tr>
<tr>
<td>CRP*</td>
<td>conDNA</td>
<td>0.06</td>
<td>28.7 ± 0.6</td>
<td>22.9</td>
<td>28.2</td>
<td>27 ± 2</td>
<td>29.1</td>
</tr>
<tr>
<td>cAMP-ligated CRP</td>
<td>syncon</td>
<td>0.04</td>
<td>28 ± 1</td>
<td>22.9</td>
<td>28.2</td>
<td>53 ± 3</td>
<td></td>
</tr>
<tr>
<td>cAMP-ligated CRP*</td>
<td>syncon</td>
<td>0.04</td>
<td>27 ± 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* The Rg values from \( P(r) \) were obtained from the first moment of the distance distribution function, \( P(r) \), using all the data points. Those from X-ray were calculated directly from the protein portion of the minimized Parkinson et al. (12) X-ray crystal structure shown in Figure 7, and those from the model were calculated from the protein component of the minimized distorted structure shown in Figure 7 and described in the text.
If the additional 64 base pairs are added symmetrically on each side of the DNA in Figure 7, the calculated $R_g$ is 48 Å. Since the $R_g$ value for a straight 104 bp duplex is 103 Å, the experimental results are clearly in better agreement with the bent DNA models. The U-shaped DNA model agrees best, but the $R_g$ value is a bit smaller (48 Å) than the observed $53 \pm 3$ Å experimental result, while the cane shaped DNA model gives an $R_g$ that is too large (69 Å). However, since the long tail portion of the cane was modeled to be straight, it is likely that this tail portion can be bent in such a way as to mimic the $R_g$ value obtained from experiment. This was not done because, without additional structural data, there are simply too many possible ways to model this additional structural bend.

**DISCUSSION**

Simulations of the SANS results show that the 104 bp *syncon* promoter is bent by CRP/mutants in an U-shaped...
conformation that is partially cane shaped since the Rg of 53 ± 3 Å is between the calculated Rg = 48 Å for the exclusive U shape, and the calculated Rg = 69 Å for the exclusive cane shape bend in solution. Bending of the promoter by CRP/mutant would undoubtedly facilitate the contacts between the CRP/mutants and the RNA polymerase as well as contacts between the promoter and the RNA polymerase. It is unlikely that part of the interaction thermodynamic changes may be attributed to these additional contacts between the promoter and the RNA polymerase since the interaction parameters are different between cAMP-ligated CRP and cAMP-ligated CRP*, while the bend in the promoter remains the same with an Rg = 53 Å. In addition, simulations of the SANS data on the shorter 40 bp conDNA and lacDNA duplexes bound to CRP/mutants indicate that the DNA duplexes have a larger cross-sectional area than the expected 10 Å diameter for a DNA duplex, which would imply that CRP/mutants induce a partial unwinding of the bound DNA. Since partial unwinding of the DNA is characteristic of the transcriptionally active open form of RNA polymerase, CRP/mutants facilitate the changes associated with isomerization of the RNA polymerase promoter complex from the closed to the open form.

CRP also increases the binding affinity of the RNA polymerase to the promoter via additional indirect contacts through the bound CRP/mutants. The free energy of this interaction between the CRP and the bound RNA polymerase is entropically driven and ranges from −4.5 to −1.4 kJ mol⁻¹ as shown in Table 2. This range is comparable to the estimated free energy range of −5.6 to −3.2 kJ mol⁻¹, determined from gel electrophoresis mobility shift assays, for the interaction between promoter-bound CRP at the −61.5 site and the lac repressor protein bound at a downstream site centered near the transcription start point (24). Although the proteins are distally bound, bending of the promoter by the CRP, as observed here in the SANS measurements, would facilitate this interaction. Estimates for a similar interaction between the lac repressor protein and RNA polymerase bound to sites that overlap are, however, higher and range from −12 to −8 kJ mol⁻¹ (24). This higher range may result from different amino acid residue interactions between the lac repressor protein and the RNA polymerase as well as the closer proximity between the two because of the overlap of their binding sites on the promoter. The interactions between CRP/mutants and the RNA polymerase are entropically driven and exhibit entropy−enthalpy compensation, where changes in the enthalpy are compensated by changes in the entropy. Enthalpy−entropy compensation implies the involvement of water in this interaction. Although there are small endothermic contributions to the interaction enthalpies between RNA polymerase and the CRP/mutants, the overall entropically driven nature of the interaction implies a classical protein−protein hydrophobic interaction. This interaction enhances the binding affinity of the RNA polymerase to the promoter by factors ranging from 1.8 for cAMP-ligated S128A to 6.0 for cAMP-ligated CRP* (Table 2) that contribute to the observed enhancement of transcriptional activation by CRP/mutant. However, this increase can only partially account for the enhancement of transcriptional activation, which ranges from factors of 11−13 (5). This further supports the involvement of other mechanisms such as the unwinding of the promoter by CRP/mutants.

