Equilibrium DNA Binding of Sac7d Protein from the Hyperthermophile *Sulfolobus acidocaldarius*: Fluorescence and Circular Dichroism Studies[†,‡]

James G. McAfee,[§] Stephen P. Edmondson,* Irene Zegar, and John W. Shriver*

Department of Medical Biochemistry, School of Medicine, Southern Illinois University, Carbondale, Illinois 62901-4413

Received October 26, 1995; Revised Manuscript Received January 19, 1996[⊗]

ABSTRACT: The thermodynamics of the binding of the Sac7d protein of *Sulfolobus acidocaldarius* to double-stranded DNA has been characterized using spectroscopic signals arising from both the protein and the DNA. Ligand binding density function analysis has been used to demonstrate that the fractional change in protein intrinsic tryptophan fluorescence quenching that occurs upon DNA binding is equal to the fraction of protein bound. Reverse titration data have been fit directly to the McGhee−von Hippel model [McGhee, J., & von Hippel, P. (1974) *J. Mol. Biol. 86*, 469−489] using nonlinear regression. Sac7d binds noncooperatively to poly(dGdC)·poly(dGdC) with an intrinsic affinity of 6.5×10^6 M^{-1} and a site size of 4 base pairs in 1 mM KH_2PO_4 and 50 mM KCl (pH 6.8). Some binding sequence preference is noted, with the binding to poly(dIdC)·poly(dIdC) over 10-fold stronger than to poly(dAdT)·poly(dAdT). The binding is largely driven by the polyelectrolyte effect and is consistent with a release of 4.4 monovalent cations from DNA upon complex formation or the formation of 5 ion pairs at the protein−DNA interface. Extrapolation of salt back-titration data to 1 M KCl indicates a −2.2 kcal/mol nonelectrostatic contribution to the binding free energy. A van't Hoff analysis of poly(dGdC)·poly(dGdC) binding shows that the binding enthalpy is approximately zero and the process is entropically driven. The affinity decreases slightly between pH 5.4 and 8.0. There is no significant difference between the binding parameters of recombinant Sac7d and native Sac7 proteins, indicating that methylation of the native protein has no effect on the DNA binding function. The binding of Sac7d to various DNAs leads to a significant increase in the DNA long-wavelength circular dichroism (CD) band, the intensity of which shows a sigmoidal dependence on Sac7d concentration. The sigmoidal CD binding isotherm can be quantitatively modeled by a conformational transition in the DNA that is cooperatively induced when protein monomers are bound within a given number of base pairs, ranging from zero for poly(dIdC)·poly(dIdC) to 8 or less for poly(dAdG)·poly(dCdT).

The Archaeon *Sulfolobus acidocaldarius* is a hyperthermophile which grows in acidic hot springs with a minimum growth temperature of 60 °C and a maximum of 85 °C (Stetter et al., 1990). It produces a number of small basic proteins ranging in molecular weight from 7000 to 10 000 (Kimura et al., 1984; Grote et al., 1986; Choli et al., 1988a) which are thought to play a role in DNA compaction and helix stabilization at the high growth temperature (Kimura et al., 1984; Grote et al., 1986; Lurz et al., 1986; McAfee, 1995; Choli et al., 1988a,b; Reddy & Suryanarayana, 1988, 1989). The 7 kDa proteins of *S. acidocaldarius* consist of five species, designated Sac7a−e[1] in order of increasing basicity (Kimura et al., 1984; Grote et al., 1986; Choli et al., 1988b; McAfee et al., 1995). Sac7d and Sac7e, which differ by six amino acid residues, are coded for by distinct genes, and the Sac7a and Sac7b proteins are carboxy terminal truncated forms of the Sac7d protein (McAfee et al., 1995). The native Sac7 proteins are all monomethylated at specific lysines to some extent. Sac7d (66 residues, 7600 MW) has been overexpressed in *Escherichia coli*, and fluorescence, circular dichroism, NMR, and binding assays indicate that the recombinant and native proteins have essentially identical structures. The solution structure of the Sac7d protein has been reported (Edmondson et al., 1995).

The binding of the 7 kDa proteins to DNA has been investigated using electron microscopy, filter binding assays, and fluorescence titrations (Kimura et al., 1984; Dijk & Reinhardt, 1986; Grote et al., 1986; Lurz et al., 1986; Choli et al., 1988a). Filter binding assays have suggested that the affinity for double-stranded DNA is relatively low (Dijk & Reinhardt, 1986; Grote et al., 1986), but equilibrium binding constants were not reported. Baumann et al. (1994) have recently followed the DNA binding of the homologous Sso7 proteins from *Sulfolobus solfataricus* using fluorescence quenching. They estimated a binding site size of 3−6 base pairs (bp) for double-stranded DNA (0.02 M TRIS, pH 7.4) and argued by visual inspection that the affinity in low salt was approximately $0.5-1 \times 10^6$ M^{-1}.

All of the binding studies described above have been qualitative. The availability of the recombinant protein allows for a quantitative characterization of the binding using a homogeneous protein preparation without the heterogeneity present in the native due to multiple isoforms and lysine

[†] This work was supported by the National Institutes of Health (GM 49686).

[‡] The Brookhaven Protein Data Bank accession code for Sac7d is 1sap.

* Authors to whom correspondence should be sent. Phone: 618-453-6479. Fax: 618-453-6408. E-mail: jshriver@som.siu.edu and sedmondson@som.siu.edu.

[§] Current address: Department of Molecular Biology, Vanderbilt University, 1161 21st Avenue, Nashville, TN 37235.

[⊗] Abstract published in *Advance ACS Abstracts,* March 15, 1996.

[1] Abbreviations: CD, circular dichroism; Sac7, a group of 7 kDa DNA-binding proteins from *Sulfolobus acidocaldarius*, individually referred to as Sac7a−e in order of increasing basicity; Sso7, a group of 7 kDa DNA-binding proteins from *Sulfolobus solfataricus*.

0006-2960/96/0435-4034$12.00/0

methylation. We present here a study of the DNA binding of the recombinant form of Sac7d, the isoform which predominates *in vivo*. Our interest is in the thermodynamics of binding and the relative contributions of electrostatic and nonelectrostatic factors to the binding energy. The apparent simplicity of the 7 kDa DNA binding proteins of *Sulfolobus* makes them useful systems for investigations into the basic principles associated with non-sequence specific DNA binding.

We use the model independent ligand binding density function method of Bujalowski and Lohman (1987) to show that the fractional change in fluorescence quenching of the single tryptophan in Sac7d observed upon DNA binding is equal to the fraction of protein bound. With this, we are able to construct binding isotherms and fit the fluorescence data using nonlinear regression to obtain intrinsic binding affinities, site sizes, and maximal fluorescence quenching as a function of salt concentration, temperature, and pH. The binding of the protein can be largely attributed to the polyelectrolyte effect and is entropically driven. The binding of the recombinant Sac7d protein is compared to that of native Sac7 protein, and it is concluded that lysine monomethylation of the native protein has no effect on the binding function of the protein. Finally, we present CD data which indicate that Sac7d induces a structural transition in the DNA that occurs cooperatively with increasing protein concentration. The apparent cooperativity is not observed in protein binding followed by either forward or reverse fluorescence titrations. We present a simple model which quantitatively accounts for the CD data as a DNA unwinding or bending that occurs when protein monomers are bound within a specified distance on the DNA.

MATERIALS AND METHODS

Growth of Microorganisms. E. coli and *S. acidocaldarius* were grown as described elsewhere (McAfee et al., 1995).

Protein Purification. Native Sac7 and recombinant Sac7d proteins were isolated and purified as described previously (McAfee, 1995). Protein concentrations were determined using an extinction coefficient of 1.1 mL mg^{-1} cm^{-1} for Sac7 and Sac7d (McAfee, et al., 1995).

Nucleic Acids. Synthetic DNA alternating copolymers were purchased from Pharmacia (Piscataway, NJ), and calf thymus DNA was obtained from Sigma Chemical Co. (St. Louis, MO). All were used without additional purification following dialysis against the appropriate buffer. Extinction coefficients and approximate lengths were supplied by Pharmacia: poly(dAdT)·poly(dAdT), $\epsilon_{262} = 6600$ $M^{-1}cm^{-1}$, 5 kilobases; poly(dGdC)·poly(dGdC), $\epsilon_{254} = 8400$ M^{-1} cm^{-1}, 0.8 kilobases; poly(dIdC)·poly(dIdC), $\epsilon_{251} = 6900$ M^{-1} cm^{-1}, 6 kilobases; poly(dAdC)·poly(dGdT), $\epsilon_{259} = 6500$ M^{-1} cm^{-1}, 5.6 kilobases; poly(dAdG)·poly(dCdT), $\epsilon_{260} = 6500$ M^{-1} cm^{-1}, 1 kilobase. The calf thymus DNA extinction coefficient was $\epsilon_{260} = 6600$ M^{-1} cm^{-1}.

Fluorescence Measurements. Fluorescence titrations were performed on a SLM 8000C spectrofluorimeter in the ratiometric mode with excitation at 295 nm (4 nm slit width) and emission monitored at 345 nm (8 nm slit width). Reverse titrations were performed in a 4 mL quartz fluorescence cuvette by addition of aliquots (e.g. 5 μL) of appropriate concentrated DNA stock solutions (e.g. 1 mM) to dilute protein solutions (typically 5 μM) using a carefully calibrated Unimetrics syringe with a utility stop allowing reproducible repetitive deliveries of a constant volume. The cuvette was stirred continuously with an internal cylindrical Teflon-coated magnetic stirrer using a magnetic drive mounted below the cuvette holder. Experimental fluorescence intensities were corrected for dilution and Raman light scattering where appropriate. We have previously noted (McAfee et al., 1995) that inner filtering effects appear to be balanced by a scattering contribution. This is evident under tight binding conditions where the observed quenching reaches a constant value and shows no change with increasing DNA concentration (e.g. Figure 4). Therefore, inner filter corrections have not been applied in most of the data herein. Photobleaching was not observed. Unless otherwise indicated, all binding studies were performed at 25 °C using a circulating water bath flowing through the cuvette holder.

