

Thermodynamic and Kinetic Methods of Analyses of Protein–Nucleic Acid Interactions. From Simpler to More Complex Systems

Włodzimierz Bujalowski*

Department of Biochemistry and Molecular Biology, Department of Obstetrics and Gynecology, the Sealy Center for Structural Biology, and the Sealy Center for Cancer Cell Biology, The University of Texas Medical Branch at Galveston, 301 University Boulevard, Galveston, Texas 77555-1053

Received March 30, 2005

Contents

1. Introduction	556	3.3.3. Correlation between the Fractional Signal Change, ΔS_{obs} , and the Average Degree of Binding, $\Sigma \Theta_i$	579
2. Direct versus Indirect Methods in Studying Protein–Nucleic Acid Interactions	557	3.3.4. Analysis of a Binding Isotherm Using a Single Titration Curve When $\Delta S_{\text{obs}}/\Delta S_{\text{max}} = L_B/L_T$	579
3. Thermodynamic Bases of Quantitative Equilibrium Spectroscopic Titrations	558	3.4. Empirical Function Approach	580
3.1. The Spectroscopic Signal Used to Monitor Protein–Nucleic Acid Interactions Originates from the Nucleic Acid (Macromolecule)	559	3.5. Stopped-Flow Kinetic Studies of Protein–Nucleic Acid Interactions	582
3.1.1. Analysis of Spectroscopic Titration Curves When the Signal Originates from the Nucleic Acid. The Case of a Single Binding Site	559	3.5.1. General Analysis of Relaxation Times and Amplitudes Using the Matrix Projection Operator Technique	582
3.1.2. The Case of Two Binding Sites	561	3.5.2. Protein–Nucleic Acid System with Slow Bimolecular Step. Kinetics of the ssDNA 20-mer Binding to the DnaB Helicase.	586
3.1.3. Two Different Binding Modes of the Protein–Nucleic Acid Complex	563	3.5.3. Protein–Nucleic Acid System with Fast Bimolecular Step with Undetectable Amplitude. Kinetics of PriA Helicase–ssDNA Interactions	590
3.1.4. Protein Binding to a One-Dimensional Homogeneous Lattice in a Single Binding Mode	566	3.5.4. Protein–Nucleic Acid System with Fast Bimolecular Step Having Detectable Amplitude. Kinetics of the Human Pol β Binding to the ssDNA	595
3.1.5. Protein Binding to a One-Dimensional Homogeneous Lattice in Two Binding Modes Differing in the Number of Occluded Nucleotides	567	4. Summary	604
3.2. Macromolecular Competition Titration Method (MCT)	571	5. Acknowledgment	604
3.2.1. Thermodynamic Bases	571	6. References	605
3.2.2. Specific Application of the MCT Method	573		
3.2.3. Direct Analysis of the Experimental Isotherm of Protein Ligand Binding to Two Competing Nucleic Acid Lattices	574		
3.2.4. Competition Titration Method Using a Single Concentration of a Nonfluorescent Nucleic Acid	575		
3.2.5. Competition Fluorescence Titrations of Short Fluorescent Oligonucleotides in the Presence of a Competing Polymer Nucleic Acid	576		
3.3. Signal Used to Monitor the Interactions Originates from the Ligand	576		
3.3.1. Thermodynamic Bases	576		
3.3.2. Analysis of Spectroscopic Titration Curves When the Signal Used to Monitor the Interactions Originates from the Ligand	577		

1. Introduction

Thermodynamic and kinetic studies of the protein–nucleic acid interactions provide information that is required to elucidate the magnitude and nature of forces that drive the formation of the complexes, the specificity of the interactions, and the role of key intermediates of the reactions.^{1–5} Knowledge of energetics and mechanisms of the interactions is also indispensable in revealing and characterizing functionally relevant structural changes of both the protein and the nucleic acid that accompany the formation of the studied complexes. Quantitative analyses are designed to provide the answers to such fundamental questions as: What is the stoichiometry of the formed complexes? How strong or how specific are the interactions? Are the binding sites intrinsically heterogeneous? Are there any cooperative interactions among the binding sites, the bound ligand molecules, or both? What is the mechanism of the complex formation? How many intermediates are involved in the reaction? Is any intermediate(s) dominating the energetics, the population distribution, or both of the formed complex? What are the effective molecular forces involved in the formation of the studied complexes, or in other words, how do the equilibrium

* To whom correspondence should be addressed. Mailing address: Department of Biochemistry and Molecular Biology, The University of Texas Medical Branch at Galveston, 301 University Boulevard, Galveston, TX 77555-1053. Tel: (409) 772-5634. Fax: (409) 772-1790. E-mail: wbujaalow@utmb.edu.



Dr. Włodzimierz Bujalowski received his Ph.D. in 1982 from Poznan University, Poland, in Physical Biochemistry of nucleic acid structure and protein–nucleic acid interactions. After a year as a visiting scientist in the Department of Chemistry at the State University of New York, Stony Brook, in the laboratory of Dr. Benjamin Chu, he undertook postdoctoral studies in the laboratory of Nobel laureate Dr. Manfred Eigen and Dr. Dietmar Porschke at the Max Plank Institute for Biophysical Chemistry in Göttingen, Germany, where he was involved in chemical relaxation studies of the dynamics of the RNA structure and codon–anticodon recognition processes. In 1986, he joined the group of Dr. Timothy Lohman at Texas A&M University in College Station, where his work concentrated on the statistical mechanical analyses of the single-strand-binding protein interactions with the DNA and development of general quantitative methods to examine multiple ligand–macromolecule binding processes using optical spectroscopic methods. In 1990, Dr. Bujalowski joined the Faculty of the University of Texas Medical Branch at Galveston, where he currently is a professor in the Department of Biochemistry and Molecular Biology, Department of Obstetrics and Gynecology, and senior scientist in the Sealy Center for Structural Biology and Sealy Center for Cancer Cell Biology. His research concentrates on the quantitative understanding of the structure–function relationships in protein–nucleic acid interactions in solution and the mechanism of the free energy transduction in these complexes. The particular interest of his group is directed toward elucidation of the molecular mechanism of substrate recognition and catalysis by the motor proteins of DNA metabolism, DNA helicases and polymerases. Part of his work is devoted toward development of quantitative methods to study thermodynamics, kinetics, and structure of macromolecular complexes in solution using spectroscopic techniques. The methods include relaxation kinetics, time-dependent fluorescence spectroscopic methods, static and dynamic fluorescence energy transfer approaches, applications of analytical ultracentrifugation, and dynamic light scattering methods.

binding and kinetic parameters depend on solution variables (temperature, pressure, pH, salt concentration, type of salt, etc.)?

Binding isotherms of a protein association with a nucleic acid or, in general, of ligand binding to a macromolecule represent a direct relationship between the degree of binding (moles of ligands bound per mole of a macromolecule) and the free ligand concentration.^{6–10} Analogously, in the case of the protein (ligand) binding to a long one-dimensional nucleic acid lattice, the equilibrium binding isotherm represents the direct relationship between the binding density (moles of ligand bound per mole of bases or base pairs) and the free protein concentration.^{9–16} *A true thermodynamic binding isotherm is model-independent and reflects only this relationship.* Only then, when such a relationship is available, can one proceed to extract physically meaningful interaction parameters that characterize the examined interacting systems. This is accomplished by comparing the experimental isotherms to theoretical predictions based on specific statistical thermodynamic models that incorporate known molecular aspects of the system, such as intrinsic binding constants,

cooperativity parameters, allosteric equilibrium constants, discrete character of the binding sites or overlap of potential binding sites, etc. Only then, can one make rational molecular interpretations of the nature and mechanisms of the observed phenomena.

In the first part of this review, we describe thermodynamically rigorous analyses of spectroscopic approaches to study protein–nucleic acid interactions, including multiple-ligand binding phenomena, to obtain model-independent binding isotherms. In the second part, we discuss analyses of spectroscopic stopped-flow kinetic approaches used to examine mechanisms of protein–nucleic acid interactions with emphasis on the properties of the reaction intermediates available from amplitude analysis of the observed relaxation processes. Our discussion focuses on the fundamental problem of obtaining thermodynamic, spectroscopic, and kinetic parameters free of assumptions about the relationship between the observed signal and the degree of protein or nucleic acid saturation. The molecular interpretation of these parameters is treated only in an auxiliary manner with understanding that any further conclusions about any system are utterly dependent upon the quantitative determination of the thermodynamic and kinetic parameters, on which such conclusions are based. Although, in binding or kinetic studies, one generally monitors some spectroscopic signal (absorbance, fluorescence, fluorescence anisotropy, circular dichroism, NMR line width, chemical shift, etc.) from either the protein or the nucleic acid that changes upon complex formation, we will discuss the quantitative analyses as applied to the use of fluorescence as the most widely used spectroscopic technique.^{17–34} However, the derived relationships are general and applicable to any physicochemical signal used to monitor the ligand–macromolecule interactions.

The discussed approaches allow the experimenter to use a spectroscopic signal to obtain a thermodynamic binding isotherm even when direct proportionality between the monitored spectroscopic signal and the degree of binding does not exist.^{5,6,9,10,35–50} The only constraint on these analyses is that they are not valid if the ligand or the macromolecule undergoes an aggregation of self-assembly process within the experimental concentration ranges used. Therefore, as in any case, knowledge of the self-assembly properties of the ligand and the macromolecule under study are essential before any rigorous analysis of a binding process can be undertaken. The discussed specific cases are selected from the works performed in our laboratory because we can provide unique and hands-on insights into how these data were obtained and analyzed, thus achieving the best illustration of the applicability and feasibility of the described methods. The analyses include such diverse systems as the *Escherichia coli* DnaB hexameric helicase, the *E. coli* plasmidic PriA helicase, the RepA hexameric helicase of plasmid RSF1010, the isolated 8-kDa domain of the rat polymerase β , and the intact rat and human polymerases β .

2. Direct versus Indirect Methods in Studying Protein–Nucleic Acid Interactions

As pointed out above, any method used to analyze ligand binding to a macromolecule must relate the extent of the complex formation to the free ligand concentration in solution. Numerous techniques have been developed or applied to study equilibrium properties of specific and nonspecific protein–nucleic acid interactions, in which

binding is directly monitored, including analytical ultracentrifugation, column chromatography, affinity chromatography, filter binding assays, and gel electrophoresis.^{19,51–61} These *direct* methods are straightforward; however, they are usually time-consuming and some, like filter binding or gel shift assays, are, in general, nonequilibrium techniques, which require many controls before reliable equilibrium binding data can be obtained.^{51,53} Therefore, these direct methods are usually applied to suitable systems or where *indirect* spectroscopic approaches cannot be used due to the lack of an adequate spectroscopic signal change accompanying the formation of the complex.

In the case of indirect methods, the binding of the ligand to a macromolecule is determined by monitoring changes in a physicochemical parameter of the macromolecule–ligand system accompanying the formation of the complex. For the protein–nucleic acid systems, the common techniques used include, for example, absorbance, circular dichroism, NMR chemical shift, and most often, fluorescence intensity.^{62–69} The change in the physicochemical parameter is then correlated with the concentration of the free and bound ligand or with the fractional saturation of the macromolecule. The major advantage of using spectroscopic measurements is that they can be performed without perturbing the examined equilibrium, are available in most laboratories, and are relatively easy to apply.

However, in indirect methods the functional relationship between the observed spectroscopic signal and the degree of binding or binding density is never a priori known.^{6,9,10,35–50} Often, possibly too often, in analyses of spectroscopic titrations, it is assumed that a linear relationship exists between the fractional change of the monitored spectroscopic signal and the fractional saturation of the ligand or the macromolecule. Although this is true for the systems where, at saturation, only a single ligand molecule binds to the macromolecule, in the general case, in which multiple ligand molecules can participate in the binding process, the observed fractional change of the spectroscopic signal and the extent of binding may not and most frequently will not have such a simple linear relationship.^{6,10,35–50}

There are several reasons why the observed spectroscopic signal may not change in a linear fashion as a function of the degree of binding or fractional saturation of a macromolecule or a ligand. For instance, this may occur if there are structurally or functionally different sites on the macromolecule, each possessing different spectroscopic properties or differently affecting spectroscopic properties of the ligand. If there are cooperative interactions, they may affect the physical state of the bound ligand molecules or binding sites and the density of cooperative interactions may not be a linear function of the extent of the degree of binding. A ligand may bind to the macromolecule forming different binding modes characterized by different spectroscopic responses of the monitored spectroscopic signal accompanying the formation of each individual binding mode.^{28,43,44,47,48,70} More complex situations may include different combinations of all mentioned above cases.