The interaction between CRP/mutants and the RNA polymerase may be facilitated by the conformational change in CRP/mutants upon binding to conDNA and lacDNA as evident in the increase in the Rg of CRP and CRP* from about 22 to 30 Å. This implies that the conformation of the CRP bound to the DNA alone remains the same as when it is bound to the promoter adjacent to the bound RNA polymerase. That this is indeed the case has been demonstrated recently where it was shown that the X-ray crystal structure of CRP in a complex with a 44 bp DNA duplex and the CTD unit of RNA polymerase was superimposable on the structure of CRP bound to the DNA alone (25). It should be emphasized that these structures in the crystal state do not necessarily rule out the conformational change observed in the SANS results for CRP in going from the free state to the DNA-bound state in solution. A conformational change resulting from increasing the torsional angle between the C and the B α-helices agrees with the experimental SANS data is shown in Figure 7. In a recent investigation of the X-ray crystal structures of protein−DNA complexes (26), it was shown that 11 out of 24 of the systems studied involve rigid domain movements in the protein upon DNA binding, similar to those described here by the distortion in the CRP energy-minimized structure. This distorted CRP structure results in a wider separation between the subunits of promoter-bound CRP, for example, from 8.9 ± 2.5 to 13 Å near the midpoint of the interface at S128, which would allow a closer approach of the proximal CRP subunit to the adjacent bound RNA polymerase. Since residues 156−164 located within the C-terminal domain of the subunit adjacent to the bound RNA polymerase have been implicated as the contacts with RNA polymerase (I) and not the residues in the N-terminal domain that becomes more exposed in the distorted structure in Figure 7 (I), it is more likely that the wider separation of the subunits in the distorted structure would facilitate contact with the bound RNA polymerase. Since DNA-bound CRP contains both subunits in the closed form, the distortion induced by these mutations may explain why CRP* enhances transcription even without ligation to cAMP as is necessary for the wild-type CRP. The exposure of this region in the N-terminal domain of CRP, however, may directly contribute to the enhancement of transcription of the promoter by CRP bound to a second site centered at −41.5 through its contacts with RNA polymerase. More specifically, a surface loop at K52 in the N-terminal domain of CRP bound at the −41.5 site has been shown to make contact with RNA polymerase (I), and the interaction of the K52 loop would be directly affected by the conformational change in CRP shown in Figure 7. However, distortions within CRP that may affect its interaction with RNA polymerase are evident in the CRP* mutant, which exhibits the maximum binding interaction with RNA polymerase in Table 2. A recent X-ray crystal structure of cAMP-ligated CRP* at a resolution of 2.2 Å shows that the interface mutations, T127→L and S128→A on the C α-helix, do indeed introduce some distortion in the CRP as evident by additional bending of the C-terminal domain of one of its subunits, the open subunit, more like the other closed subunit than is observed in the corresponding open subunit of the wild-type cAMP-ligated CRP (27).
A comparison of the CRP/mutant binding enthalpies to the syncon promoter in the absence of bound RNA polymerase and to conDNA show that the lower endothermic binding enthalpies for the promoter cannot be attributed to exothermic contributions from additional long-range electrostatic forces on the promoter. The lower binding enthalpy for the longer promoter may be attributed to the energy required for bending of the promoter that is observed in the SANS measurements. The work contribution from bending of the DNA (28) is

$$W_{\text{bend}} = 0.5B(\theta^2)/L$$

so that the total enthalpy change upon binding and bending of the DNA is

$$\Delta H_B = W_{\text{bend}} + \Delta H_I$$

where $L$ is the length of DNA undergoing the bending, $\theta$ is the angle of the bend, and $\Delta H_I$ refers to the exothermic enthalpy change resulting from interactions between the amino acid residues on CRP/mutants and the nucleotide bases on the promoter. Since the number of the consensus bases in both the conDNA and the syncon promoter that interact with CRP is the same, then $\Delta H_I$ would be expected to remain the same for both DNA duplexes. However, for the same degree of bending extended over a promoter length of 104 bp, the energy required is only 38% of the energy required for extension of the same bend over the shorter 40 bp conDNA duplex. That this bend is extended over a longer base sequence in the 104 bp syncon promoter is evident from the more predominantly U shape of the promoter than the expected cane shape bend that extends over part of the promoter. This would be reflected in a reduction of the endothermic enthalpy required for CRP/mutant binding to the promoter relative to binding to the shorter conDNA, as is indeed observed in Table 1. This difference in the enthalpy would only marginally affect the binding free energy to the promoter since enthalpy–entropy compensation minimizes changes in the binding free energy.

ACKNOWLEDGMENT

S.G. would like to thank the DOE/Sloan foundation for a postdoctoral fellowship in computational molecular biology. B.D.W. is grateful to the Cottrell College Science Program and support through Research Corporation.

REFERENCES