Fluorescence Reverse Titration Data Analysis. Fluorescence reverse titration data were analyzed by directly fitting the dependence of the observed intrinsic tryptophan fluorescence quenching, Q_{obs}, on the total DNA concentration by nonlinear regression. Q_{obs} is defined by $(F - F_{obs})/F$, where F and F_{obs} are the fluorescence intensity of the protein in the absence and presence of DNA, respectively. Root mean square errors in the optimization were evaluated using an iterative computer algorithm which numerically searches for the value of the binding density, $\nu(i)$, which corresponds to the ith experimental total nucleic acid concentration, $D_{t\ exp}(i)$. To start the search, an initial value for the concentration of protein bound to DNA, $L_b(i)$, is obtained from

$$L_b(i) = [Q_{obs}(i)/Q_{max}]L_t(i) \quad (1)$$

since $L_b/L_t = Q_{obs}/Q_{max}$ (which is proven by the ligand binding density function analysis herein), where $Q_{obs}(i)$ is initially taken to be the experimentally observed value. $L_b(i)$ and $L_t(i)$ are the bound and total concentrations of protein for data point i, respectively, and Q_{max} is the maximal fluorescence quenching. Using the definition of the binding density, $\nu(i)$ (i.e. moles of bound ligand per mole of total lattice residues), we can obtain an initial value for $\nu(i)$.

$$\nu(i) = L_b(i)/D_{t\ exp} \quad (2)$$

This is substituted into the McGhee−von Hippel model (McGhee & von Hippel, 1974), as modified by Lohman and Mascotti (1992) to include both noncooperative and cooperative binding, to obtain the concentration of free protein, $L_f(i)$:

$$L_f(i) = \nu(i)/[K_{obs}[1 - n\nu(i)]\{2\omega[1 - n\nu(i)]/[(2\omega - 1) \times [1 - n\nu(i)] + \nu(i) + R]\}^{n-1}\{[1 - (n+1)\nu(i) + R]/[2[1 - n\nu(i)]]\}^2] \quad (3)$$

where

$$R = \{[1 - (n+1)\nu(i)]^2 + 4\omega\nu(i)[1 - n\nu(i)]\}^{1/2} \quad (4)$$

where n is the site size (i.e. the number of bases occluded by binding), ω is the cooperativity parameter, and K_{obs} is the intrinsic binding constant observed at the specified pH and salt concentrations. A new value for $L_b(i)$ is calculated from the total protein concentration,

$$L_b(i) = L_t(i) - L_f(i) \quad (5)$$

and the corresponding value for the total DNA concentration, $D_{t\ calc}$, is calculated using the definition of the binding density

$$D_{t\ calc}(i) = L_b(i)/\nu(i) \quad (6)$$

$D_{t\ calc}(i)$ is compared to the experimental value $D_{t\ exp}(i)$, and $\nu(i)$ is iteratively incremented until the difference between the calculated and experimental D_t values is acceptably small (typically less than 0.01% error). The value of $Q_{obs}(i)$ which corresponds to the final $D_{t\ calc}(i)$ [i.e. $D_{t\ exp}(i)$], is calculated by rearrangement of eq 1:

$$Q_{obs}(i) = Q_{max}L_b(i)/L_t(i) \quad (7)$$

Thus, $Q_{obs}(i)$ has been calculated for a given value of $D_{t\ exp}(i)$. The function requires four parameters, K_{obs}, n, ω, and Q_{max}, which were optimized by nonlinear regression (Bevington & Robinson, 1992). In practice, the data were fit initially by constraining ω to 1 (no cooperativity) and performing a three-parameter (K_{obs}, n, and Q_{max}) fit. The quality of this fit was compared to that obtained by assuming cooperativity, constraining the site size to that determined by the break point in a reverse titration under tight binding conditions (which is accurate if there is cooperativity), and fitting K_{obs}, ω, and $Q_{max.}$ Initial estimates of the three parameters were obtained using a computer program which allows graphic overlays of simulations onto the experimental data. The standard state used for the DNA was 1 M bases. Note that the derived thermodynamic parameters ΔG°_{obs} and ΔH°_{obs} are defined in terms of the protein and DNA concentrations only and do not include linkages to potential hydrogen ion and salt interactions.

Errors in the parameters were obtained using the Monte Carlo method for confidence interval estimation in nonlinear regression (Kamath & Shriver, 1989; Press et al., 1989; Straume & Johnson, 1992). Random data sets were constructed using a Gaussian random number generator (Miller, 1987) with the mean equal to the "true" Q_{obs} obtained from the optimized values for the parameters and the standard deviation equal to the standard deviation of the fit. Errors in the three optimized parameters are reported as standard deviations of the individual parameter distributions obtained from 50 Monte Carlo data sets.

Salt Back-Titrations Followed by Fluorescence. Salt back-titrations were performed by titration of a solution of known concentrations of DNA (typically 50 μM) and protein (typically 5 μM) with increasing concentrations of salt. Typically, 5 μL aliquots of a 4.0 M salt solution were added to 3.0 mL of protein−DNA complex. Experimental fluorescence intensities were corrected for dilution. The observed quenching for titration point i, $Q_{obs}(i)$, was converted into an intrinsic binding constant, K_{obs}, using the noncooperative McGhee−von Hippel model (McGhee & von Hippel, 1974):

$$K_{obs} = \frac{\nu(i)}{L_f(i)[1 - n\nu(i)]\left[\frac{1 - n\nu(i)}{1 - (n-1)\nu(i)}\right]^{n-1}} \quad (8)$$

where

$$L_f(i) = L_t(i) - L_b(i) = L_t(i) - \left[\frac{Q_{obs}(i)}{Q_{max}}\right]L_t(i) \quad (9)$$

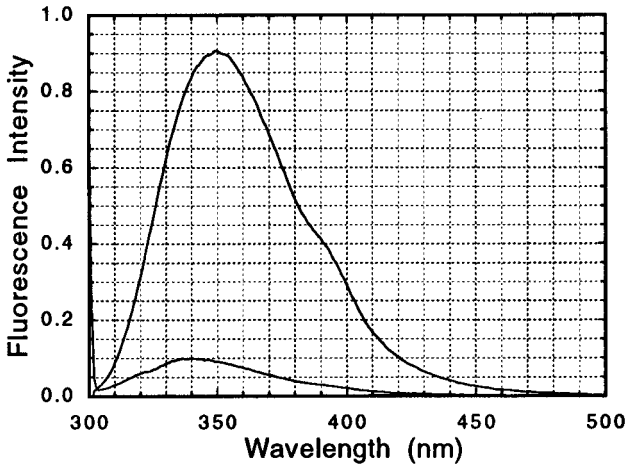

FIGURE 1: Quenching of the intrinsic tryptophan fluorescence of recombinant Sac7d upon binding to poly(dGdC)·poly(dGdC) in 10 mM K_2HPO_4 (pH 6.1) at 25 °C. Emission spectra of the free protein (upper curve, 3.8 μM) and protein−DNA complex (lower curve, 3.8 μM/68 μM) were collected with excitation at 295 nm using 4 nm excitation and emission slit widths. The Raman contribution has been removed by subtraction of the spectrum of buffer alone from each.

and

$$\nu(i) = \frac{L_b(i)}{D_t(i)} = \frac{Q_{obs}(i)L_t(i)}{Q_{max}\ D_t(i)} \quad (10)$$

Use of the noncooperative model is justified by the nonlinear least squares analysis of reverse titrations (see below).

Circular Dichroism Measurements. CD spectra were measured using an Aviv 62DS spectropolarimeter at 20 °C as previously described (McAfee et al., 1995). CD titrations were performed gravimetrically by addition of aliquots of a concentrated protein solution to a 1 cm path length cuvette containing a known concentration of DNA. The spectra were scaled to $\Delta\epsilon$ of the DNA. All DNA and protein solutions used for CD measurements were dialyzed against 0.01 M KH_2PO_4 (pH 7.0).

RESULTS

Reverse Titrations Followed by Fluorescence. The fluorescence emission of the single tryptophan (Trp 24) in Sac7d protein is quenched by nearly 90% upon protein binding to poly(dGdC)·poly(dGdC) in 10 mM KCl and 1 mM KH_2PO_4 (pH 6.8) with a blue shift of the emission maximum from 350 to approximately 340 nm (Figure 1). The high signal-to-noise ratio of the fluorescence change facilitates a precise quantitative analysis of the binding.

The graphical method of Bujalowski and Lohman (1987) has been used to investigate the relationship between the magnitude of the fluorescence signal change and the fraction of protein bound. An array of plots of the binding density function, $Q_{obs}(L_t/D_t)$, as a function of poly(dAdT)·poly(dAdT) concentration for various total Sac7d protein concentrations is shown in Figure 2. To investigate the dependence of Q_{obs} at low fractional binding (e.g. 0.2), it was necessary to accurately define the binding density function in the plateau region at low DNA concentrations, a region very susceptible to errors in protein and DNA concentration measurements. This was achieved by performing all of the titrations on the same day using the same nucleic acid stock solution. This

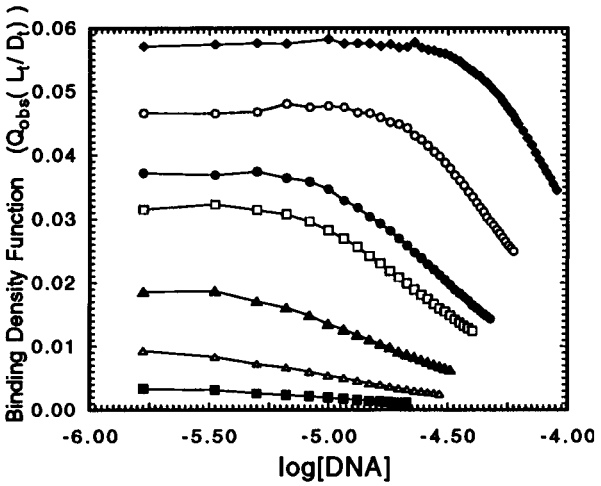

FIGURE 2: Binding density function, $Q(L_t/D_t)$, dependence on the total concentration of poly(dAdT)·poly(dAdT) in 10 mM K_2HPO_4 (pH 6.8) at 25 °C and 50 mM KCl. The total protein concentrations were as follows: 0.041 μM (■), 0.114 μM (△), 0.305 μM (▲), 0.695 μM (□), 0.93 μM (●), 1.97 μM (○), 4.01 μM (◆).

necessarily limits the number of titrations which can be performed, and also the extent of the titrations at high protein concentrations, but the higher quality of the data more than compensates for the lesser number of data points.