Because the extent of any deviation from the assumed ideal linear behavior of the spectroscopic signal is unknown a priori, the magnitude of error introduced into the isotherm and the resulting binding parameters is also unknown. In other words, if a binding isotherm is obtained by indirect methods that involve a linear assumption about the relationship between the observed spectroscopic signal and the extent

of ligand binding or macromolecule saturation, then the interaction parameters that one obtains will be no more accurate than the accepted assumption. This may cause particular problems if the titration curve is being used to differentiate between alternative models for the interaction, since if one model does or does not “fit” the titration curve, this may be due either to the failure of the model or to the failure of the assumptions on which the calculated curve is based. Therefore, it is imperative that in quantitative studies a thermodynamically rigorous binding isotherm is obtained, independent of any assumptions as to the relationship between the degree of binding or the fractional saturation of the macromolecule and the observed spectroscopic signal.^{5,6,10,35–50} Determination of a model-independent, thermodynamic isotherm constitutes the first step in a correct analysis of energetics of protein–nucleic interactions or, to that matter, of any ligand–macromolecule system. In these analyses, the entity that binds multiple molecules of another chemical species is treated as a macromolecule. In the context of the protein–nucleic acid complexes, the nucleic acid is usually, although not exclusively, treated as the macromolecule, while the protein is considered as the ligand (see below).

3. Thermodynamic Bases of Quantitative Equilibrium Spectroscopic Titrations

Analysis of the binding of a protein to a nucleic acid can be performed using two different types of equilibrium spectroscopic titrations.^{6,10,13,14,18,23,24,27} In the first type, the nucleic acid (macromolecule) is titrated with the protein (ligand). This approach is referred to as a “normal” titration, since the total average degree of binding, $\sum \Theta_i$ (average number of moles of protein bound, L_B , per mole of the nucleic acid, N_T), $\sum \Theta_i = (L_B/N_T)$, increases as the titration progresses.^{6,10,35,36,38,40} Notice, that the symbol, $\sum \Theta_i$, is used, instead of just Θ , to describe the total average degree of binding. This is because, in the general case of a multiple ligand binding system, the bound ligand molecules can be distributed over “ i ” possible different bound states, all contributing to the total average degree of binding, $\sum \Theta_i$.

In the other type of titration, the protein (ligand) is titrated with the nucleic acid (macromolecule). This type of titration is referred to as a “reverse” titration, since the degree of binding decreases throughout the titration.^{6,10,13,14,17,18,23,24,28} Generally, the type of titration that is performed will depend on whether the signal that is monitored is from the macromolecule (normal) or the ligand (reverse). As pointed out above, the first task in examining ligand–macromolecule interactions is to convert a spectroscopic titration curve, that is, a change in the monitored signal as a function of the concentration of the titrant, into a model-independent, thermodynamic binding isotherm, which can then be analyzed, using an appropriate binding model to extract binding parameters.

The thermodynamic basis of the method is that the total average degree of binding, $\sum \Theta_i$, of the ligand on a macromolecule, including all different distributions of ligands bound in all possible different states i , is a sole function of the free ligand concentration, L_F , at equilibrium.^{6,10,35–50,71} In other words, any intensive property of the macromolecule can be used to monitor the binding if this property is affected by the state of ligation of the macromolecule. At a given free ligand concentration, L_F , the value of $\sum \Theta_i$ will be the same, independent of the concentration of the macro-

molecule, N_T . The unique values of L_F and $\sum \Theta_i$ and the corresponding total ligand, L_T , and total macromolecule, N_T , concentrations must satisfy the mass conservation equation

$$L_T = (\sum \Theta_i)N_T + L_F \quad (1)$$

Therefore, if the set of concentrations (L_T , N_T) can be found for which $\sum \Theta_i$ and L_F are constant, then $\sum \Theta_i$ and L_F can be determined from the slope and intercept, respectively, of a plot of L_T vs N_T , on the basis of eq 1. In theory and in practice, only two sets of concentrations of N_T and L_T are often enough to obtain $\sum \Theta_i$ and L_F over $\sim 85\%$ of the entire binding isotherm^{35–50} (see below).

3.1. The Spectroscopic Signal Used to Monitor Protein–Nucleic Acid Interactions Originates from the Nucleic Acid (Macromolecule)

In this case, the observed signal monitors the progress of the saturation of a nucleic acid with a protein and a “normal” titration (addition of a protein to a sample with a constant nucleic acid concentration) is performed. As mentioned above, for the total nucleic acid concentration, N_T , the equilibrium distribution of the nucleic acid among its different ligation states, N_i , is determined solely by the free protein concentration, L_F .^{6,10,35–50} Therefore, at each L_F , the observed spectroscopic signal, S_{obs} , is the algebraic sum of the concentrations of the nucleic acid in each state, N_i , each weighted by the value of the intensive spectroscopic property of that state, S_i . In general, a nucleic acid will have the ability to bind n ligands; hence the “signal conservation” equation for the observed signal, S_{obs} , of a sample containing the ligand at a total concentration L_T and the nucleic acid at a total concentration N_T is given by^{6,10,42}

$$S_{\text{obs}} = S_F N_F + \sum S_i N_i \quad (2)$$

where S_F is the molar signal of the free nucleic acid and S_i is the molar signal of the complex, N_i , which represents the nucleic acid with i bound ligands ($i = 1$ to n). The presence of, for example, cooperative interactions or the presence of different binding modes do not affect eq 2, because it defines a general distribution of all bound species of the nucleic acid grouped according to the number of the bound ligand molecules. The mass conservation equation, which relates N_F and N_i to N_T , is given by

$$N_T = N_F + \sum N_i \quad (3)$$

The partial degree of binding, Θ_i (i moles of protein bound per mole of nucleic acid), corresponding to all complexes with a given number i of bound protein molecules is given by

$$\Theta_i = \frac{iN_i}{N_T} \quad (4)$$

Therefore, the expression for N_i , the concentration of macromolecule with i protein molecules bound, is

$$N_i = \left(\frac{\Theta_i}{i}\right)N_T \quad (5)$$

Introducing eqs 4 and 5 into eq 2 provides a general

relationship for the observed spectroscopic signal, S_{obs} , as

$$S_{\text{obs}} = S_F N_T + \left[\sum (S_i - S_F) \left(\frac{\Theta_i}{i}\right) \right] N_T \quad (6)$$

Subsequently, by rearranging eq 6, one can define the normalized quantity ΔS_{obs} as

$$\Delta S_{\text{obs}} = \frac{(S_{\text{obs}} - S_F N_T)}{S_F N_T} \quad (7a)$$

and

$$\Delta S_{\text{obs}} = \sum \left(\frac{\Delta S_i}{i}\right) (\Theta_i) \quad (7b)$$

Notice, $\Delta S_{\text{obs}} = (S_{\text{obs}} - S_F N_T)/(S_F N_T)$ is the experimentally determined fractional signal change observed at the total protein and nucleic acid concentrations L_T and N_T , and $(\Delta S_i/i) = [(S_i - S_F)/S_F]/i$ is the average molar signal change per bound protein in the complex containing i protein molecules. The quantity $\Delta S_{\text{obs}} = (S_{\text{obs}} - S_F N_T)/(S_F N_T)$ as a function of the total average degree of binding, $\sum \Theta_i$, is referred to as the macromolecular binding density function (MBDF).^{6,9,42}

Since $\Delta S_i/i$ is an intrinsic molecular property of the protein–nucleic acid complex with i protein molecules bound, eq 7 indicates that ΔS_{obs} is only a function of the degree of binding distribution, $\sum \Theta_i$. Therefore, the total average degree of saturation of the nucleic acid and the total average degree of binding of the protein, $\sum \Theta_i$, must be the same for any value of L_T and N_T for which ΔS_{obs} is constant. Thus, when one performs a spectroscopic titration of a nucleic acid with a protein, at different total nucleic acid concentrations, N_T , the same value of ΔS_{obs} indicates the same physical state of the nucleic acid, that is, the same degree of the nucleic acid saturation with the protein and the same $\sum \Theta_i$. Since $\sum \Theta_i$ is a unique function of the free protein concentration, L_F , then the value of L_F at the same degree of saturation must also be the same. The above derivation is rigorous and independent of any binding model and, as such, can be applied to any binding system, with or without cooperative interactions or with overlapping of binding sites.^{6,10,35–50}

3.1.1. Analysis of Spectroscopic Titration Curves When the Signal Originates from the Nucleic Acid. The Case of a Single Binding Site

As an example of the quantitative analysis of a simple system of the protein binding to a nucleic acid, we consider the data from studies of the binding of the *E. coli* replicative helicase DnaB protein to the fluorescent etheno derivative of the ssDNA 20-mer, $d\epsilon A(\text{peA})_{19}$.^{35,36} The DnaB protein is a primary replicative helicase of the *E. coli* cell, that is, the factor that is responsible for unwinding duplex DNA in front of the replication fork.^{72–74} The DnaB monomer is built of the two domains, a small 12-kDa and a large 31-kDa domain at the N- and C-terminus of the protein.⁷⁵ Hydrodynamic and electron microscopy studies showed that the enzyme forms a stable ringlike hexamer built of six chemically identical subunits, as depicted in Figure 1.^{76–80} The diameter of the cross channel of the hexamer is ~ 40 Å. The use of the etheno derivative of the ssDNA is dictated by the fact that binding of the helicase to the unmodified ssDNA does not induce any adequate changes in the protein fluorescence.

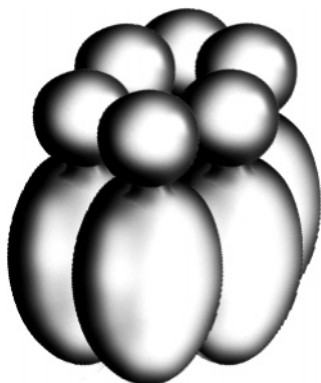


Figure 1. Schematic model of the *E. coli* DnaB hexameric helicase based on biochemical, hydrodynamic, and electron microscopy data.

This is a common problem in various protein–nucleic acid systems, which can be overcome by using a suitable fluorescent derivative of the nucleic acid lattice. Subsequently, interactions with unmodified DNA or RNA can be addressed in competition experiments with the fluorescent lattice (see below). Thus, association of the DnaB helicase with the fluorescent etheno derivative of the ssDNAs induces a strong increase of the nucleic acid fluorescence; hence, this signal, originating from the DNA, can be used to monitor the complex formation.^{35,36}

Fluorescence titrations of the ssDNA 20-mer, $d\epsilon A(p\epsilon A)_{19}$ (macromolecule), with the DnaB protein (ligand) at three different 20-mer concentrations are shown in Figure 2a.³⁵ The values of ΔS_{obs} , as defined by eq 7, are plotted as a function of the logarithm of the total DnaB protein concentration, L_T . At higher nucleic acid concentration, a given relative fluorescence increase, ΔS_{obs} , is reached at higher protein concentrations. This results from the fact that at higher DNA concentrations, more protein is required to saturate the increased concentration of the nucleic acid and to reach the same total average degree of binding, $\sum \Theta_i$. Also depicted in Figure 2a is the approach by which a set of values of $(\sum \Theta_i)_j$ and $(L_F)_j$ for the selected “ j ” value of $(\Delta S_{\text{obs}})_j$ can be obtained from these data. A horizontal line that intersects both curves at the same value of $(\Delta S_{\text{obs}})_j$ is drawn. The point of intersection of this horizontal line with each titration curve defines the set of values of L_{Tj} and N_{Tj} for which $(L_F)_j$ and $(\sum \Theta_i)_j$ are constant, as discussed above. The example in Figure 2a has three titration curves; hence sets of three values of L_{Tj} and N_{Tj} can be obtained at each selected $(\Delta S_{\text{obs}})_j$. On the other hand, in the considered case, the separation of the titration curves on the DnaB concentration scale is large and only two titration curves can be selected to obtain a quantitative estimate of $\sum \Theta_i$ and L_F .³⁵