Horizontal lines corresponding to constant values of the binding density function (ranging from 0.0025 to 0.045) were graphically constructed through the family of curves shown in Figure 2. Sets of L_t and D_t concentrations for which the binding density function is constant satisfy the linear relationship

$$L_t = L_f + (\sum_j \nu_j) D_t = L_f + L_b \qquad (11)$$

where the sum is over all possible binding states (Bujalowski & Lohman, 1987). The slopes obtained from plots of L_t vs D_t provided the average binding density, $\sum \nu_j$, from which the concentration of bound Sac7d was obtained by multiplication by D_t. Determination of L_b from the slope rather than subtraction of the *y*-intercept (L_f) from L_t results in significantly lower noise levels in the following calculations.

With the concentration of bound protein known for all experimental measurements, the dependence of quenching magnitude on fractional binding (i.e. L_b/L_t) was investigated. A plot of the observed quenching with poly(dAdT)·poly(dAdT) binding as a function of the ratio of the bound Sac7d to the total Sac7d concentrations is linear with a *y*-intercept of 0.0043, a slope of 0.891, and a correlation coefficient of 0.996 (Figure 3A). Within experimental error, the *y*-intercept is zero. The value of Q_{max}, i.e. Q when $L_b/L_t = 1.00$, is 0.895. The linearity demonstrates that the fraction of protein bound equals the fractional change in the fluorescence signal. If multiple binding modes exist, they produce equal fluorescence quenching.

Similar analyses of the binding of Sac7d to poly(dGdC)·poly(dGdC) and calf thymus DNA also show linear dependencies of Q on L_b/L_t (Figure 3). For poly(dGdC)·poly(dGdC), the *y*-intercept, slope, and correlation coefficient were −0.0176, 0.856, and 0.994, respectively. For calf thymus DNA, the corresponding values were −0.0014, 0.889, and 0.999, respectively. The high quality of latter

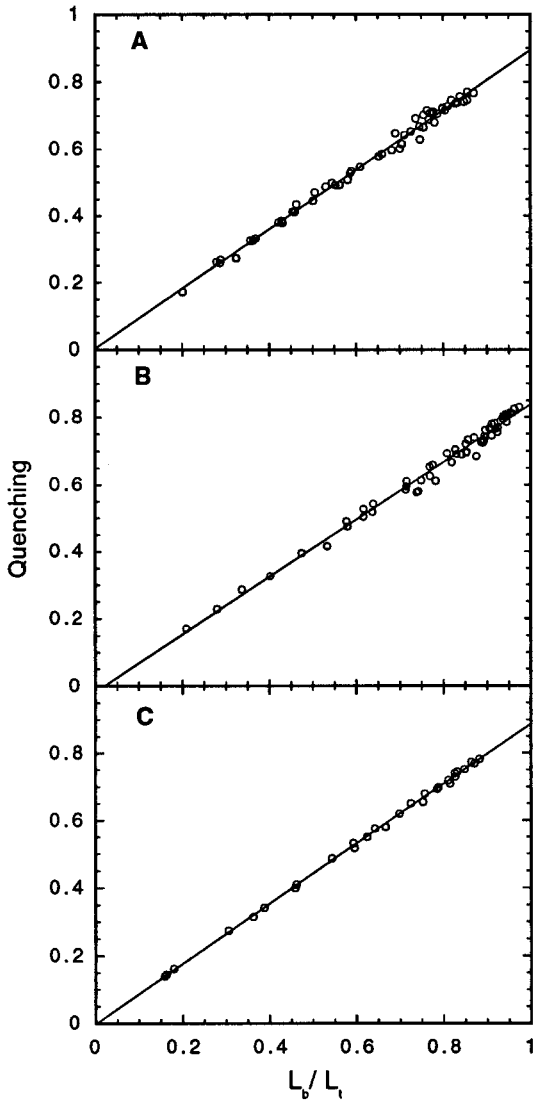

FIGURE 3: Intrinsic tryptophan fluorescence quenching, Q_{obs}, as a function of the fraction of Sac7d protein bound to (A) poly(dAdT)·poly(dAdT), (B) poly(dGdC)·poly(dGdC), and (C) calf thymus DNA in 10 mM K_2HPO_4 (pH 6.8), at 25 °C, and 50 mM KCl.

data, and the essentially identical parameters, is interesting given the expected heterogeneity.

Data Analysis of Reverse Titrations. Reverse titrations of the protein with nucleic acid were initially fit interactively by computer by overlaying on the experimental data theoretical simulations using the McGhee−von Hippel model (McGhee & von Hippel, 1974). This method is limited by the subjectivity of the analysis and the inability to evaluate errors in the parameters. Therefore, the binding parameters obtained from these overlays were used as initial estimates in a nonlinear least squares optimization.

It is mathematically impossible to solve the McGhee−von Hippel equation (eq 3) for the binding density. Kowalczykowski et al. (1986) have described a grid search procedure with the best values for the binding density as a function of ligand concentration found by interpolation. We use an alternative method here in which the value of $\nu(i)$ (and therefore Q_{obs}) which corresponds to each experimental D_t concentration is calculated numerically using an algorithm which iteratively refines $\nu(i)$ until the agreement between the experimental and calculated D_t concentrations is acceptable (see Materials and Methods). In our hands, this

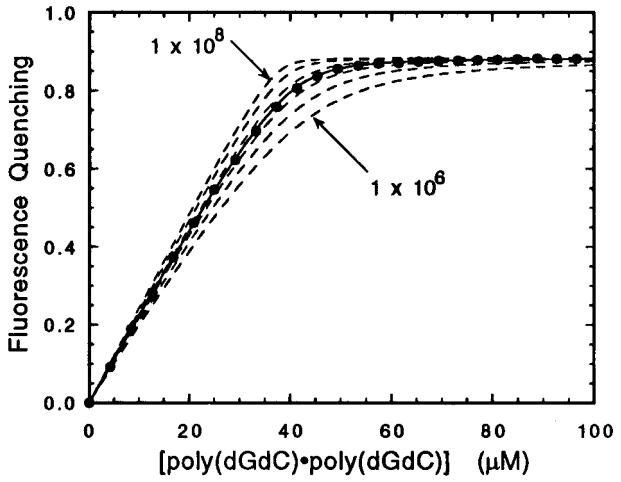

FIGURE 4: Reverse titration of recombinant Sac7d with poly(dGdC)·poly(dGdC) in 0.001 M KH_2PO_4 and 50 mM KCl (pH 6.8) followed by quenching of intrinsic tryptophan fluorescence intensity at 345 nm. The solid curve through the data (●) is the nonlinear least-squares fit obtained using the McGhee−von Hippel model with $K_{obs} = 6.51 \times 10^6$ M^{-1}, a site size of 4.3 base pairs, a Q_{max} of 0.884, and $\omega = 1$. Six simulated binding isotherms are shown for comparison using the same site size and Q_{max} but with K_{obs} equal to 1×10^6, 2×10^6, 4×10^6, 1×10^7, 4×10^7, and 1×10^8 M^{-1}. The Sac7d concentration was 3.8 μM. Note that the apparent site size of approximately 10 bases is greater than that derived from the fit (8.6 bases) due to site overlap.

algorithm is fairly robust and fast, typically requiring approximately 40 increments of $\nu(i)$ for each D_t, and less than 1 s for the calculation of a complete titration data set of approximately 25 points on a personal computer. Parameter optimization using this routine with nonlinear regression is also quite fast (typically 2 min). However, the minimization is susceptible to local minima problems and requires good initial estimates of the parameters which were obtained from graphical overlays. The occurrence of local minima was detected by overlaying the fitted isotherm onto the original data.

Figure 4 shows the nonlinear least squares fit of the binding of recombinant Sac7d to poly(dGdC)·poly(dGdC) using the McGhee−von Hippel model (McGhee & von Hippel, 1974) with the cooperativity parameter ω equal to 1. The fitted parameters (and associated errors) are 6.51 ($\pm$0.11) $\times$ 10^6 M^{-1} for K, 8.61 ($\pm$0.01) bases for n, and 0.884 ($\pm$0.001) for Q_{max}. The errors are derived from a Monte Carlo analysis and represent the quality of a fit of the data under the conditions used; they do not represent the effects of experimental errors in measuring protein and DNA concentrations (see below). Overlays of simulations using binding constants ranging from 1×10^6 to 1×10^8 M^{-1} are shown in Figure 4 to graphically demonstrate the reliability of the analysis under the conditions used here.

We were unable to obtain good fits of the data using the McGhee−von Hippel model (McGhee & von Hippel, 1974) with a cooperativity parameter greater than 1. On the basis of the quality of the fit of this and all other data herein using the noncooperative model, we conclude that there is no justification for inclusion of cooperativity in the analysis of the fluorescence titration data.

A nonlinear fit of the reverse titration data for *native* Sac7 binding to poly(dGdC)·poly(dGdC) (0.001 M NaH_2PO_4, 50 mM KCl, pH 6.8) yields an intrinsic binding constant of 4.56 ($\pm$0.07) $\times$ 10^6 M^{-1}, a site size of 8.42 ($\pm$0.01) bases, and a Q_{max} of 0.881 ($\pm$0.001).

Table 1: Sac7d Binding Parameters for Various Nucleic Acids[a]

DNA	$K \times 10^{-5}$ (M^{-1})	site size (bp)	Q_{max}
poly(dAdT)·poly(dAdT)	2.00 ($\pm$0.01)[b]	4.32 ($\pm$0.01)	0.876 ($\pm$0.001)
poly(dAdG)·poly(dCdT)	5.04 ($\pm$0.06)	4.30 ($\pm$0.02)	0.912 ($\pm$0.001)
calf thymus DNA	5.37 ($\pm$0.1)	4.55 ($\pm$0.03)	0.885 ($\pm$0.003)
poly(dAdC)·poly(dGdT)	6.23 ($\pm$0.08)	3.40 ($\pm$0.01)	0.863 ($\pm$0.001)
poly(dGdC)·poly(dGdC)	10.7 ($\pm$0.1)	3.88 ($\pm$0.01)	0.874 ($\pm$0.001)
poly(dIdC)·poly(dIdC)	34.2 ($\pm$0.9)	3.35 ($\pm$0.01)	0.910 ($\pm$0.001)

[a] Binding was measured by reverse titrations using fluorescence quenching at 25 °C in 10 mM KH_2PO_4 (pH 6.8) and 50 mM KCl. [b] Errors in parentheses are standard deviations of the distributions of the parameters obtained from fifty Monte Carlo data sets (see Materials and Methods) and represent the precision of the parameters from the nonlinear regression. Experimental errors due to inaccuracies in concentrations are estimated to be $\pm$10, $\pm$5, and $\pm$0.3% for K_{obs}, site size, and Q_{max}, respectively.

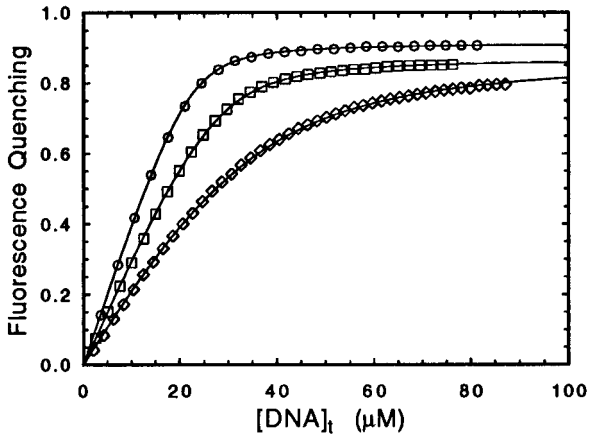

FIGURE 5: Reverse titrations of Sac7d with poly(dIdC)·poly(dIdC) (○), poly(dGdC)·poly(dGdC) (□), and poly(dAdT)·poly(dAdT) (◇) in 10 mM K_2HPO_4 (pH 6.8) at 25 °C and 50 mM KCl. The solid curves through the data are nonlinear least squares fits of the data using a noncooperative McGhee−von Hippel model with binding affinities, site sizes, and Q_{max} values given in Table 1.

Binding Parameters for Various Nucleic Acids. The binding affinity of Sac7d shows some sequence dependence as indicated by reverse titrations monitored by quenching of the protein intrinsic tryptophan fluorescence (Table 1 and Figure 5) (sequence dependence includes differences in structure of the DNA due to differences in sequence, and possible differences in specific protein−base interactions). The lowest affinity is observed for poly(dAdT)·poly(dAdT) and the strongest for poly(dIdC)·poly(dIdC) at pH 6.8 in 10 mM KH_2PO_4 and 50 mM KCl. Some variation in site size is observed, with an average of 3.97 base pairs and a trend toward decreasing site size with increasing affinity. The binding affinity for calf thymus DNA was very similar to that of poly(dAdG)·poly(dCdT) and poly(dAdC)·poly(dGdT).

Salt Dependence of Binding. The binding of Sac7d to DNA as a function of salt concentration was investigated using both reverse titrations at various salt concentrations and by salt back-titrations. Reverse titrations monitored with fluorescence using poly(dGdC)·poly(dGdC) in 1 mM K_2HPO_4 (pH 6.8) were performed at various KCl concentrations (Table 2). No significant differences were observed between parameters obtained in NaCl and KCl. The binding is strongly salt dependent with the affinity decreasing from 6.51

Table 2: Salt Dependence of the Binding of Native Sac7 and Recombinant Sac7d to Poly(dGdC)·Poly(dGdC)

protein	salt concentration[a]	$K \times 10^{-6}$ (M^{-1})	site size (bp)	Q_{max}
native Sac7	50 mM KCl	4.56 (±0.07)[b]	4.21 (0.01)	0.881 (0.001)
Sac7d	10 mM KCl	50.7 (±0.2)	3.81 (±0.01)	0.883 (±0.001)
	20 mM KCl	50.8 (±0.8)	3.85 (±0.01)	0.883 (±0.001)
	30 mM KCl	21.8 (±0.4)	4.11 (±0.01)	0.884 (±0.001)
	40 mM KCl	9.17 (±0.2)	4.03 (±0.01)	0.886 (±0.001)
	50 mM NaCl	7.64 (±0.2)	4.30 (±0.02)	0.883 (±0.001)
	50 mM KCl	6.51 (±0.1)	4.31 (±0.01)	0.884 (±0.001)
	60 mM KCl	2.83 (±0.04)	4.30 (±0.01)	0.888 (±0.001)
	80 mM KCl	0.693 (±0.009)	4.15 (±0.01)	0.891 (±0.001)
	100 mM KCl	0.266 (±0.002)	4.27 (±0.01)	0.891 (±0.001)

[a] All solutions contained 0.001 M KH_2PO_4 (pH 6.8). [b] See Footnote *b* in Table 1.

$\times 10^6$ M^{-1} to 2.66×10^5 M^{-1} with increasing KCl concentration from 50 to 100 mM. The theory of protein-nucleic acid binding developed by de Haseth et al. (1977) indicates that

$$-\left(\frac{\partial \log K_{obs}}{\partial \log[M^+]}\right) = m'\Psi + \frac{aK_X[X^-]}{1 + K_X[X^-]} \qquad (12)$$

where M^+ and X^- are the monovalent salt cation and anions, respectively, m′ is the number of ion pairs between protein and DNA, Ψ is the number of M^+ ions released from the DNA for each ion pair formed with the protein, and *a* is the number of anion binding sites on the protein with the intrinsic binding constant K_X. A plot of the log of K_{obs} vs the log of KCl concentration is linear (not shown), implying that anion release from the protein is thermodynamically insignificant at salt concentrations above 50 mM ($\partial \log K_{obs}/\partial \log[KCl]$ equal to −4.7 above 50 mM KCl). If Ψ is taken to be 0.88 for duplex DNA (Record et al., 1976), the number of ion pairs formed upon complex formation is predicted to be 5.3.

The error in the determination of K_{obs} as a result of errors in buffers and concentration measurements can be obtained from the magnitude of the residuals of a fit of K_{obs} vs [KCl] with $K_{obs} = x[KCl]^y$ for the data from 50 to 100 mM KCl. The residuals are less than 10% of K_{obs}, which is approximately 5 times larger than the Monte Carlo-derived error for each reverse titration fit. Over the same range, there is no significant change in site size, the average being 4.26 base pairs with an error of ±0.07. The Q_{max} also shows little variation with salt concentration with an average of 0.889 ± 0.003.

Salt Back-Titrations. The independence of Q_{max} with salt concentration permits a more precise determination of the salt dependence of the binding affinity using salt back-titrations. Figure 6A shows the observed fluorescence quenching as a function of KCl concentration. A plot of the log of K_{obs} vs the log of the salt concentration obtained from a salt back-titration is linear within experimental error with $\partial \log K_{obs}/\partial \log[KCl]$ equal to −4.4 (Figure 6B), in good agreement with the independent binding constant measurements from reverse titrations. There is negligible curvature in log K_{obs} vs log[KCl], again indicating that anion release from the protein is thermodynamically insignificant. The number of ion pairs formed is predicted to be 5.0.

The dependence of the binding free energy on log[KCl] is also linear (Figure 7). The polyelectrolyte effect contribution to the interaction energy goes to zero at 1 M K^+ (de Haseth et al., 1977), so extrapolation of the binding energy dependence to 1 M salt allows an estimate of the ionic and nonelectrostatic interaction free energy contribution to binding to be −1.2 kcal/mol. Identical results were obtained for the native Sac7 protein (Figure 7).

NaCl salt back-titrations of recombinant Sac7d complexed with poly(dGdC)·poly(dGdC) coincided within experimental

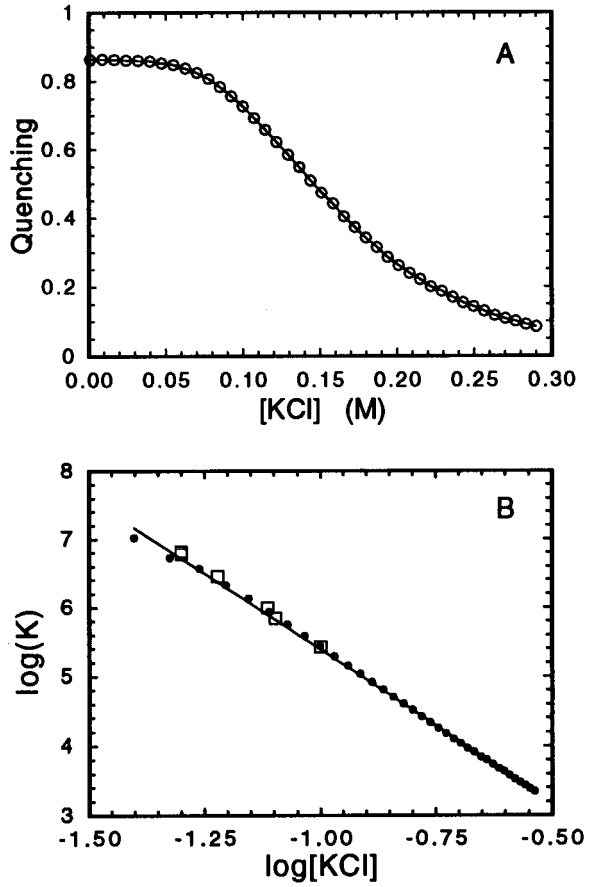

FIGURE 6: Salt back-titration of Sac7d−poly(dGdC)·poly(dGdC) in 1 mM K_2HPO_4 (pH 6.8) at 25 °C. (A) Fluorescence quenching of Sac7d by poly(dGdC)·poly(dGdC) as a function of KCl concentration. (B) Salt dependence of the binding affinity of Sac7d for poly(dGdC)·poly(dGdC) determined by salt back-titration (●). Independent reverse titration measurements of K_{obs} at various salt concentrations are indicated (□) for comparison.