For the two titrations at macromolecular (20-mer) concentrations, for example, N_{T1} and N_{T2} ($N_{T2} > N_{T1}$), two mass conservation equations can be written in the form of eqs 1 and 3 for each set of L_{Tj} and N_{Tj} . These two equations can then be solved for $(\sum \Theta_i)_j$ and $(L_F)_j$ as

$$(\sum \Theta_i)_j = \frac{(L_{T2} - L_{T1})}{(N_{T1} - N_{T2})} \quad (8)$$

and

$$(L_F)_j = (L_T)_j - (\sum \Theta_i)_j(N_T) \quad (9)$$

In this manner, model-independent values of $(L_F)_j$ and $(\sum \Theta_i)_j$

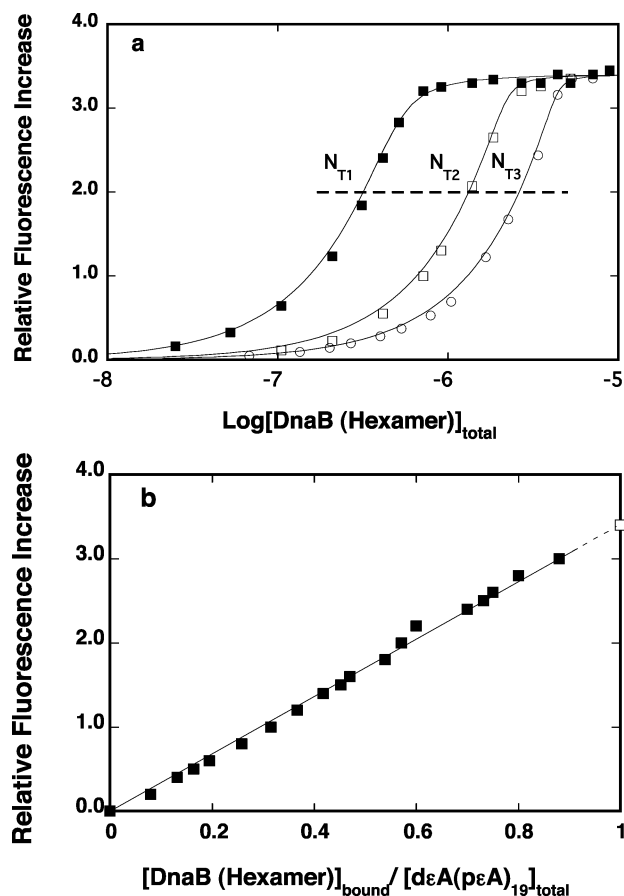


Figure 2. (a) Fluorescence titrations of the ssDNA 20-mer $d\epsilon A(p\epsilon A)_{19}$ with the DnaB protein ($\lambda_{\text{ex}} = 325$ nm, $\lambda_{\text{em}} = 410$ nm) in 50 mM Tris/HCl (pH 8.1, 10 °C) containing 50 mM NaCl, 5 mM MgCl₂, and 1 mM AMP-PNP at three different nucleic acid concentrations: (■) 4.9×10^{-7} M; (□) 2.2×10^{-6} M; (●) 4.6×10^{-6} M. The solid lines are nonlinear least-squares fits of the titration curves using the single-binding site isotherm with a single binding constant, $K_{20} = 4.8 \times 10^7$ M⁻¹, and spectroscopic parameter, $\Delta S_{\text{max}} = 3.4$. (b) Dependence of the relative increase of the $d\epsilon A(p\epsilon A)_{19}$ fluorescence upon the average number of DnaB hexamers bound per oligomer. The dashed line is the extrapolation to $\Delta S_{\text{max}} = 3.4$ (□). Reprinted with permission from ref 35. Copyright 1995 American Chemical Society.

can be obtained at any selected j value of $(\Delta S_{\text{obs}})_j$, yielding a set of values for $(L_F)_j$ and $(\sum \Theta_i)_j$.

Practically, the accuracy of the determination of $(L_F)_j$ and $(\sum \Theta_i)_j$ will depend on the used concentrations of the nucleic acid (macromolecule). Thus, the most accurate estimates are obtained in the region of the titration curves where the concentration of a bound protein is comparable to its total concentration, L_T . Thus, concentration of the bound protein must constitute a significant fraction of the total protein concentration in solution. In our practice, this limits the accurate determination of the degree of binding, $(\sum \Theta_i)_j$, to the region of the titration curves where the concentration of the bound ligand is at least $\sim 10\%$ of the L_T . Therefore, a selection of proper concentrations of the macromolecule is crucial for obtaining $(L_F)_j$ and $(\sum \Theta_i)_j$ over the largest possible region of the titration curves, although the accuracy of the determination of $(\sum \Theta_i)_j$ is mostly affected in the region of the high concentrations of the ligand approaching the maximum saturation. Such a selection of macromolecule concentrations is usually based on preliminary titrations that provide initial estimates of the expected affinity of the protein for the nucleic acid.

In the first step of the analysis, the obtained values of $(\sum\Theta_i)_j$, corresponding to given values of $(\Delta S_{\text{obs}})_j$, are used to (1) determine the maximum stoichiometry of the nucleic acid–protein complex and (2) determine the relationship between the signal change, ΔS_{obs} , and the average number of bound protein molecules, in the considered case, the average number of bound DnaB molecules per ssDNA 20-mer. Both objectives can be achieved by plotting ΔS_{obs} as a function of $\sum\Theta_i$. The dependence of the relative increase of the ssDNA 20-mer fluorescence upon the total average degree of binding of the DnaB protein is shown in Figure 2b. The selected concentrations of the nucleic acid allowed us to obtain an accurate ($\pm 5\%$) estimate of the average degree of binding up to ~ 0.9 DnaB molecules bound. Although the maximum value of $\sum\Theta_i$ cannot be directly determined, due to the discussed inaccuracy at the high protein ligand concentration region, the plateau of the fluorescence quenching, ΔS_{max} , corresponding to the maximum saturation, can be determined with the accuracy of $\pm 5\%$ (see Figure 2a). Therefore, knowing the maximum increase of the nucleic acid fluorescence ($\Delta S_{\text{max}} = 3.4$), one can perform a short extrapolation of the plot (ΔS_{obs} vs $\sum\Theta_i$) to this maximum value. Such extrapolation of the data presented in Figure 2b shows that, at saturation, one DnaB protein hexamer binds to the ssDNA 20-mer, establishing the maximum stoichiometry of one DnaB hexamer per ssDNA 20-mer in the complex.³⁵

Notice that the plot of ΔS_{obs} as a function of $\sum\Theta_i$, shown in Figure 2b, is within experimental accuracy linear. This is expected for the simple system where only one ligand binds to the macromolecule (see above). Also, as shown in Figure 2a,b, only two titration curves are needed to obtain reliable values of $\sum\Theta_i$ and L_F that cover a very large range of the total degree of binding. This is possible for the data in Figure 2a, since the signal change of $d\epsilon A(p\epsilon A)_{19}$ (fluorescence increase) is large and the affinity under the solution conditions used is fairly high. However, if necessary, the values of $(\sum\Theta_i)_j$ and $(L_F)_j$ can be obtained for more than two titrations, and the data are then graphically analyzed using eq 1. For a case in which “ p ” titrations are performed, then p sets of values of L_{Tj} and N_{Tj} will be obtained, one for each titration curve that is intersected by each horizontal line. A plot of L_T vs N_T can then be constructed, which will result in a straight line (eq 1), from which the values of $(L_F)_j$ and $(\sum\Theta_i)_j$ can be obtained from the intercept and the slope. If a plot of L_T vs N_T determined in this manner is not linear, this may indicate one or more inconsistent sets of titration data or aggregation phenomena associated with the protein or the nucleic acid, and these data should be viewed with caution.^{6,10,42}

Because only one DnaB protein molecule binds to the ssDNA 20-mer, the dependence of the relative nucleic acid fluorescence, ΔS_{obs} , is defined in terms of a single binding constant, K_{20} , and the free protein concentration, L_F , as

$$\Delta S_{\text{obs}} = \frac{\Delta S_{\text{max}} K_{20} L_F}{1 + K_{20} L_F} \quad (10)$$

where $\Delta S_{\text{max}} = 3.4$ (see above). The solid lines in Figure 2a are the nonlinear least-squares fits of the fluorescence titration curves using eq 10 and using K_{20} as a single fitting parameter.

The interactions of the DnaB hexamer with ssDNA were very intensively examined over the past decade.^{35,36,38–40,77,81,82} Here, we would like to stress that even such a relatively

simple thermodynamic analysis of the ssDNA 20-mer binding to the DnaB hexamer, as discussed above, provides a profound insight into the functioning of the enzyme. For instance, the fact that an order of magnitude increase of the concentration of the 20-mer does not change the 1:1 stoichiometry of the complex (Figure 2a) indicates that the hexamer has only a single effective binding site for the ssDNA. The low stoichiometry indicates that only a limited number of subunits are engaged in interactions with the nucleic acid. Subsequent photo-cross-linking experiments showed that only one subunit is engaged in ssDNA binding.^{35,36} Moreover, in the context of the ringlike hexamer structure (Figure 1), the presence of only a single site and the very low stoichiometry also suggest that the binding site is located inside the ringlike structure of the hexamer where steric hindrances would not allow additional 20-mers to bind to the enzyme. This suggestion was later directly confirmed by direct and quantitative fluorescence energy transfer studies.^{81,82}

3.1.2. The Case of Two Binding Sites

A more complex situation, in terms of the stoichiometry of the formed complex and the nature of the association process, occurs in the binding of, for example, rat polymerase β (pol β) to the double-stranded (ds) DNA 10-mer.⁸³ Polymerase β (pol β) is one of a number of recognized DNA-directed polymerases of the eukaryotic nucleus that plays very specialized functions in the mammalian cell DNA-repair machinery.^{84–87} The enzyme possesses two functional domains, a large 31-kDa catalytic domain and a small 8-kDa domain, and both the 8-kDa and 31-kDa domains have DNA-binding capability.⁸⁴ However, only the DNA-binding subsite on the 8-kDa domain of pol β has been found to have similar and significant affinity for both ss- and dsDNA.^{43–46,83} Thus, the DNA-binding site on the 8-kDa domain is the primary DNA-binding site of the enzyme, which serves to initiate the contact with the DNA and provides the dominant part of the free energy of interactions of the enzyme with the nucleic acid. Not surprisingly (see above), binding of the enzyme to the dsDNA is not accompanied by large enough changes of the protein fluorescence that would allow an experimenter to examine the complex binding process. However, association of the enzyme with a dsDNA oligomer labeled at the 5' end of one of the ssDNA strands with the coumarin derivative CP is accompanied by a strong ($\sim 150\%$) increase of the nucleic acid fluorescence.⁸³ Such a large emission change provides an excellent signal to perform high-resolution measurements of the enzyme–dsDNA complex formation. Moreover, competition studies with unmodified DNA oligomers clearly show that the fluorescent label does not affect, to any detectable extent, the energetics of the interactions (see below).⁸³

Fluorescence titrations of the CP-labeled dsDNA 10-mer with rat pol β at two different nucleic acid concentrations are shown in Figure 3a. The selected DNA concentrations provide separation of binding isotherms up to $\Delta S_{\text{obs}} \approx 1.35$. Figure 3b shows the dependence of the observed relative fluorescence increase, ΔS_{obs} , as a function of the total average degree of binding, $\sum\Theta_i$, of rat pol β . The plot is clearly nonlinear and shows two binding phases. In the first phase, a single molecule of the polymerase binds with the relative fluorescence increase, ΔS_{obs} , reaching the value of ~ 0.85 at $\sum\Theta_i \approx 0.9$. Extrapolation of the second phase to the maximum value of the fluorescence increase, $\Delta S_{\text{max}} = 1.55$

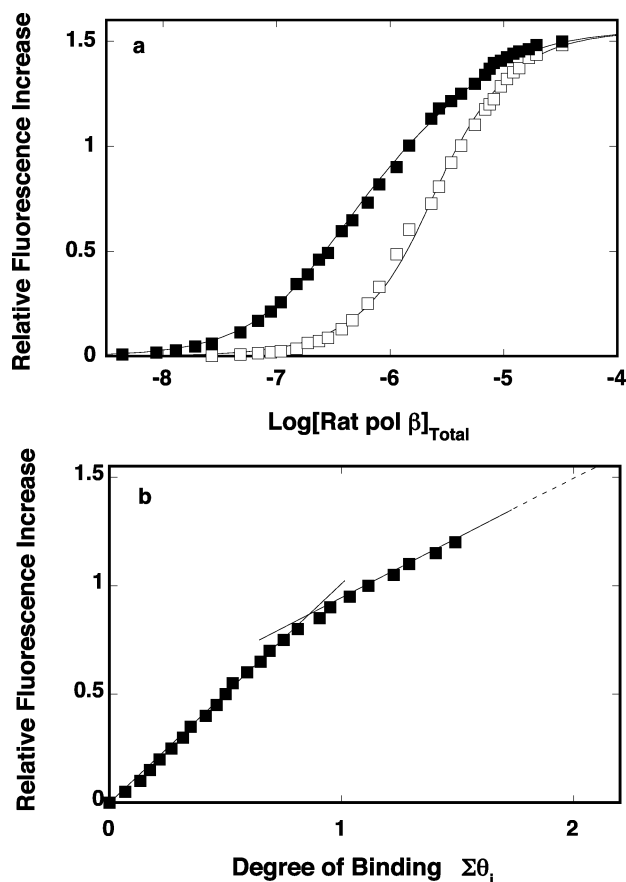


Figure 3. (a) Fluorescence titrations of the dsDNA 10-mer with the rat polymerase β ($\lambda_{\text{ex}} = 435$ nm, $\lambda_{\text{em}} = 480$ nm) in 10 mM sodium cacodylate/HCl (pH 7.0, 10 °C) containing 100 mM NaCl at two different nucleic acid concentrations: (■) 1.11×10^{-7} M; (□) 2.22×10^{-6} M (oligomer). The solid lines are nonlinear least-squares fits of the titration curves for the cooperative binding of two ligand molecules to the nucleic acid lattice with the site size of the complex $n = 5$ base pairs, the intrinsic binding constant $K = 7.8 \times 10^5$ M $^{-1}$, the cooperative interaction parameter $\omega = 2.3$, and the relative fluorescence changes $\Delta S_1 = 0.95$ and $\Delta S_2 = 0.6$ (see text for details). (b) Dependence of the relative fluorescence of the dsDNA oligomer, ΔS_{obs} , upon the average number of bound pol β molecules (■). The solid line follows the experimental points and has no theoretical basis. The dashed line is the extrapolation of ΔS_{obs} to the maximum value of $\Delta S_{\text{max}} = 1.55$. Reprinted with permission from ref 49. Copyright 2003 American Chemical Society.

± 0.1 , gives the maximum value of $\Sigma\Theta_i = 2.2 \pm 0.1$. Thus, the data indicate that at saturation, two rat pol β molecules bind to the dsDNA 10-mer. Analogous data using dsDNA oligomers containing an extra 5 and 10 bp (15- and 20-mer) indicate consistent and corresponding increases of the maximum number of the polymerase molecules bound to the nucleic acid.⁸³ Such consistency would not be observed if the enzyme formed different binding modes with the dsDNA with a different number of occluded base pairs by the protein in each mode (see below). Therefore, these results indicate that the polymerase forms a single binding mode in the complex with the dsDNA; that is, it forms a single type of complex with a site size of $n = 5 \pm 1$ base pairs independent of the available length of the nucleic acid.⁸³

3.1.2.1. Application of a Binding Model To Analyze the Binding Isotherm. The analysis of the maximum stoichiometry of the pol β -dsDNA 10-mer complex described in the previous section is completely model-independent; that

is, no specific model of the pol β binding to the dsDNA oligomer has been assumed. Once this analysis has been performed, one can postulate, on the bases of the physical properties of the system, a specific statistical thermodynamic model, which describes the binding isotherms and allows the experimenter to obtain intrinsic affinities and parameters characterizing the cooperative interactions within the system. Briefly, pol β does not possess any significant base or base pair specificity.^{43-46,84-87} The reader should consult the original papers for a full discussion of this aspect of the enzyme interactions with the nucleic acid.^{43-46,84-87} The data in Figure 3b indicate that the site size of the protein-dsDNA complex is $n = 5 \pm 1$ base pairs and the binding of the first polymerase molecule induces a larger increase of the labeled nucleic acid fluorescence than the binding of the second enzyme molecule. Therefore, the simplest statistical thermodynamic model that describes the rat pol β binding to the dsDNA 10-mer is defined by the partition function, Z_D , as⁸³

$$Z_D = 1 + (N - n + 1)KL_F + \omega(KL_F)^2 \quad (11)$$

where K is the intrinsic binding constant, N is the total number of base pairs in the oligomer (in the considered case, $N = 10$), n is the site size of the pol β dsDNA complex, ω is the parameter characterizing the cooperative interactions between the bound protein molecules, and L_F is the free pol β concentration. The factor $(N - n + 1)$ results from the fact that the first polymerase molecule experiences the presence of $N - n + 1$ potential binding sites on the DNA oligomer.^{6,10,11,12,88} The total degree of binding, $\Sigma\Theta_i$, is then described by

$$\Sigma\Theta_i = \frac{[(N - n + 1)KL_F + 2\omega(KL_F)^2]}{Z_D} \quad (12)$$

The observed relative fluorescence increase, ΔS_{obs} , of the nucleic acid is then

$$\Delta S_{\text{obs}} = \Delta S_1 \left[\frac{(N - n + 1)KL_F}{Z_D} \right] + (\Delta S_1 + \Delta S_2) \left[\frac{\omega(KL_F)^2}{Z_D} \right] \quad (13)$$

where ΔS_1 and ΔS_2 are relative molar fluorescence increases accompanying the binding of the first and second rat pol β molecule to the dsDNA 10-mer.

The determination of all interaction and spectroscopic parameters of this binding system can be achieved by applying the following strategy, which relies on the fact that two sets of data, original fluorescence titration curves, and the plot of ΔS_{obs} as a function of the total average degree of binding, $\Sigma\Theta_i$, are available and the maximum stoichiometry of the complex is known. The value of ΔS_1 can be obtained as the slope, $\Delta S_1 = \partial\Delta S_{\text{obs}}/\partial(\Sigma\Theta_i)$, of the initial part of the plot in Figure 3b, which provides $\Delta S_1 = 0.95 \pm 0.05$.⁸³ The determination of ΔS_2 is based on the fact that the final complex, at saturation, must contain two rat pol β molecules bound to the dsDNA, that is, $\Delta S_{\text{max}} = \Delta S_1 + \Delta S_2$. Therefore, the determined value of $\Delta S_{\text{max}} = 1.55 \pm 0.1$ provides $\Delta S_2 = 0.6 \pm 0.05$. Thus, there are only two remaining binding parameters that must be determined, K and ω . The solid lines in Figure 3a are the nonlinear least-squares fits of the experimental titration curves to eq 13 with intrinsic binding constant K and cooperativity parameter ω as the fitting

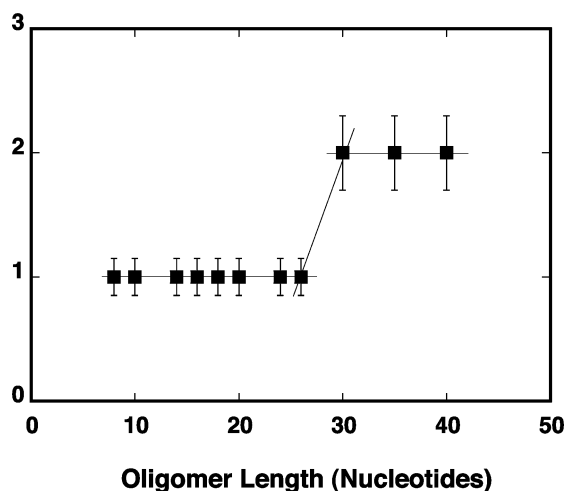


Figure 4. The maximum number of bound PriA molecules as a function of the length of the ssDNA oligomer in 10 mM sodium cacodylate/HCl (pH 7.0, 10 °C) containing 100 mM NaCl. The solid lines follow the experimental points and have no theoretical bases. The number of the bound PriA molecules has been determined using the quantitative approach described in the text. Reprinted with permission from ref 47. Copyright 2000 American Society for Biochemistry and Molecular Biology.

parameters. It is clear that the model (eqs 11–13) provides an excellent description of the experimentally observed binding process.⁸³

3.1.3. Two Different Binding Modes of the Protein–Nucleic Acid Complex

3.1.3.1. Anatomy of the Total DNA-Binding Site. The total site size of a large protein ligand–DNA complex corresponds to the DNA fragment, which includes nucleotides directly involved in interactions with the protein, its proper DNA-binding site, and, in general, nucleotides that may not be engaged in direct interactions or may interact differently with the protein matrix.^{6,10,11,12,47,48} The latter are prevented from interacting with another protein molecule by the protruding protein matrix of the previously bound protein molecule over nucleotides adjacent to the proper DNA-binding site. Binding of the monomeric *E. coli* PriA helicase to the ssDNA provides an example of such a complex structure of the total ssDNA-binding site of the protein interacting with the nucleic acid.^{47,48} Quantitative analysis of the maximum stoichiometry of PriA–ssDNA complexes has been performed for a series of etheno derivatives of ssDNA oligomers. The dependence of the maximum number of bound PriA molecules per ssDNA oligomer, determined using the quantitative approach discussed above, upon the length of the oligomer is shown in Figure 4. Oligomers from 8 to 26 nucleotides bind a single PriA molecule. Transition from a single enzyme bound per oligomer to two bound protein molecules per ssDNA occurs between 26- and 30-mers. However, a further increase in the length of the oligomer, up to 40 residues, does not lead to an increased number of bound PriA molecules per ssDNA oligomer. These results provide the first indication that the total site size of the PriA–ssDNA complex is less than 24 nucleotides, but it must contain at least 15 nucleotides per bound protein molecule (see below). Also, the fact that both 8- and 26-mers can bind only a single enzyme molecule yet the length of the 26-mer is more than three times longer than the length of the 8-mer clearly indicates that only a part of the total

DNA-binding site is involved in direct interactions with the nucleic acid.⁴⁷

Further analysis requires the evaluation of the number of nucleotides directly engaged in interactions with the proper ssDNA-binding site of the PriA helicase. Recall that a single molecule of the enzyme binds to 26-, 14-, 10-, and 8-mers (Figure 4). Moreover, the binding to these oligomers is characterized by very similar intrinsic affinities, that is, the macroscopic affinities corrected for the statistical effect, resulting from the presence of the potential binding sites on the DNA.^{47,48} A very dramatic drop in the intrinsic affinity occurs when the number of nucleotides is lower than eight with the affinity of the 6-mer being undetectable. These data indicate that the number of nucleotides directly involved in interactions with the proper ssDNA-binding site of the PriA helicase within the total site size of the protein–ssDNA complex is 8 ± 1 . Moreover, the fact that the helicase binds only a single 8-mer molecule also indicates that the enzyme possesses only one proper ssDNA-binding site.

With these data, we can address the structure of the total ssDNA-binding site of the PriA protein. If the proper ssDNA-binding site of the protein is located on one side of the molecule, with a part of the enzyme protruding over the extra seven residues, and the total site size is 15 residues, then the 26- and 40-mer would be able to accommodate two and three PriA molecules, respectively, as depicted in Figure 5. This is not what is experimentally observed. One PriA molecule binds to the 26-mer, and only two enzyme molecules bind to the 40-mer (Figure 4). Therefore, the model in Figure 5 cannot represent the PriA–ssDNA complex. Next, we consider an alternative model where the proper ssDNA-binding site, which engages eight nucleotides, is located in the central part of the protein, as depicted in Figure 6. Also, the protein molecule now has two parts that are protruding over six nucleotides on both sides of the proper ssDNA-binding site. In this model, only one PriA molecule can bind to the 24- and 26-mers. This is because the first bound molecule can now block at least 14 nucleotides. For efficient binding, the second protein molecule also needs a fragment of at least 14 nucleotides, which is larger than the remaining 11 and 12 residues of the 24- and 26-mers, respectively. The two bound PriA molecules require at least 28 nucleotides of the ssDNA. On the other hand, this allows two molecules of the enzyme to bind to the 30-, 35-, and 40-mers (Figure 4). In the case of the 40-mer, the remaining fragment of eight residues is six nucleotides too short, that is, it does not provide efficient interacting space to allow the third PriA protein to associate with the oligomer. This is exactly what is experimentally observed (Figure 4). Therefore, the model of the total ssDNA-binding site in Figure 6 adequately describes all experimentally determined stoichiometries of the PriA with the series of ssDNA oligomers. Moreover, these data and the analyses indicate that the actual total site size of the PriA–ssDNA complex is 20 ± 3 nucleotides and the complex includes eight residues encompassed by the proper ssDNA-binding site of the enzyme, as well as ~ 12 residues occluded by the protruding protein matrix (Figure 6).^{47,48}

3.1.3.2. Determination of the Binding Parameters for the PriA Protein–ssDNA Complex. The major aspect of the experimental strategy described above is the application of the large series of ssDNA oligomers of well-defined length. Oligomers range from a length shorter than the number of nucleotides encompassed by the proper ssDNA-

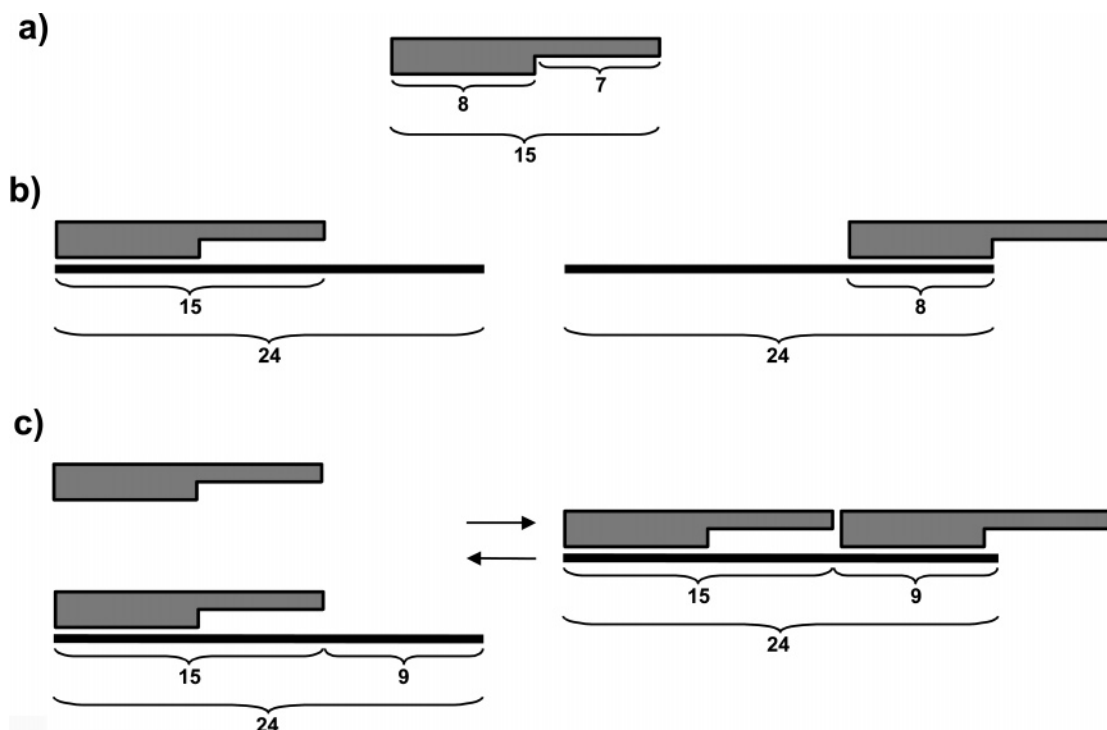


Figure 5. Schematic model for the binding of the PriA helicase to the ssDNA based on the minimum total site size of the protein–nucleic acid complex, $n = 15$, and the size of the binding site engaged in protein ssDNA interactions, $m = 8$. The helicase binds the ssDNA in a single orientation with respect to the polarity of the sugar–phosphate backbone of the ssDNA. The proper ssDNA-binding site, which encompasses eight nucleotides, is located on one side of the enzyme molecule with the rest of the protein matrix protruding over the extra seven nucleotides without engaging in thermodynamically significant interactions with the DNA (black ribbon) (a). When bound at the ends of the nucleic acid, or in its center, the protein can occlude 8 or 15 nucleotides (b). This model would allow the binding of two molecules of the PriA helicase to the ssDNA 24-mer and three molecules of the enzyme to the 40-mer (c). This is not experimentally observed. Reprinted with permission from ref 47. Copyright 2000 American Society for Biochemistry and Molecular Biology.