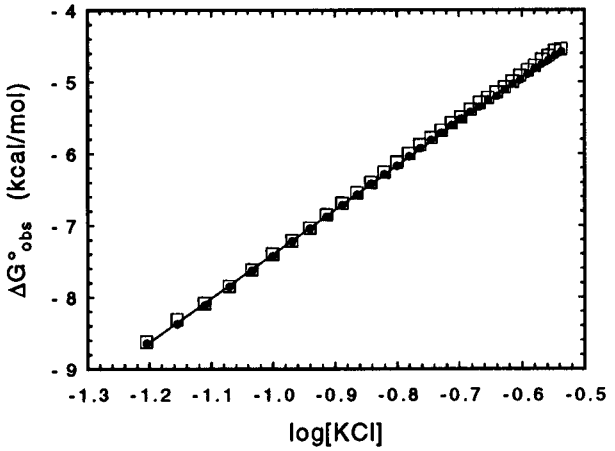

FIGURE 7: Dependence of the free energy of binding of native Sac7 (□) and recombinant Sac7d (●) to poly(dGdC)·poly(dGdC) as a function of KCl concentration in 1 mM K_2HPO_4 (pH 6.8) at 25 °C.

Table 3: Temperature Dependence of Sac7d Binding to Poly(dGdC)·Poly(dGdC)[a]

temp (°C)	$K \times 10^{-7}$ (M^{-1})	site size (bp)	Q_{max}
10	4.98 (±0.30)[b]	4.19 (±0.02)	0.851 (±0.002)
20	5.09 (±0.25)	4.05 (±0.03)	0.893 (±0.003)
30	4.99 (±0.08)	3.91 (±0.01)	0.907 (±0.001)
40	5.07 (±0.07)	3.74 (±0.04)	0.920 (±0.001)

[a] Binding was measured by reverse titrations using fluorescence in 10 mM KH_2PO_4 (pH 6.8). [b] See Footnote *b* in Table 1.

error with that observed for KCl (data not shown), implying no difference in Na^+ and K^+ affinities for the DNA.

Temperature Dependence of Binding. The binding of the recombinant Sac7d protein to poly(dGdC)·poly(dGdC) in 10 mM KH_2PO_4 (pH 6.8) was investigated at 10, 20, 30, and 40 °C using reverse titrations followed by fluorescence (higher temperatures were not accessible on our fluorimeter). K_{obs} was essentially constant over this range with an average of 5.03×10^7 M^{-1} (Table 3). The apparent experimental error in the mean ($\pm 0.06 \times 10^7$) is slightly smaller than the errors in the individual fits, which are on the order of 5% or less. The site size decreased from 4.2 to 3.8 base pairs with an increase in temperature from 10 to 40 °C, and Q_{max} increased from 0.85 to 0.92 over the same range, consistent with an increase in dynamics with temperature.

pH Dependence of Binding. The binding of the Sac7d protein to poly(dGdC)·poly(dGdC) was measured as a function of pH from 5.4 to 8.0 in 10 mM KH_2PO_4 by reverse titrations followed by fluorescence (not shown). The affinity showed only a slight decrease with pH over this range with $\partial \log K_{obs}/\partial \text{ pH} = -0.18$. There was negligible change in site size and Q_{max} from pH 6 to 8; the error in the site size over this range was 0.14 and in Q_{max} it was 0.003. The apparent affinity decreased dramatically below pH 5.4 due to aggregation upon addition of the DNA to the protein.

Overall Error Estimation. The errors in K_{obs}, site size, and Q_{max} obtained from fits of individual titrations do not include concentration measurement errors but are comparable to the errors obtained from repeated experiments under the same conditions. Actual experimental errors in the measurements of K_{obs} are less than ±10%, on the basis of the residuals of the fits of the dependence of K_{obs} on pH, temperature, and salt concentration. This is larger in most cases than the error of a single titration, which is approximately ±1.5%. The experimental error in the site size is approximately ±5%, estimated from the residuals in fits of the dependence on salt concentration and pH, or approximately 1 order of magnitude larger than the error from a single titration. The experimental error in Q_{max} is approximately ±0.003 on the basis of the data from the pH dependence, slightly larger than the error of a single titration.

Forward Titrations Followed by Circular Dichroism. Forward titrations of DNA with Sac7d and native Sac7 were monitored by circular dichroism. The CD spectrum of poly(dGdC)·poly(dGdC) is significantly altered as the nucleic acid is titrated with increasing amounts of recombinant Sac7d protein (Figure 8). Identical spectra were obtained for the native Sac7 protein (data not shown). The spectra are nearly isodichroic at 250 and 271 nm and suggest that there are two predominant DNA conformations.

The CD band at 220 nm becomes increasingly negative with increasing protein concentration in a *forward* titration

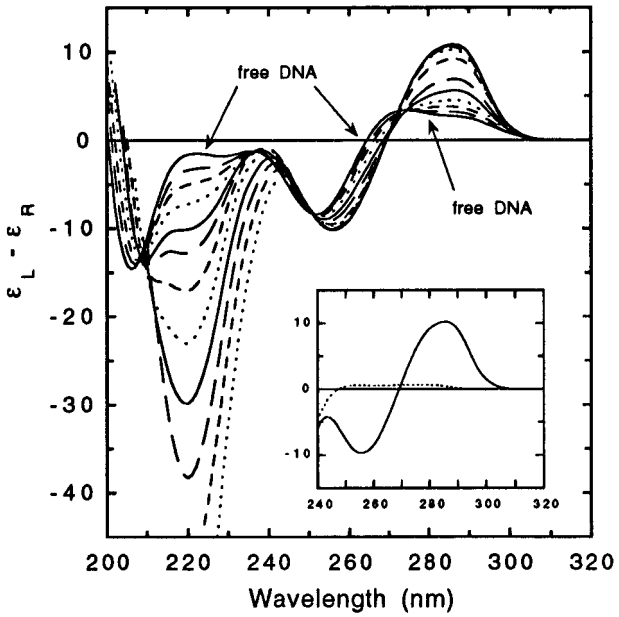

FIGURE 8: Forward titration of poly(dGdC)·poly(dGdC) with Sac7d followed by circular dichroism. The initial CD spectrum in the absence of protein (a protein/DNA ratio of 0) is labeled, and spectra for progressively increasing ratios of protein/DNA (0.01, 0.02, 0.03, 0.05, 0.07, 0.09, 0.12, 0.15, 0.19, 0.25, and 0.32) in 0.01 M KH_2PO_4 buffer (pH 7.0) are shown, scaled to the $\Delta\epsilon$ of the DNA. The initial concentration of poly(dGdC)·poly(dGdC) was 85 μM. The inset shows a comparison of the CD spectra of the Sac7d−DNA complex at the highest ratio (0.32) (—), and the free protein at the same concentration (···).

and is attributed to the Sac7d protein. At wavelengths above 250 nm, contributions from the DNA dominate the CD spectrum, and the signal change shows a plateau consistent with saturation at high protein concentrations. At the highest protein concentration used in Figure 8, the protein contribution to the CD at 280 nm is expected to be about 6% on the basis of the CD spectrum of the isolated protein. In addition, the CD spectra of the complexes in the long wavelength region are all similar to the CD spectra of the isolated DNAs (Figure 9); i.e. the DNA sequence dependent effects observed in the long wavelength region (see below) are consistent with the different base−base interactions present in repeating polynucleotides. We conclude that the CD changes in the long-wavelength region reflect conformational changes in the DNA induced by Sac7d protein binding.