binding site of the enzyme to ssDNA oligomers that can accommodate two PriA molecules. Such an approach dramatically increases the resolution of the binding experiments and allows the experimenter to determine not only the total site size of the protein–ssDNA complex, n , but also the number of nucleotides directly engaged in interactions with the proper ssDNA-binding site of the protein, m . Only when these two parameters are known can the correct statistical thermodynamic model be formulated and the intrinsic binding parameters extracted.

Fluorescence titrations of the 40-mer, $d\epsilon A(p\epsilon A)_{39}$, with the PriA helicase at two different oligomer concentrations, are shown in Figure 7a. The dependence of the observed relative fluorescence increase as a function of the total average degree of binding, $\sum \Theta_i$, of the PriA helicase on the 40-mer is shown in Figure 7b. The value of $\sum \Theta_i$ could be determined up to ~ 1.5 . Short extrapolation to the maximum value of the fluorescence increase, $\Delta F_{\max} = 3.5 \pm 0.2$, gives the maximum value of $\sum \Theta_i = 2.0 \pm 0.2$. The plot is linear over the entire binding process, indicating that the binding of the first and second protein molecules induce the same fluorescence increase of the ssDNA 40-mer.

Quantitative analysis of the two PriA molecules binding to the 40-mer must, in general, include intrinsic affinity, possible cooperative interactions between bound protein molecules, and the overlap between potential binding sites on the nucleic acid lattice.^{6,10,11,12} We know that the total site size of the PriA–ssDNA complex is $n = 20 \pm 3$. However, the number of nucleotides engaged in interactions with the proper ssDNA-binding site of the enzyme is only $m = 8 \pm 1$, and the protein protrudes over a distance of 6 \pm

1 nucleotides on both sides of the binding site (Figure 6). Therefore, the partial degree of binding that involves only the first PriA molecule bound to the ssDNA 40-mer is described by

$$\sum \Theta_i = \frac{(N - m + 1)K_{40}L_F}{1 + (N - m + 1)K_{40}L_F} \quad (14)$$

where $N = 40$, $m = 8$, and K_{40} is the intrinsic binding constant for the 40-mer. The factor $N - m + 1$ indicates that the first single PriA molecule experiences the presence of 33 potential binding sites on the 40-mer. As the protein concentration increases, this complex is replaced by the complex with two PriA molecules bound to the nucleic acid, and there are several different possible configurations of the two proteins on the 40-mer (Figures 4 and 6). To derive the part of the partition function corresponding to the binding of two PriA molecules, we apply a combinatorial analysis for the cooperative binding of a large ligand to a finite lattice.^{12,48,49} The complete partition function for the PriA–40-mer system, Z_{40} , is then

$$Z_{40} = 1 + (N - m + 1)K_{40}L_F + \sum_{j=0}^{k-1} S_N(k,j)(K_{40}L_F)^k \omega^j \quad (15)$$

where $k = 2$ and j is the number of cooperative contacts between the bound PriA molecules in a particular configuration on the lattice, and ω is the parameter characterizing the cooperative interactions. The factor $S_N(k,j)$ is the number

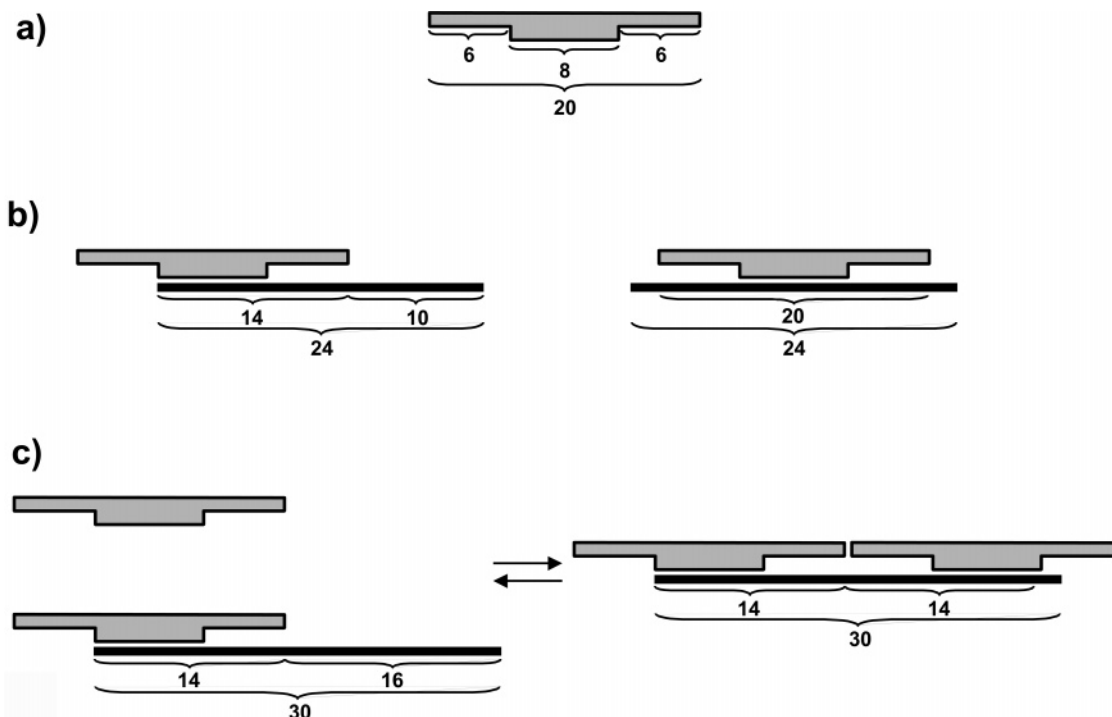


Figure 6. Schematic model for the binding of the PriA helicase to the ssDNA based on the total site size of the protein–nucleic acid complex, $n = 20$, and the size of the ssDNA-binding site engaged in protein–ssDNA interactions, $m = 8$. The helicase binds the ssDNA in a single orientation with respect to the polarity of the sugar–phosphate backbone of the ssDNA. The ssDNA-binding site of the PriA helicase, which encompasses only eight nucleotides, is located in the center of the enzyme molecule (a). The protein matrix protrudes over six nucleotides on both sides of the ssDNA-binding site without engaging in interactions with the nucleic acid (black ribbon). When bound at the 5' or the 3' end of the nucleic acid, the protein always occludes 14 nucleotides, while in the complex in the center of the ssDNA oligomer, 20 nucleotides are occluded (b). This model would allow the binding of only one molecule of the PriA helicase to the 24-mer and only two molecules of the enzyme to the 30-, 35-, and 40-mers (c). This is in agreement with all of the experimental data on stoichiometries of the protein complexes with various ssDNA oligomers. Reprinted with permission from ref 47. Copyright 2000 American Society for Biochemistry and Molecular Biology.

of distinct ways that two protein ligands bind to a lattice with j cooperative contacts and is defined by¹²

$$S_N(k,j) = \frac{[(N - (m + 6)k + 1)!(k - 1)!]}{[(N - (m + 6)k - k + j + 1)!(k - j)!j!(k - j - 1)!]} \quad (16)$$

The total average degree of binding, $\sum \Theta_i$, is defined as

$$\sum \Theta_i = \frac{[(N - m + 1)K_{40}L_F + \sum_{j=0}^{k-1} S_N(k,j)k(K_{40}L_F)^k \omega^j]}{Z_{40}} \quad (17)$$

The observed relative fluorescence increase of the nucleic acid, ΔS_{obs} , is then

$$\Delta S_{\text{obs}} = \Delta S_1 \left[\frac{(N - m + 1)K_{40}L_F}{Z_{40}} \right] + \Delta S_{\text{max}} \left[\frac{\sum_{j=0}^{k-1} S_N(k,j)(K_{40}L_F)^k \omega^j}{Z_{40}} \right] \quad (18)$$

where ΔS_1 and ΔS_{max} are relative molar fluorescence parameters that characterize the complexes with one and two PriA molecules bound to the 40-mer, respectively. Notice

that there are only two unknown parameters, K_{40} and ω , in eqs 14–18. The solid lines in Figure 7a are the nonlinear least-squares fits of the spectroscopic titration curves according to the considered model.

It should be pointed out that the binding of the PriA protein to the ssDNA constitutes an example, albeit very specific one, of a protein binding to a nucleic acid in two binding modes differing by the number of the occluded nucleotides.^{28,30,31,70} In one mode, the protein forms a complex with the site size of 20 ± 3 , that is, the total ssDNA-binding site is occluding the nucleic acid lattice, and in the other binding mode, the protein can associate with the DNA using only its proper ssDNA-binding site with the site size of 8 ± 1 residues, with an additional six residues occluded by the protruding protein matrix (Figure 6). The “specific” nature of the PriA helicase is that on the polymer ssDNA lattices only a single binding mode with the site size of 20 ± 3 residues will be observed due to the location of the proper ssDNA-binding in the central part of the protein molecule (Figure 6). Moreover, in this mode, only the proper ssDNA-binding site engages in interactions with the DNA, that is, no additional areas of interactions are being engaged, as compared to the proper ssDNA-binding site, in the examined solution conditions. In other words, the detection of the PriA binding modes was only possible through the experimental strategy of examining the stoichiometry of the protein–ssDNA complex using an extensive series of the ssDNA oligomers.^{47,48} We will return to the problem of different binding modes in subsequent sections of this review.

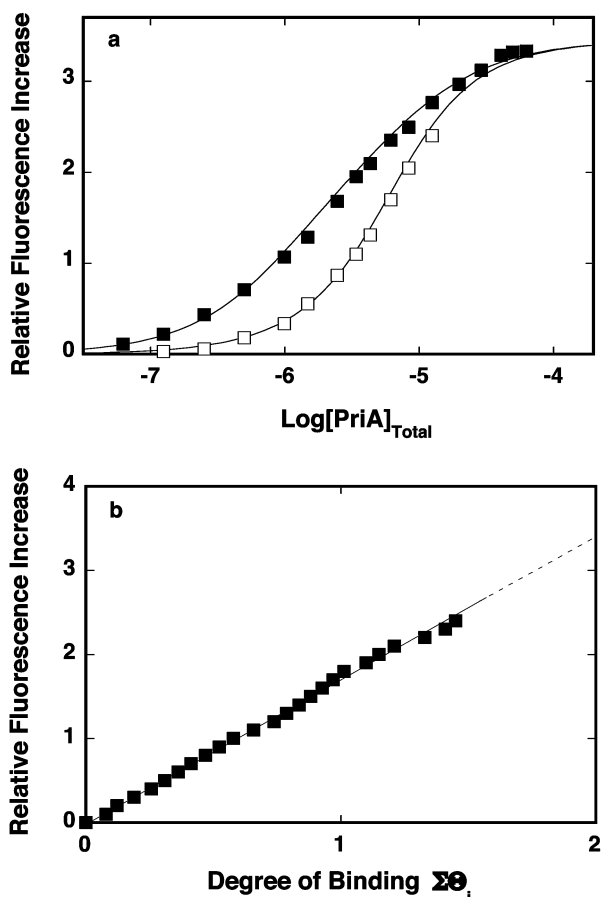


Figure 7. (a) Fluorescence titrations of the 40-mer $d\epsilon A(p\epsilon A)_{39}$ with the PriA protein ($\lambda_{\text{ex}} = 325 \text{ nm}$, $\lambda_{\text{em}} = 410 \text{ nm}$) in 10 mM sodium cacodylate/HCl (pH 7.0, 10 °C) containing 100 mM NaCl at two different nucleic acid concentrations: (■) $4.6 \times 10^{-7} \text{ M}$; (□) $4 \times 10^{-6} \text{ M}$. The solid lines are nonlinear least-squares fits of the titration curves using the statistical thermodynamic model for the binding of two PriA molecules to the 40-mer (see text for details). The intrinsic binding constant is $K_i = 6 \times 10^4 \text{ M}^{-1}$, the cooperativity parameter is $\omega = 0.8$, and the relative fluorescence changes are $\Delta S_1 = 1.7$ and $\Delta S_{\text{max}} = 3.5$. (b) Dependence of the relative fluorescence increase of the 40-mer, ΔS_{obs} , upon the average number of bound PriA proteins (■). The solid line follows the experimental points and has no theoretical basis. The dashed line is the extrapolation of ΔS_{obs} to the maximum value of $\Delta S_{\text{max}} = 3.5$. Reprinted with permission from ref 47. Copyright 2000 American Society for Biochemistry and Molecular Biology.

3.1.4. Protein Binding to a One-Dimensional Homogeneous Lattice in a Single Binding Mode

Quantitative analyses to obtain the total average degree of binding described above are fully applicable to the cooperative binding of a protein ligand to an infinite homogeneous lattice.^{6,9,10,35,36,42} This type of binding process is of great significance for elucidation the physiological role played by single-strand binding proteins (SSB), histones, and histone-like proteins in DNA and RNA metabolism, as well as, the role of cooperative interactions in various proteins involved in interactions with the nucleic acid. The large protein ligand occludes a number of nucleotides or base pairs, n (site size), in the complex with the nucleic acid. Each base or base pair can be an initial point of the binding site, that is, the binding sites overlap and the number of potential binding sites is a nonlinear function of the fractional saturation of the long nucleic acid lattice. Therefore, instead of the total degree of binding, $\sum \theta_i$, one considers the total

binding density, $\sum \nu_i$, that is, the total average number of protein molecules bound per monomer of the long lattice. The intrinsic binding constant, K , characterizes the intrinsic interactions with the site size of the complex. Moreover, the bound protein molecules can engage in cooperative interactions characterized by the cooperativity parameter, ω .

The simplest and paradigm statistical thermodynamic model that describes the binding of a large protein ligand, which occludes a number of n nucleotides or base pairs in the complex, to an infinite, homogeneous lattice is the McGhee–von Hippel model.¹¹ Although, originally, two implicit equations were obtained for the noncooperative and cooperative binding cases, these two expressions can be unified in one general equation that describes both cooperative and noncooperative binding.⁸⁸ Besides its theoretical value, by virtue of eliminating the necessity of using multiple equations, the generalized McGhee–von Hippel equation is particularly useful in any computer fitting or simulation of large ligand binding to the nucleic acid for $0 < \omega < \infty$. The binding density, $\sum \nu_i$, and the free protein concentrations are described, using the generalized equation, as

$$\sum \nu_i = K(1 - n \sum \nu_i) \{ [2\omega(1 - n \sum \nu_i) / ((2\omega - 1)(1 - n \sum \nu_i) + \sum \nu_i + R)]^{(n-1)} \{ [1 - (n + 1) \sum \nu_i + R] / [2(1 - n \sum \nu_i)] \}^2 L_F \} \quad (19a)$$

$$L_F = \sum \nu_i / [K(1 - n \sum \nu_i) \{ [2\omega(1 - n \sum \nu_i) / ((2\omega - 1)(1 - n \sum \nu_i) + \sum \nu_i + R)]^{(n-1)} \{ [1 - (n + 1) \sum \nu_i + R] / [2(1 - n \sum \nu_i)] \}^2 \}] \quad (19b)$$

where $R = \{ [1 - (n + 1) \sum \nu_i]^2 + 4\omega \sum \nu_i (1 - n \sum \nu_i) \}^{0.5}$.

The first step in the analysis of the spectroscopic titration curves of a protein binding to polymer nucleic acid is to determine the site size of the formed complex, n , that is, the maximum stoichiometry of the formed complex and the relationship between the observed signal and the binding density, $\sum \nu_i$. Once the site size, n , of the complex and the relationship between the observed signal and $\sum \nu_i$ is known, the paradigm statistical thermodynamic binding model, defined by eqs 19a and 19b, allows one to extract an intrinsic binding constant K and the parameter ω , and takes into account the overlap among potential binding sites.

Here we consider the binding of the isolated 8-kDa domain of the rat polymerase β to an etheno derivative of poly(dA), poly(dεA), where the binding is monitored by following the fluorescence increase of poly(dεA).⁴⁹ Recall, the DNA-binding subsite on the 8-kDa domain is the primary DNA-binding site of the enzyme, which provides the dominant part of the free energy of interactions of the polymerase with both the ss- and dsDNA.^{43–46} Fluorescence titrations of poly(dεA) with the rat pol β 8-kDa domain at two different nucleic acid concentrations are shown in Figure 8a. The relative increase of the nucleic acid fluorescence reaches the value of $\Delta S_{\text{max}} = 1.4 \pm 0.1$. At higher nucleic acid concentrations, a given fluorescence increase is reached at higher enzyme concentrations, due to the binding of the domain to the extra nucleic acid in the solution. The selected nucleic acid concentrations provide a separation of the titration curves up to the relative fluorescence increase of $\Delta S_{\text{obs}} \approx 1$, that is, up to $\sim 75\%$ of the relative observed signal change.

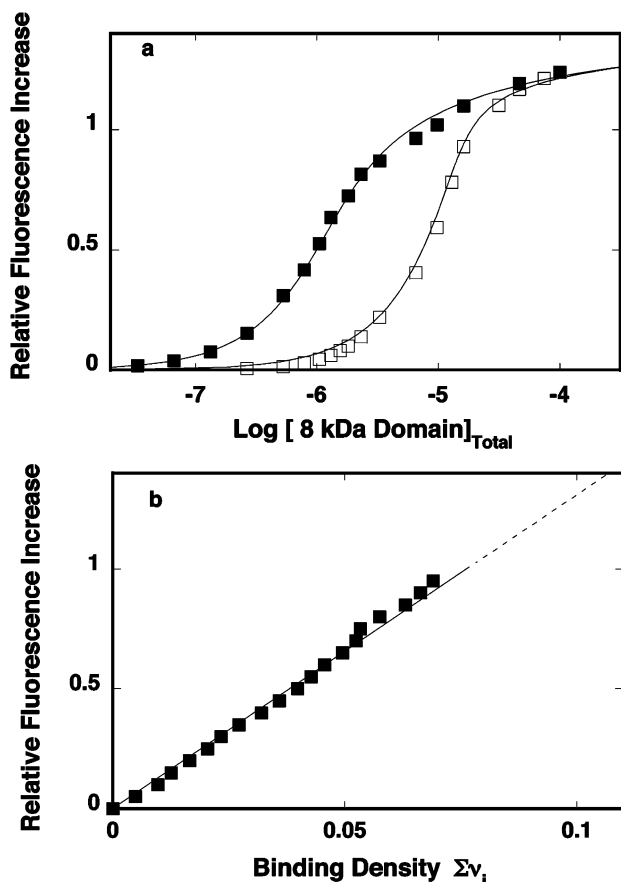


Figure 8. (a) Fluorescence titrations ($\lambda_{\text{ex}} = 325$ nm, $\lambda_{\text{em}} = 410$ nm) of poly(dεA) with the 8-kDa domain of rat pol β in sodium cacodylate/HCl (pH 7.0, 10 °C) containing 50 mM NaCl and 1 mM MgCl_2 at two different concentrations of the nucleic acid (nucleotide): (■) 2×10^{-5} M; (□) 1.94×10^{-4} M. The solid lines are nonlinear least-squares fits of the fluorescence isotherms using the generalized McGhee–von Hippel model as defined by eq 19a using a single set of parameters: $K = 3 \times 10^5$ M^{-1} , $n = 9$, $\omega = 1.5$, and $\Delta S_{\text{max}} = 1.4$. (b) The dependence of the relative fluorescence increase, ΔS_{obs} , upon the total average binding density, $\Sigma\nu_i$, of the 8-kDa domain on poly(dεA) (■). The solid straight line follows the points and does not have a theoretical basis. The dashed line is an extrapolation of the binding density to the maximum value of the relative fluorescence change, $\Delta S_{\text{max}} = 1.4$, which provides $\Sigma\nu_i = 0.115 \pm 0.005$, corresponding to the site size of the domain–ssDNA complex, $n = 9 \pm 0.6$. Reprinted with permission from ref 49. Copyright 2001 American Chemical Society.

Figure 8b shows the dependence of the observed relative fluorescence increase, ΔS_{obs} , as a function of the total average binding density, $\Sigma\nu_i$. The plot is linear and shows the existence of only a single binding phase. This linear behavior indicates that any possible cooperative interactions do not affect the spectroscopic properties of the protein–DNA complex.^{6,10,42} Extrapolation of the binding density to the maximum fluorescence increase at saturation provides $\Sigma\nu_i = 0.115 \pm 0.005$, which shows that in the examined solution conditions the isolated 8-kDa domain occludes $n = 9 \pm 0.6$ nucleotides in the complex with the ssDNA. As mentioned above, because the site size of the complex is determined and the domain binds to the ssDNA in a single phase with $\Delta F_{\text{max}} = 1.4 \pm 0.1$, we can use the generalized McGhee–von Hippel isotherm to extract the remaining two interaction parameters, the intrinsic binding constant, K , and the cooperativity parameter, ω (eq 19a). The solid lines in Figure 8a are nonlinear least-squares fits of the experimental

isotherms using eq 19a. The theoretical lines provide an excellent fit to the experimental isotherms with $K = (3 \pm 0.5) \times 10^5$ M^{-1} and $\omega = 1.5 \pm 0.5$.⁴⁹

3.1.5. Protein Binding to a One-Dimensional Homogeneous Lattice in Two Binding Modes Differing in the Number of Occluded Nucleotides

3.1.5.1. Statistical Thermodynamic Model. A much more complex situation than that considered above arises when the protein binds the nucleic acid using various binding modes that differ by the number of the occluded nucleotides or base pairs in the complex. This is more frequent behavior than previously thought and has been, for the first time, extensively and quantitatively characterized for the *E. coli* SSB protein.^{9,28,30,31,70,89} We have also already encountered similar behavior in the case of the *E. coli* PriA helicase (see above). Striking examples of the protein binding to the DNA in two different binding modes are the interactions of the human and rat polymerase β with the ssDNA.^{47,48}

We initiate our discussion with the heuristic derivation of the general closed-form parametric equations, which describe the binding of a large protein ligand to a one-dimensional homogeneous lattice in two binding modes, differing in the number of occluded nucleotides. Notice that in the McGhee–von Hippel model described above, a nucleic acid lattice site can only exist in two states, free and bound with the ligand in a single type of the complex, that is, a single binding mode. However, if the protein binds the ssDNA in two binding modes, differing by the number of occluded nucleotides, such interactions introduce multiple lattice-site states, that is, a site can be free, bound to the protein in one binding mode, or bound to the protein in another binding mode. This is a three-state lattice-binding model.⁸⁸ Such complex binding systems can be approached by a “brute” force of numerical fitting.²⁸ On the other hand, it turns out that a general analytical solution in terms of implicit, closed-form parametric expressions describing the binding isotherm for the ligand binding in two cooperative binding modes differing in the number of occluded nucleotides can be obtained.^{43,44} This can be accomplished in the most convenient way by using the sequence generating function method (SGF) that, in the case of a three-state lattice, requires only a 3×3 matrix to derive statistical thermodynamic expressions for the entire binding system.^{88,90,91,92}

Let the site size of the first and the second binding modes be n and m residues long, respectively. Then, let u_j be the statistical weight of j , the contiguous, empty lattice site. Because there are two lattice-site states with a bound protein, we assign a_j as the statistical weight of a sequence of j protein molecules bound to the lattice in the first binding mode and b_j as the statistical weight of a sequence of j protein molecules bound to the lattice in the second binding mode. Since the values of the statistical weights are relative, we designate $u_j = 1$. Consequently, $a_j = (K_n L_F) \omega_1^{j-1} = (1/\omega_1)(a\omega_1)^j$, where K_n is the intrinsic binding constant for an isolated protein in the first binding mode, L_F is the free protein concentration, ω_1 is the nearest-neighbor cooperativity parameter for the protein bound to the lattice site in the first binding mode, and $a = K_n L_F$. Analogously, for the protein bound to the lattice site in the second binding mode, $b_j = (K_m L_F) \omega_2^{j-1} = (1/\omega_2)(b\omega_2)^j$. The sequence generating functions for these three states of infinite lattice sites are

then defined in a standard way as^{88–92}

$$U(x) = \sum (x)^{-j} = \frac{1}{x-1} \quad (20)$$

$$P(x) = \left(\frac{1}{\omega_1}\right) \sum (a\omega_1)^j (x^n)^{-j} = \frac{a}{(x^n - a\omega_1)} \quad (21)$$

$$R(x) = \left(\frac{1}{\omega_2}\right) \sum (b\omega_2)^j (x^m)^{-j} = \frac{b}{(x^m - b\omega_2)} \quad (22)$$

To describe the appropriate alternations of the different possible sequences of states (U , P , R), the following matrix, \mathbf{M} , can be formed, where the elements of \mathbf{M} are the SGFs, defined as^{88,90}

$$\mathbf{M} = \begin{pmatrix} 0 & P & R \\ U & 0 & \omega_3 R \\ U & \omega_3 P & 0 \end{pmatrix} \quad (23)$$

where ω_3 is the cooperativity parameter characterizing interactions on the boundary between sequence R and P sequences, that is, between the first and second binding mode. This parameter appears as a multiplication factor in the matrix and not in the sequence generating functions, because the SGF is a sequence partition function constructed independently of any influence of other types of sequences, whereas the matrix equation serves to construct the grand partition function of the whole system.