Figure 9 shows the effects of Sac7d binding on the CD spectra of different polynucleotides where the CD spectrum of isolated protein has been subtracted from that of the DNA−protein complex (6/1 ratio). For all polynucleotides, the CD at long wavelengths increases upon complex formation, but significant sequence dependent effects are observed. The positive long wavelength band of poly(dAdT)·poly(dAdT) at 262 nm doubles in magnitude, and an additional CD band appears at about 290 nm. A similar increase is seen in the 275 nm band of poly(dAdG)·poly(dCdT). The CD of poly(dIdC)·poly(dIdC) increases over the entire spectral region, resulting in a reduction of the negative CD band observed at 283 nm for the B-form of this DNA sequence. The positive long wavelength bands of poly(dGdC)·poly(dGdC) fuse to form a single positive peak at 286 nm with about 3 times the magnitude of the B-form, and the negative CD band at 255 nm increases in magnitude. The spectral changes of poly(dAdC)·poly(dGdT) and poly-

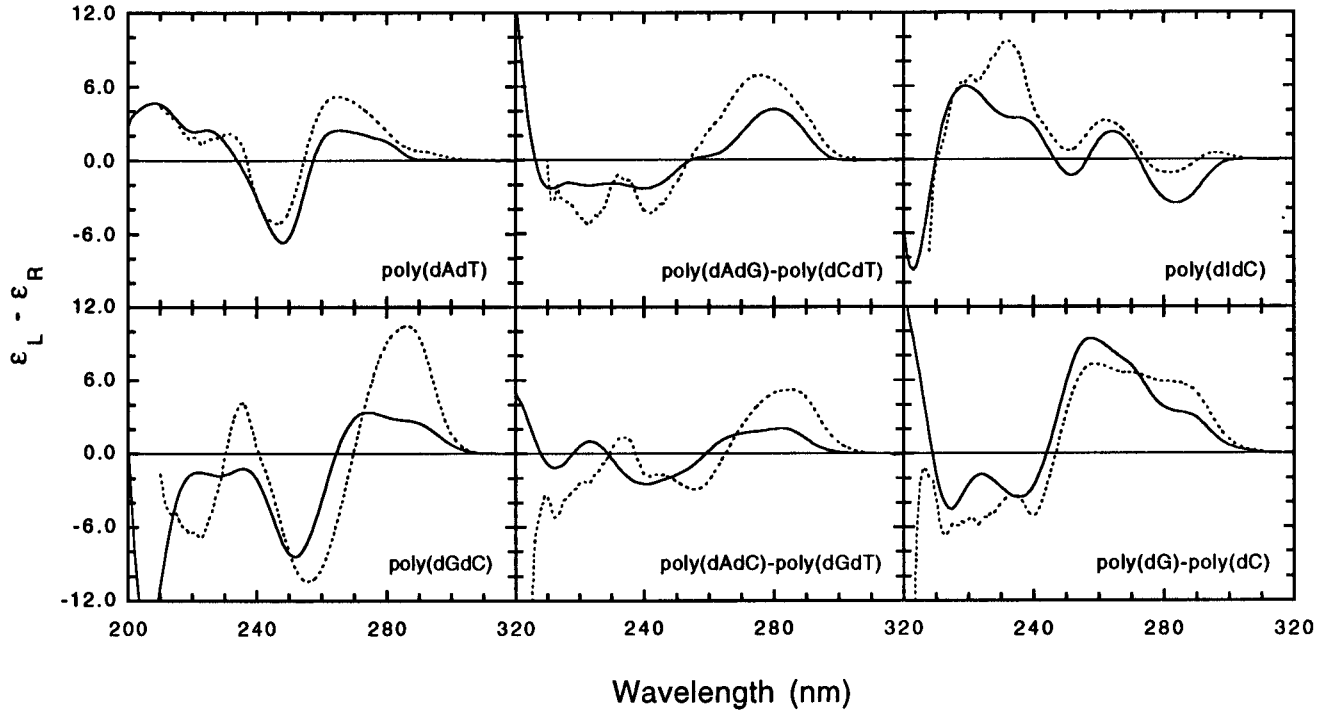

FIGURE 9: CD spectra of free, B-form DNA (—) and the DNA component of Sac7d complexes (···) calculated by subtraction of the CD contributions of Sac7d protein from the CD spectra of DNA−protein (6/1) complexes (in 0.01 M KH_2PO_4 buffer, pH 7.0).

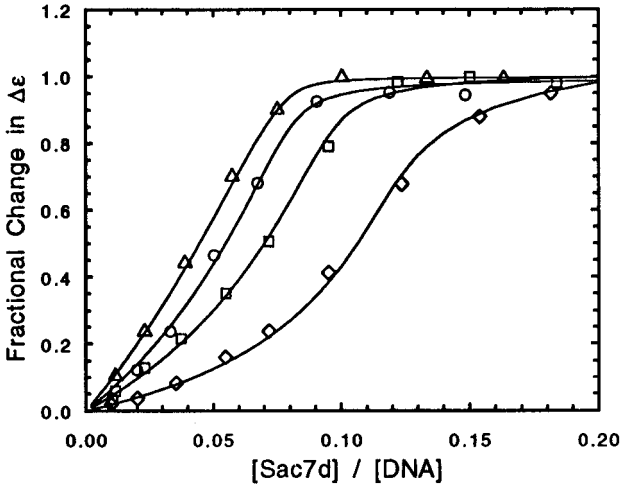

FIGURE 10: Fractional change in $\Delta\epsilon_{278}$ for poly(dAdG)·poly(dCdT) (△), and $\Delta\epsilon_{285}$ for poly(dAdC)·poly(dGdT) (○), poly(dGdC)·poly(dGdC) (□), and poly(dIdC)·poly(dIdC) (◇) with increasing concentrations of Sac7d (in 0.01 M KH_2PO_4 buffer, pH 7.0). The smooth curves through the data are simulations using the following parameters: poly(dIdC)·poly(dIdC), $n = 3.5$ base pairs, $K_{obs} = 1 \times 10^8$ M^{-1}, and $G = 0$ base pairs; poly(dGdC)·poly(dGdC), $n = 4.3$ base pairs, $K_{obs} = 5 \times 10^7$ M^{-1}, and $G = 3$ base pairs; poly(dAdC)·poly(dGdT), $n = 5$ base pairs, $K_{obs} = 5 \times 10^7$ M^{-1}, and $G = 5$ base pairs; poly(dAdG)·poly(dCdT), $n = 5$ base pairs, $K_{obs} = 5 \times 10^7$ M^{-1}, and $G = 8$ base pairs. Poly(dAdT)•poly(dAdT) (not shown) was fit with $K_{obs} = 1 \times 10^7$ M^{-1}, n = 4.3, and $G = 5$ base pairs.

(dG)·poly(dC) are similar to those of poly(dGdC)·poly(dGdC).

The intensity of the long wavelength CD band shows a sigmoidal dependence on protein concentration (Figure 10), indicating a cooperative transition in the DNA structure. No evidence of sigmoidicity was apparent in the intensity of the protein band at 220 nm. In addition, no sigmoidicity was noted in a forward titration of poly(dGdC)·poly(dGdC) with Sac7d monitored by fluorescence (data not shown). The sigmoidicity is therefore not dependent on the order of the addition of protein and DNA. We were unable to fit the CD data by assuming that the CD signal reflects the binding density using the cooperative McGhee−von Hippel model (McGhee & von Hippel, 1974), again indicating that the signal does not directly represent cooperative binding of protein. The sigmoidicity in the forward titrations followed by CD is attributed to a cooperative transition in the DNA structure which occurs with increasing protein binding density.

The degree of sigmoidicity in the CD titrations depends on the DNA sequence (Figure 10). Poly(dIdC)·poly(dIdC) exhibits the most sigmoidicity, and the CD approaches its limiting value at about 3.5 base pairs per protein molecule, which is the same as the binding site size determined by fluorescence. Poly(dAdG)·poly(dCdT), on the other hand, is barely cooperative, and its conformational change is completed at about 7 base pairs per protein, significantly larger than the 4.3 base pairs determined by fluorescence reverse titrations. The other polynucleotides studied are intermediate between these two extremes. The conformational change of poly(dGdC)·poly(dGdC) is completed at about 4.5 base pairs per protein. Apparent break points in titration curves of poly(dAdT)·poly(dAdT) and poly(dAdC)·poly(dGdT) occur at about 5−6 base pairs per protein.

DISCUSSION

The binding of Sac7d to DNA has been studied here by using intrinsic spectroscopic signals arising from both the protein and the DNA. Indirect spectroscopic measurements of binding can provide high quality data permitting precise measurements of binding parameters but suffer from the obvious drawback of not providing a direct measure of the bound (and, therefore, the free) ligand concentration. It cannot be assumed that the magnitude of the signal change is directly related to the fraction of ligand bound (Bujalowski

& Lohman, 1987). Multiple binding modes and interactions between neighboring bound ligands can lead to a nonlinear dependence of the signal change on the fraction of ligand bound. However, a graphical binding density function analysis permits a straightforward model independent determination of the free ligand concentration from indirect data and an investigation of the dependence of signal change on binding density (Bujalowski & Lohman, 1987). The ligand binding density function analysis of the fluorescence quenching data presented here demonstrates that the fractional change in the fluorescence quenching is equal to the fraction of protein bound (Figure 3), and if multiple binding modes exist, the fluorescence for each is identical. Therefore, the fluorescence signal is an ideal indicator of the amount of protein bound, and $L_b = Q_{obs}/Q_{max}L_t$.

The salt dependence of the DNA binding affinity of recombinant Sac7d indicates that the predominant interaction is electrostatic. The slope of the salt dependence, $\partial \log K/\partial \log[KCl]$, indicates that a net release of 4.4 cations is associated with the binding of Sac7d to poly(dGdC)·poly(dGdC). The absence of curvature in the log K vs log[KCl] plot suggests that anion release from the protein makes a negligible contribution to the binding free energy, at least above 50 mM salt. It appears that at least five phosphates are involved in electrostatic interactions with the Sac7d complex, comparable to that determined for the nonspecific interaction between DNA and the *lac* repressor headpiece (Schnarr et al., 1983).

If anion release from the protein is thermodynamically unimportant, we can extrapolate the binding free energy change to 1 M salt to obtain the free energy of interaction due to effects other than the polyelectrolyte effect. The remaining free energy is primarily due to the energy of charge pairs in the absence of ion release and nonelectrostatic interaction contributions (Record et al., 1976; de Haseth et al., 1977; Lohman & Mascotti, 1992):

$$\Delta G^\circ_{obs}(1\ M\ K^+) = \Delta G^\circ(\text{nonelec}) + z\Delta G^\circ(\text{ionic}) \quad (13)$$

where ΔG°(nonelec) includes H bonding, van der Waals, and hydrophobic interactions, and $z\Delta G^\circ$(ionic) is the energy of interaction of z charge pairs between the protein and nucleic acid. We observe an extrapolated binding free energy at 1 M KCl of −1.2 kcal/mol. From work with oligolysines, ΔG°(ionic) has been estimated to be +0.2 kcal/mol ionic interaction in duplex DNA (Record et al., 1976; Lohman et al., 1980). If there are five ion pairs, the total ionic interaction energy is +1.0 kcal/mol, and ΔG°(nonelec) is predicted to be −2.2 kcal/mol, or approximately 30% of the binding free energy at 0.1 M KCl (pH 6.8). A large part of this favorable free energy may by due to interactions between the DNA bases and Trp 24, which lies on the surface of the three-stranded β-sheet (Edmondson et al., 1995) and is believed to interact nonspecifically with bases in the major groove (Baumann et al., 1995). Trp 24 is conserved in all of the Sac7 and Sso7 proteins and would appear to be important for the function of the protein.