For lattices N nucleotides in length, where N approaches infinity, the grand partition function of the protein–nucleic acid lattice system can be written as⁹⁰

$$Z = x_1^N \quad (24)$$

where x_1 is the largest root of the following secular matrix equation:

$$f(x) = |\mathbf{I} - \mathbf{M}| = 0 \quad (25)$$

where \mathbf{I} is the identity matrix. To obtain the average fraction of a lattice site in any state, Ψ , it is only necessary to treat eq 25 as an implicit function so that Ψ is

$$\Psi = - \left[\frac{\partial f(x)}{\partial \ln s} \right] \left[\frac{\partial \ln f(x)}{\partial \ln x^{-1}} \right]_{x_1} \quad (26)$$

where s is a statistical weight associated with a given particular state. Subscript x_1 indicates that the derivative is evaluated at $x = x_1$, the largest root of eq 25. Analogously, the total average binding density, $\sum \nu_i$, is then defined as

$$\sum \nu_i = - \left[\frac{\partial f(x)}{\partial \ln L_F} \right] \left[\frac{\partial \ln f(x)}{\partial \ln x^{-1}} \right]_{x_1} \quad (27)$$

After expansion, secular eq 25 takes the form

$$L_F^2 \{ K_m K_n [(\omega_1 \omega_2 - \omega_3^2)x + (\omega_1 + \omega_2 - 2\omega_3 - \omega_1 \omega_2 + \omega_3^2)] \} + L_F \{ K_m [(\omega_2 - 1)x^n - \omega_2 x^{n+1}] + K_n [(\omega_1 - 1)x^m - \omega_1 x^{m+1}] \} + x^{n+m+1} - x^{n+m} = 0 \quad (28)$$

This is a second-degree polynomial with respect to the free protein concentration, L_F , which has a physically

acceptable analytical solution as

$$L_F = \left\{ \frac{[-p - (p^2 - 4qr)^{0.5}]}{2q} \right\}_{x_1} \quad (29)$$

evaluated at x_1 , where

$$p = K_n [(\omega_1 - 1)x^m - \omega_1 x^{m+1}] + K_m [(\omega_2 - 1)x^n - \omega_2 x^{n+1}] \quad (30a)$$

$$q = K_m K_n [(\omega_1 \omega_2 - \omega_3^2)x + (\omega_1 + \omega_2 - 2\omega_3 - \omega_1 \omega_2 + \omega_3^2)] \quad (30b)$$

and

$$r = x^{n+m+1} - x^{n+m} \quad (30c)$$

For a given set of interaction parameters (site sizes, intrinsic binding constants, and cooperativity parameters), L_F in eq 29 is a sole function of x_1 . Also, substituting eq 29 into eq 27 renders $\sum \nu_i$ as a sole function of x_1 . Thus, eq 27 (with L_F replaced by eq 29) and eq 29 constitute a set of two parametric equations that completely define the considered binding process with a single variable parameter, x_1 . Both parametric equations provide an easy way to perform computer simulations or fittings of the binding isotherms by treating x_1 as an independent variable.

The physical nature and the range of values of x_1 can be obtained by inspection of eq 24, which shows that in the limit of an infinite lattice, x_1 is an *effective partition function of a single monomer of the lattice* (nucleotide). Therefore, x_1 , as an effective partition function, can only assume real values from $(1, \infty)$, where $x_1 = 1$ corresponds to the free lattice, that is, when $L_F = 0$. Thus, computer simulations can be performed by first introducing x_1 from $(1, \infty)$ into eq 29 and then calculating the corresponding value of L_F . Next, one can introduce x_1 and the corresponding L_F into eq 27 and, thus, obtain the required value of the total average binding density, $\sum \nu_i$. In general, introducing x_1 and L_F into eq 26 provides any average property of the lattice–ligand system, Ψ , in which the lattice residues can exist in three different states.⁹²

To illustrate the behavior of a ligand–lattice system in which the large protein ligand can bind the nucleic acid lattice in two binding modes differing in the number of occluded nucleotides, a series of theoretical binding isotherms for different values of the intrinsic binding constants in both binding modes is shown in Figure 9a. The binding isotherms have been generated using eqs 27 and 29 in the manner described above. The protein is assumed to have the site size of the first binding mode, $n = 16$, and the site size of the second binding mode, $m = 5$ (see below). The binding is characterized by weak positive cooperativity with $\omega_1 = 3$, $\omega_2 = 3$, and $\omega_3 = 5$. The value of $K_m = 10^5 \text{ M}^{-1}$ and K_n assumes different values. The plots show that the high site size binding mode dominates the binding process at low protein concentrations, even when its intrinsic binding constant, K_n , is close to or lower than the intrinsic binding constant of the low site size binding mode. As a result, the plot has clear biphasic character. When the intrinsic binding constant for the high site size binding mode is higher than the intrinsic binding constant for the low site size mode the

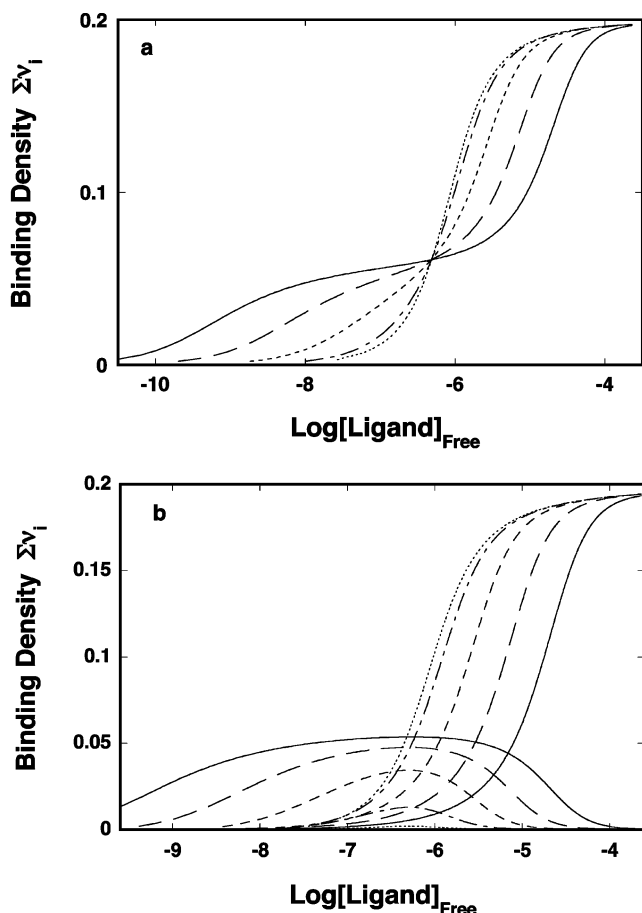


Figure 9. (a) Theoretical binding isotherms of a large ligand in two cooperative binding modes differing by the number of the occluded nucleotides to an infinite homogeneous nucleic acid generated using eqs 27 and 29 in the text. Binding of the ligand in the first mode is characterized by the site size $n = 16$, cooperativity parameter $\omega_1 = 3$, and different values of the intrinsic binding constant, K_{16} : (—) $K_{16} = 10^8 \text{ M}^{-1}$; (---) $K_{16} = 10^7 \text{ M}^{-1}$; (- - -) $K_{16} = 10^6 \text{ M}^{-1}$; (- · -) $K_{16} = 10^5 \text{ M}^{-1}$; (···) $K_{16} = 10^4 \text{ M}^{-1}$. Binding of the ligand in the second mode is characterized by the site size $m = 5$, intrinsic binding constant $K_5 = 10^5 \text{ M}^{-1}$, and cooperativity parameter $\omega_2 = 3$. The cooperativity parameter characterizing the interactions between the ligand molecules bound in different binding modes is $\omega_3 = 5$. (b) The dependence of the high site size and low site size binding modes upon the increasing free concentration of the ligand in solution using the same intrinsic binding constants and cooperativity parameters as in panel a. Reprinted with permission from ref 43. Copyright 1998 Elsevier.

protein initially binds predominantly in the high site size binding mode.

The dependence of the partial binding densities corresponding to each binding mode upon the logarithm of the free protein concentration for the different values of intrinsic binding constants is shown in Figure 9b. It is evident that, as the protein concentration increases, the high site size binding mode is gradually replaced by the low site size mode, independent of the value of the intrinsic binding constant for the high site size mode. This results from the large negative entropy factor, which increases with the degree of saturation of the lattice with the ligand and is associated with the difficulty of saturating the lattice with the ligand that has the larger number of occluded nucleotides within its site size.^{11,12} In other words, in the system where a protein can bind the nucleic acid lattice in two different binding modes with a different number of occluded nucleotides at saturation,

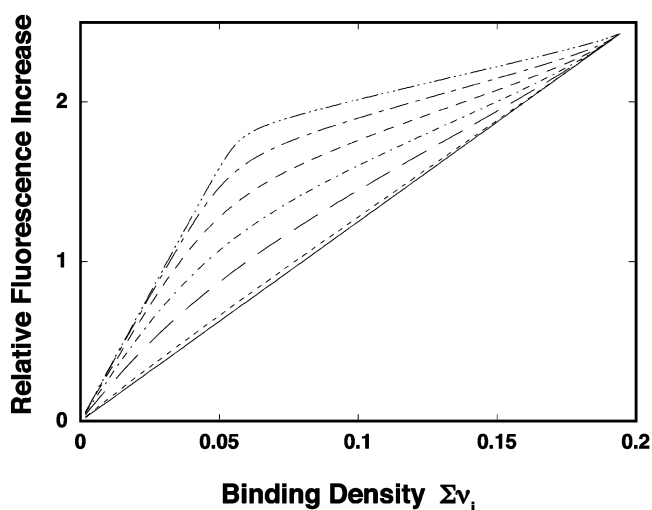


Figure 10. Theoretical dependence of the observed relative fluorescence change, ΔS_{obs} , upon the total binding density, Σv_i , for the binding of a large ligand in two cooperative binding modes differing by the number of the occluded nucleotides to an infinite homogeneous nucleic acid. Binding of the ligand in the first mode is characterized by the site size $n = 16$, cooperativity parameter $\omega_1 = 3$, and different values of the intrinsic binding constant K_{16} : (—) $K_{16} = 10^8 \text{ M}^{-1}$; (---) $K_{16} = 5 \times 10^6 \text{ M}^{-1}$; (- - -) $K_{16} = 10^6 \text{ M}^{-1}$; (- · -) $K_{16} = 3 \times 10^5 \text{ M}^{-1}$; (---) $K_{16} = 10^5 \text{ M}^{-1}$; (···) $K_{16} = 10^4 \text{ M}^{-1}$. Binding of the ligand in the second mode is characterized by the site size, $m = 5$, intrinsic binding constant $K_5 = 10^5 \text{ M}^{-1}$, and cooperativity parameter $\omega_2 = 10$. The cooperativity parameter characterizing the interactions between the ligand molecules bound in different binding modes is $\omega_3 = 5$. The solid line corresponds to the binding of the ligand exclusively in the low site size binding mode. Reprinted with permission from ref 43. Copyright 1998 Elsevier.

the low site size binding mode will always predominate at high protein concentration, independently of the ratio of the intrinsic binding constants.

It is enlightening to examine the theoretical dependence of the observed fluorescence change accompanying the binding of a protein in two different binding modes to the ssDNA, that is, the fractional signal generated from the macromolecule, upon the total average binding density Σv_i of the protein on the nucleic acid. Such dependence is shown in Figure 10. The protein is assumed to bind in two weakly cooperative binding modes with site sizes of $n = 16$ and $m = 5$ and with different intrinsic affinities of high site size binding modes. The selected maximum fluorescence change for the high site size binding mode, ΔS_n , is 2, and the change for the low site size mode, ΔS_m , is 2.5. The plots show characteristic nonlinear behavior. All curves are composed of two binding phases, particularly for the higher intrinsic affinities of the high site size binding mode, where the inflection point between the two phases occurs at the binding density value corresponding to the site size of the high site size binding mode, that is, $\Sigma v_i \approx 0.07$. In fact, the inflection point is still clearly visible, allowing for an estimate of the approximate stoichiometry of the high site size mode, even when the intrinsic binding constant of the high site size mode is only higher by a factor of ~ 3 than the binding constant of the low site size mode (Figure 10). This can be achieved by determining the intersection of the line tangent to the slope of the high site size mode with the line tangent to the slope of the low site size mode.