The binding affinity of the Sac7d protein for DNA is temperature independent over the range of study. A van't Hoff analysis provides a ΔH_{obs} of −0.06 kcal/mol, or zero within experimental error. The binding of Sac7d is therefore entropically driven, which is expected for protein binding driven predominantly by the polyelectrolyte effect (Record, 1988). Similar observations have been made with polyamines and the λ repressor headpiece (Lohman & Mascotti, 1992). The absence of any curvature in the temperature dependence of ΔH_{obs} implies a negligible change in heat capacity for binding. Thus, an extrapolation of the measured binding affinity to the growth temperature can be made with reasonable precision, and the values for K, ΔG_{obs}, ΔH_{obs}, and ΔS_{obs} measured here can be expected to be reasonably accurate at 75 °C. An explanation of the slight temperature dependence of the site size will need to await a structure of the complex. Although the fundamental interactions between DNA and protein may be temperature independent, the effective site size may reflect a partial occlusion of neighboring DNA sites by a flexible, nonbinding loop that becomes less structured with increasing temperature, which would allow an effective decrease in site size.

The large fluorescence quenching observed upon Sac7d binding to DNA as well as the slight temperature dependence are difficult to explain quantitatively. For the sake of comparison, it is interesting to note that the binding of autoimmune anti-DNA antibodies to single-stranded DNA is often associated with a significant quenching of intrinsic fluorescence [e.g. 62% in the case of HEd10 (Lee et al., 1982)], while there is negligible change upon binding to double-stranded DNA (Bruan & Lee, 1987). The structure of a complex of the autoantibody BV04−01 with single-stranded DNA shows intercalation of a thymine base between tryptophan and tyrosine rings (Herron et al., 1991). Of course, such an interaction is not likely in a complex with double-stranded DNA, nor is it required for significant quenching of fluorescence as witnessed by the 64% quenching observed upon nonspecific binding of *lac* repressor headpiece to DNA (Schnarr et al., 1983).

There are no significant differences between K_{obs}, site size, Q_{max}, and salt dependence (and therefore apparent number of cations released and electrostatic and nonelectrostatic contributions to the binding free energy) of the native Sac7 and recombinant Sac7d proteins for poly(dGdC)·poly(dGdC) (Figure 7). The CD spectra of the native and recombinant protein complexes with DNA are also identical. This extends our earlier report of similar properties for native and recombinant proteins (McAfee et al., 1995) and indicates that the folding of the native and recombinant Sac7 proteins are identical, at least in the vicinity of the DNA binding site. We conclude that lysine monomethylation has no effect on the DNA binding function of Sac7. Lysine methylation appears to be important primarily in increasing the stability of the protein (McAfee et al., 1995).

The binding site sizes reported here for Sac7d are in the middle of the range estimated by Baumann et al. (1994) for the homologous native Sso7 proteins. Under tight binding conditions (20 mM TRIS/HCl, pH 7.4), the site size for Sso7 binding to poly(dGdC)·poly(dGdC) was reported to be approximately 6 base pairs with a conservative estimate extending down to 3 base pairs. Some overestimate on their part can be attributed to the neglect of the overlap of sites that must occur with protein binding to linear DNA lattices (McGhee & von Hippel, 1974). Site sizes cannot be derived directly from apparent saturation points in titration curves under stoichiometric binding conditions (i.e. tight binding) unless the binding is moderately to highly cooperative. The *relative* binding affinities reported here for various polynucleotides are in good agreement with the estimates

presented by Baumann et al. (1994) for Sso7.

Apparent saturation of the DNA observed in the CD forward titration data occurs at a protein/DNA ratio equal to or slightly less than that determined by fluorescence reverse titrations. Complete titration of the protein essentially coincides with complete titration of the DNA. This demonstrates that Sac7d binds to vacant sites on the DNA and not to other Sac7d molecules. This would appear to differ from the electron microscopy results of Lurz et al. (1986) which showed that Sac7d aggregates or conglomerates in localized regions on DNA. Formation of aggregates would dramatically decrease the apparent site size determined by CD forward titrations relative to that determined by fluorescence reverse titrations. Noncooperative binding also indicates that the protein binds evenly and randomly along the length of DNA.

The binding density function analysis presented here demonstrates that the percent change in fluorescence quenching accurately reflects the amount of protein bound. The same binding parameters (K, n, and ω) must hold whether the titrations are followed by fluorescence or CD. We conclude that the CD data indicates that binding of the protein perturbs the DNA structure beyond the actual number of bases occluded by binding.

The susceptibility of the DNA to the protein-induced conformational change does not appear to be related to the binding affinity and probably reflects the sequence dependent conformational preferences of DNA. That is, some DNA sequences are more facile in undergoing the conformational switch than others. For poly(dIdC)·poly(dIdC), the maximum conformational change occurs when the DNA is nearly saturated with protein. At low binding densities, Sac7d has less of an effect than would be expected given the amount of protein bound. Other polynucleotides, however, are more easily affected, and less bound protein is needed to induce the maximal structural change. For these polynucleotides, the apparent site size observed by CD is greater than the binding site size determined by fluorescence.

The sigmoidal CD forward titration data can be quantitatively explained by assuming that a local structural transition in the DNA cannot be induced by a single bound protein molecule but requires mutual reinforcement of at least two proteins bound within a specified distance. A fit of the CD data requires an expression describing the relative separations of protein molecules on the DNA as a function of the protein concentration. McGhee and von Hippel (1974) have shown that the probability of a gap of g residues between two bound proteins on a one dimensional lattice is given by

$$P_g = \left[\frac{1 - n\nu}{1 - (n-1)\nu}\right]^g \left[\frac{\nu}{1 - (n-1)\nu}\right] \qquad (14)$$

The probability of two proteins being bound within G base pairs of each other is equivalent to that of observing a gap of G residues or less and is given by the sum of the probabilities for gaps from O to G

$$\sum_0^G P_g = P_0 + P_1 + ... + P_G = \Pi_G \qquad (15)$$

Using a simple model where the fractional change in CD signal is proportional to $\sum_g$, sigmoidal CD titrations are predicted for small values of G, and the degree of sigmoidicity decreases with increasing G. The CD data can be adequately fit with values for K_{obs}, n, and G which are in good agreement with the values obtained from fluorescence reverse titrations (Figure 10). The poly(dIdC)·poly(dIdC) data is best fit with a gap size of zero; i.e. the DNA structural transition can only occur if two proteins are immediately adjacent. The other DNAs are more easily affected by the bound protein, with the structural transition in poly(dAdG)·poly(dCdT) occurring when protein molecules are bound within 8 base pairs. Note that for nucleic acids requiring a larger G value, such as poly(dAdG)·poly(dCdT), a good fit requires a site size somewhat larger than that obtained by fluorescence. This is most likely due to the neglect of the possibility of a finite probability of inducing a structural transition in the DNAs by a single protein molecule or the mutual interactions between more than two bound proteins.

The long-wavelength CD bands of DNA are very sensitive to conformation and may indicate the nature of the structural transition induced in the DNA by Sac7d. Theoretical calculations indicate that the CD at 275 nm is most sensitive to the helix winding angle and the base pair propeller twist (Johnson et al., 1981), and a correlation with the helical twist (or winding) of duplex DNA has been established experimentally (Baase & Johnson, 1979). Increasing ionic strength and decreasing temperature increase the helical twist of DNA in solution and result in a decrease of the CD at 275 nm (Gennis & Cantor, 1972; Zimmer & Luck, 1974). Similar CD changes are produced by nucleosome formation (Cowman & Fasman, 1978) and the packaging of DNA in bacteriophages (Maestre et al., 1971).

Sac7d binding results in an increase in the long wavelength CD bands of DNA. Similar spectral changes are observed for duplex DNA with increasing temperature prior to the start of the melting transition (the premelting phenomenon), reflecting a uniform decrease in helical twist as the temperature increases (Sprecher & Johnson, 1982). Assuming the same relationship between $\Delta\epsilon_{275}$ and the winding angle as observed by Baase and Johnson (1979), we calculate an unwinding of about $1.4 \pm 0.4°$ per base pair for saturation of synthetic polynucleotides with Sac7d protein. There is little change in the hypochromicity of DNA upon complex formation, and thus, base stacking is not significantly disrupted. Further, Sac7d does not facilitate the strand separation of DNA but rather stabilizes the duplex form by about 40 °C (McAfee et al., 1995).

The CD changes we observe upon complex formation may reflect more than unwinding of duplex DNA. In particular, the red shift in the long-wavelength CD band of poly(dGdC)·poly(dGdC) suggests a much greater perturbation for this DNA sequence upon binding Sac7. The CD maximum of the complex at 283 nm is reminiscent of the formation of hemiprotonated C−C$^+$ base pairs observed for cytosine-containing polymers in acidic solutions (Gray et al., 1988). It seems unlikely that Sac7d could promote C−C$^+$ base pairing in poly(dGdC)·poly(dGdC), but the conformational change produced by Sac7d may enhance base−base interactions in poly(dGdC)·poly(dGdC) and thereby have a similar effect on the optical properties. Difference spectra (not shown) indicate that all the DNA sequences exhibit an enhancement of a CD band at 285-290 nm, although the enhancement is weakest for poly(dAdT)·poly(dAdT) and poly(dAdG)·poly(dCdT).

The sequence specific DNA-binding proteins *gal* repressor, *tet* repressor, LexA repressor, and *lac* repressor cause increases in the long wavelength CD of their operator DNAs, similar to that found for Sac7d (Altschmied & Hillen, 1984; Wartell & Adhya, 1988; Culord & Maurizot, 1981). The amino acid sequence homologies among these repressors suggest similar helix−turn−helix protein structures, and they probably induce similar structural changes in their operator DNAs upon binding. The DNA methyltransferase protein M.*Eco*R124I, although it does not contain a helix−turn−helix motif, has a similar effect on its specific target DNA sequence (Taylor et al., 1994). Except for LexA and *lac* repressor (see below), these proteins have no effect on DNA sequences that differ from their target sequence.

LexA and *lac* repressor also exhibit nonspecific binding to DNA. LexA increases the positive CD band of the "pseudo-operator" poly(dAdT)·poly(dAdT) by a factor of about 2 and induces the appearance of an additional CD band at 290 nm (Schnarr et al., 1983), almost identical to that observed for Sac7d. With random sequence DNA, LexA causes a 3-fold increase in the CD with no change in the spectral shape (Hurstel et al., 1986). Nonspecific binding of *lac* repressor is similar, increasing the positive CD bands of poly(dAdT)·poly(dAdT) by a factor of 2 and poly(dGdC)·poly(dGdC) by a factor of 3 (Durand & Maurizot, 1980). The NMR structure of the *lac* repressor complex indicates that the *lac* repressor causes a significant bending (about 25° on the basis of the PDB coordinates) of the operator DNA (Chuprina et al., 1993).

Sac7d differs from the other DNA-binding proteins discussed above for which there are CD spectra of the protein−DNA complex. Sac7d has an architecture different from the helix−turn−helix motif of the repressors and appears to bind DNA through its β-sheet (Baumann et al., 1995). Sac7d exhibits little sequence specificity, and we are not aware of any other nonspecific DNA-binding protein that causes a large increase in the long-wavelength CD bands of duplex DNA. Except for the interaction of LexA with poly(dAdT)·poly(dAdT), the sequence specific proteins have little effect on the shape of the CD bands. This is in contrast to the effect of Sac7d binding to poly(dGdC)·poly(dGdC), where there is a significant red shift of the CD band to 283 nm. A detailed understanding of the structural transition induced in DNA by Sac7d binding must await a determination of the structure of the complex.

The results presented here can be compared to those for the predominantly β-sheet autoimmune anti-double-stranded DNA antibodies which bind to DNA with essentially no sequence specificity (Eilat & Anderson, 1994). Interestingly, the binding of the murine monoclonal anti-DNA antibody (BV17−45) which binds preferentially to double-stranded DNA shows negligible salt dependence (Ballard et al., 1984). Stollar et al. (1986) described another murine autoantibody (H241) with a DNA binding site size of approximately 4 base pairs. The binding data appear to indicate that the protein binds across the sugar−phosphate backbone with contacts in both the major and minor grooves. H241 shows only a slight preference for double-stranded over single-stranded DNA. Braun and Lee (1987) presented data which indicates that the Fab fragment of another anti-double-stranded DNA antibody (Jel 241) binds with a site size of 6 base pairs and an intrinsic association constant of 1.5×10^6 M^{-1} in 50 mM NaCl. The preference for double stranded DNA is significant as noticed by the thermal stabilization of duplex DNA. The binding of Jel 241 is strongly salt dependent and indicates that four phosphates are involved in binding. Comparison of binding affinites for various nucleic acids implied that Jel 241 binds across the major groove (Braun & Lee, 1986). Unfortunately, there is no detailed structural information on this or any other autoimmune DNA−antibody complex.

REFERENCES

Altschmied, L., & Hillen, W. (1984) *Nucleic Acids Res. 12*, 2171−2180.

Baase, W. A., & Johnson, W. C. J. (1979) *Nucleic Acids Res. 6*, 797−814.

Ballard, D., Lynn, S., Gardner, J., & Voss, E. W., Jr. (1984) *J. Biol. Chem. 259*, 3492−3498.

Baumann, H., Knapp, S., Karshikoff, A., Ladenstein, R., & Härd, T. (1995) *J. Mol. Biol. 247*, 840−846.

Bevington, P. R., & Robinson, D. K. (1992) *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill, New York.

Braun, R., & Lee, J. S. (1986) *Nucleic Acids Res. 14*, 5049−5065.

Braun, R., & Lee, J. S. (1987) *J. Immunology 139*, 175−179.

Bujalowski, W., & Lohman, T. (1987) *Biochemistry 26*, 3099−3106.

Choli, T., Henning, P., Wittmann-Liebold, B., & Reinhardt, R. (1988a) *Biochim. Biophys. Acta 950*, 193−203.

Choli, T., Wittmann-Liebold, B., & Reinhardt, R. (1988b) *J. Biol. Chem. 263*, 7087−7093.

Chuprina, V. P., Rullmann, J. A. C., Lamerichs, R. M. J. N., van Boom, J. H., Boelens, R., & Kaptein, R. (1993) *J. Mol. Biol. 234*, 446−462.

Cowman, M. K., & Fasman, G. D. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 4759−4763.

Culard, F., & Maurizot, J. C. (1981) *Nucleic Acids Res. 9*, 5175−5184.

de Haseth, P., Lohman, T., & Record, M. T., Jr. (1977) *Biochemistry 16*, 4783−4790.

Dijk, J., & Reinhardt, R. (1986) in *Bacterial Chromatin* (Gualerzi, C. O., & Pon, C. O., Eds.) Springer-Verlag, Berlin.

Durand, M., & Maurizot, J.-C. (1980) *Biochimie 62*, 503−507.

Edmondson, S. P., Qiu, L., & Shriver, J. W. (1995) *Biochemistry 34*, 13289−13304.

Eilat, D., & Anderson, W. F. (1994) *Mol. Immunol. 31*, 1377−1390.

Gennis, R. B., & Cantor, C. R. (1972) *J. Mol. Biol. 65*, 381−399.

Gray, D. M., Ratliff, R. L., Antao, V. A., & Gray, C. W. (1988) in *Structure and Expression, Volume 2*: *DNA and Its Drug Complexes* (Sarma, R. H., & Sarma, M. H., Eds.) Adenine Press, Guilderland, NY.

Grote, M., Dijk, J., & Reinhardt, R. (1986) *Biochim. Biophys. Acta 873*, 405−413.

Herron, J. N., He, X. M., Ballard, D., Blier, P., Pace, P., Bothwell, A., Voss, E., Jr., & Edmundson, A. (1991) *Proteins 11*, 159−175.

Hurstel, S., Granger-Schnarr, M., Daune, M., & Schnarr, M. (1986) *EMBO J. 5*, 793−798.

Johnson, B. B., Dahl, K. S., Tinoco, J. I., Ivanov, V. I., & Zhurkin, V. B. (1981) *Biochemistry 20*, 73−78.

Kamath, U., & Shriver, J. (1989) *J. Biol. Chem. 264*, 5586−5592.

Kimura, M., Kumura, J., Phillips, D., Reinhardt, R., & Dijk, J. (1984) *FEBS Lett. 176*, 176−178.

Kowalczykowski, S., Paul, L., Lonberg, N., Newport, J., McSwiggen, J., & von Hippel, P. (1986) *Biochemistry 25*, 1226−1240.

Lee, J., Dombroski, D., & Mosmann, T. (1982) *Biochemistry 21*, 4940−4945.

Lohman, T., & Mascotti, D. (1992) *Methods Enzymol. 212*, 400−424.

Lohman, T., deHaseth, P. H., & Record, M. T., Jr. (1980) *Biochemistry 19*, 3522.

Lurz, R., Grote, M., Dijk, J., Reinhardt, R., & Dobrinski, B. (1986) *EMBO J. 5*, 3715−3721.

Maestre, M. F., Gray, D. M., & Cook, R. B. (1971) *Biopolymers 10*, 2537−2553.
McAfee, J. G., Edmondson, S. P., Datta, P. K., Shriver, J. W., & Gupta, R. (1995) *Biochemistry 34*, 10063−10077.
McGhee, J., & von Hippel, P. (1974) *J. Mol. Biol. 86*, 469−489.
Miller, A. (1987) *Turbo BASIC Programs for Scientists and Engineers*, SYBEX, San Francisco.
Press, W., Flannery, B., Teukolsky, S., & Vetterling, W. (1989) *Numerical Recipes: The Art of Scientific Computing (Fortran Version)*, Cambridge University Press, Cambridge, U.K.
Record, M. T., Jr. (1988) in *Unusual DNA Structures* (Wells, R. D., & Harvey, S. C., Eds.) Springer-Verlag, New York.
Record, M. T., Jr., Lohman, T., & de Haseth, P. (1976) *J. Mol. Biol. 107*, 145−158.
Reddy, T. R., & Suryanarayana, T. (1988) *Biochim. Biophys. Acta 949*, 87−96.
Reddy, T. R., & Suryanarayana, T. (1989) *J. Biol. Chem. 264*, 17298−17308.
Schnarr, M., Durand, M., & Maurizot, J.-C. (1983) *Biochemistry 22*, 3563−3570.
Sprecher, C. A., & Johnson, W. C. J. (1982) *Biopolymers 21*, 321−329.
Stetter, K. O., Fiala, G., Huber, G., Huber, R., & Segerer, A. (1990) *FEMS Microbiol. Rev. 75*, 117−124.
Stollar, B. D., Zon, G., & Pastor, R. W. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 4469−4473.
Straume, M., & Johnson, M. (1992) *Methods Enzymol. 210*, 117−129.
Taylor, I. A., Davis, K. G., Watts, D., & Kneale, G. G. (1994) *EMBO J. 13*, 5772−5778.
Wartell, R. M., & Adhya, S. (1988) *Nucleci Acids Res. 16*, 11531−11541.
Zimmer, C., & Luck, G. (1974) *Biochim. Biophys. Acta 361*, 11−32.

BI952555Q