On the other hand, extrapolation of the high binding density region of the plots in Figure 10 provides the

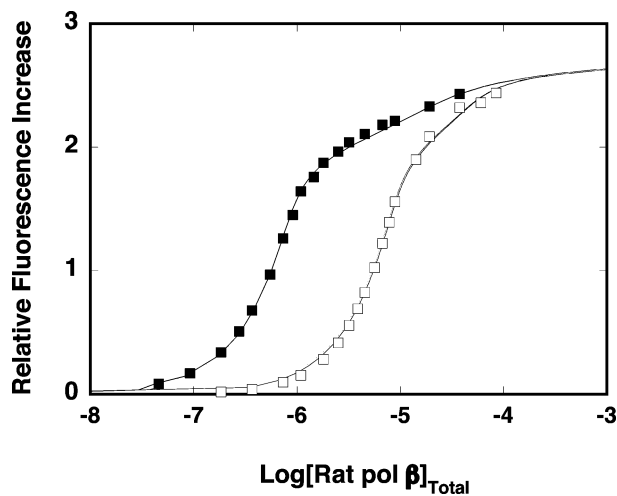


Figure 11. Fluorescence titrations ($\lambda_{\text{ex}} = 325$ nm, $\lambda_{\text{em}} = 410$ nm) of poly(dεA) with rat pol β in sodium cacodylate/HCl (pH 7.0, 10 °C) containing 50 mM NaCl and 1 mM MgCl₂ at two different concentrations of the nucleic acid (nucleotide): (■) 2×10^{-5} M; (□) 1.92×10^{-4} M. The solid lines are nonlinear least-squares fits of the experimental fluorescence binding isotherms, according to the model in which rat pol β binds the ssDNA in two binding modes differing in the number of nucleotides occluded in the protein–nucleic acid complex using the approach described in the text with $n = 16$, $m = 5$, $K_{16} = 2 \times 10^7$ M⁻¹, $K_5 = 3.5 \times 10^5$ M⁻¹, $\omega_1 = 1$, $\omega_2 = 2.5$, and $\omega_3 = 3$ (eqs 27 and 29 in the text). Reprinted with permission from ref 43. Copyright 1998 Elsevier.

stoichiometry of the low site size binding mode. Notice that in the considered case, the slope of the plot corresponding to this mode is not strictly linear, due to the differences in the relative fluorescence changes induced upon the complex formation in both binding modes with the nucleic acid. However, the obtained estimate of the site size is within $\pm 5\%$ of its true value. In the case of the experimental isotherms, such an error is completely absorbed by the inherent error of the data, which allows the number of occluded residues in the low site size binding mode to be determined within $\pm 20\%$ of its true value (see below).

3.1.5.2. Binding of Rat Polymerase β to the ssDNA in Two Binding Modes Differing in the Number of Occluded Nucleotides. Fluorescence titrations of poly(dεA) with rat pol β at two different nucleic acid concentrations are shown in Figure 11. In these solution conditions, the relative maximum increase of the nucleic acid fluorescence reaches the value of 2.6 ± 0.2 . The selected nucleic acid concentrations (nucleotides) provide a clear separation of the binding isotherms up to the relative fluorescence increase of ~ 2.3 , that is, $\sim 85\%$ of the total observed fluorescence change. Figure 12a shows the dependence of the observed relative fluorescence increase as a function of the total average binding density, $\Sigma \nu_i$, of rat pol β . Unlike the binding of the isolated 8-kDa domain of the enzyme (see above), the plot is nonlinear and shows two well-separated binding phases. In the first phase, occurring at low protein concentrations, the binding density reaches the value of 0.06 ± 0.005 , which corresponds to the site size of the enzyme–ssDNA complex of 16 ± 2 nucleotides per bound polymerase molecule. In the second phase, at higher protein concentrations, the binding density extrapolated over the remaining $\sim 15\%$ of the fluorescence change to its maximum value, $\Delta S_{\text{max}} = 2.6 \pm 0.2$ (dashed line), reaches a value of 0.21 ± 0.03 , which corresponds to the site size of 5 ± 1 nucleotides occluded by the protein in the complex.

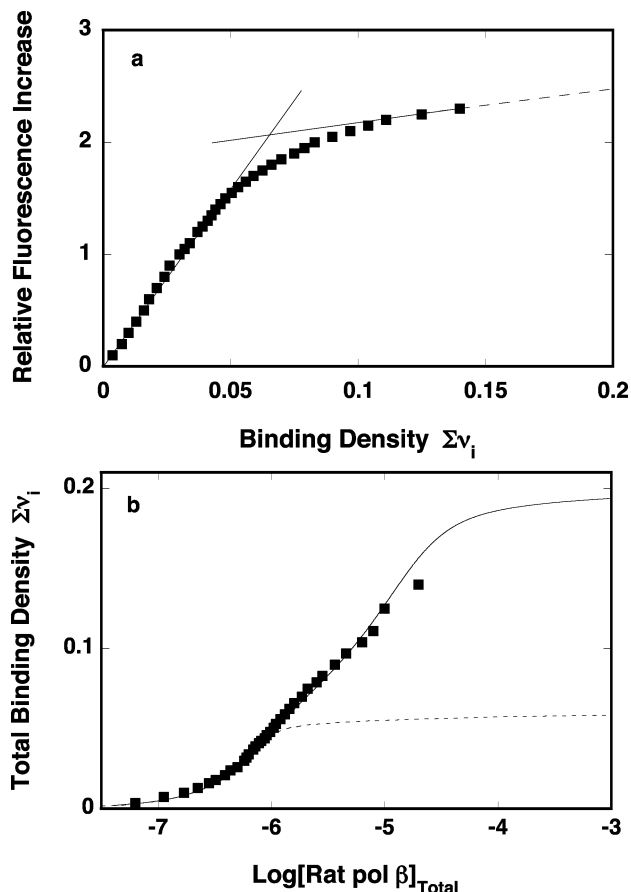


Figure 12. (a) The dependence of the relative fluorescence change upon the total average binding density, $\Sigma \nu_i$, of rat pol β on poly(dεA) in sodium cacodylate/HCl (pH 7.0, 10 °C) containing 50 mM NaCl and 1 mM MgCl₂ (■). The solid lines are tangent to the slopes corresponding to both the low and high binding density phases of the isotherm. The intersection of the lines occurs at $\Sigma \nu_i = 0.06 \pm 0.005$ and indicates the stoichiometry of the high site size binding mode, (pol β)₁₆. The dashed line is an extrapolation of the low affinity binding density phase to the maximum value of the relative fluorescence change, which provides the stoichiometry of the low site size binding mode, (pol β)₅ ($\Sigma \nu_i = 0.22 \pm 0.005$). (b) The dependence of the total average binding density, $\Sigma \nu_i$, of pol β on poly(dεA) as a function of the logarithm of the total enzyme concentration (■). The solid line is the nonlinear least-squares fit of the isotherm obtained by applying the theory and methodology described in the text using the model of two cooperative binding modes, (pol β)₁₆ and (pol β)₅, (eqs 27 and 29) with $K_{16} = 2 \times 10^7$ M⁻¹, $K_5 = 3.5 \times 10^5$ M⁻¹, $\omega_1 = 1$, $\omega_2 = 2.5$, and $\omega_3 = 3$. The dashed line is the computer fit of the initial part of the binding isotherm corresponding to the (pol β)₁₆ mode using the generalized McGhee–von Hippel model (eq 19a) with $K_{16} = 2 \times 10^7$ M⁻¹ and $\omega_1 = 1$. Reprinted with permission from ref 43. Copyright 1998 Elsevier.

These data and analogous studies with the ssDNA oligomers clearly show that rat pol β binds the ssDNA in two binding modes, which differ in the affinities and the number of occluded nucleotides.^{43,44} We designate the higher site size mode as the (pol β)₁₆ and the lower site size mode as the (pol β)₅ binding mode. Additional direct evidence of the presence of two binding modes comes from the thermodynamic binding isotherm, that is, from the plot of the total binding density, $\Sigma \nu_i$, as a function of the logarithm of the total pol β concentration, shown in Figure 12b. The data points are more scattered than in the original fluorescence titration curves due to an error in the determination of $\Sigma \nu_i$. However, there is a discrete intermediate plateau in the

isotherm around $\sum \nu_i \approx 0.07$, indicating heterogeneity of the binding process and reflecting the formation of the $(\text{pol } \beta)_{16}$ binding mode. As the total protein concentration increases, the $(\text{pol } \beta)_{16}$ mode is replaced by the lower site size binding mode, $(\text{pol } \beta)_5$ (see below).

To estimate the intrinsic affinities and cooperativities of rat pol β binding to the ssDNA in different binding modes, the titration curves have been analyzed using the statistical thermodynamic model described above (eqs 27–29). There are nine independent parameters in the model, n , m , K_{16} , K_5 , ω_1 , ω_2 , ω_3 , $\Delta S_{\text{max}16}$, and $\Delta S_{\text{max}5}$, the determination of which from the fluorescence titration curves alone is a hopeless task. The power of the approaches discussed here results from the fact that we have available not only the original fluorescence titration curves (Figure 11) but also the thermodynamic binding isotherm (Figure 12b). Moreover, the site sizes of the two binding modes of the pol β –ssDNA complexes, n and m , are also independently estimated (Figure 12a), leaving the interaction parameters, intrinsic binding constants, K_{16} and K_5 , the cooperativity parameters, ω_1 , ω_2 , and ω_3 , and the fluorescence changes accompanying the binding in two different modes, $\Delta S_{\text{max}16}$ and $\Delta S_{\text{max}5}$, to be determined. This is still a formidable number of parameters, which precludes any attempt to obtain all these quantities in a single fitting procedure of the binding isotherms.

The following strategy can be applied to determine all interaction and spectroscopic parameters for this very complex binding system. Inspection of the thermodynamic isotherm in Figure 12a,b shows that the $(\text{pol } \beta)_{16}$ binding mode is, due to its higher affinity, significantly separated from the $(\text{pol } \beta)_5$ mode with respect to the free protein concentration scale. Such separation of the binding modes allows us to independently obtain estimates of K_{16} and ω_1 , because initially binding in the high site size binding mode completely dominates the association process (see also Figure 9a,b). This is achieved by analyzing the initial part of the thermodynamic isotherm, such as in Figure 12b, where there is an exclusive binding in the $(\text{pol } \beta)_{16}$ mode, using the generalized McGhee–von Hippel isotherm, as defined by eq 19a. The dashed line in Figure 12b is the nonlinear least-squares fit of the initial part of the binding isotherm using the McGhee–von Hippel model with $n = 16$ and K_{16} and ω_1 as fitting parameters (eq 19a). To increase the accuracy, the fitting analysis can be performed using several different isotherms obtained at several different nucleic acid concentrations for the same set of solution conditions.⁴³ With the estimates of K_{16} and ω_1 , the isotherm in Figure 12b is further subjected to a nonlinear least-squares fit to obtain the estimates of the only three remaining interaction parameters, K_5 , ω_2 , and ω_3 , with the determined K_{16} and ω_1 kept as constants. Once all interaction parameters are determined, the spectroscopic parameters are then obtained by directly fitting the fluorescence titration curves with $\Delta S_{\text{max}16}$ and $\Delta S_{\text{max}5}$ as the only remaining fitting parameters. The solid lines in Figure 11 are the nonlinear least-squares fits of the fluorescence titration curves using the interaction parameters obtained from the analysis of the thermodynamic isotherm described above and with $\Delta S_{\text{max}16}$ and $\Delta S_{\text{max}5}$ as fitting parameters.

Although an experienced experimenter would possibly notice that the fluorescence titration curves in Figure 11 already hint at the presence of two binding phases, the existence of the two binding modes of polymerase β and their site sizes could only be firmly established through the

quantitative analysis of the original spectroscopic titrations, leading to the thermodynamic data shown in Figure 12. Further studies using thermodynamic, kinetic, and fluorescence energy transfer approaches provided additional confirmation and characterization of the two binding modes.^{93–97}

The origin of the different binding modes is based in the complex structure of the nucleic acid binding site of the protein.^{28,30,31,70} In the case of pol β , the existence of the two binding modes is a result of the spatial separation of the two DNA-binding subsites of the polymerase with different DNA binding capabilities, located on the two different structural domains of the protein and forming the total DNA-binding site of the enzyme.^{43,44,84} Thus, in the $(\text{pol } \beta)_{16}$ binding mode, the total DNA-binding site of the enzyme is engaged in the complex, that is, both the 8-kDa and the 31-kDa domains are involved in interactions with the ssDNA. In the $(\text{pol } \beta)_5$ binding mode, only the 8-kDa domain is engaged in interactions with the nucleic acid.^{43,44} The structure of pol β is a paradigm of the structure of the DNA-repair polymerase.⁸⁴ Similar organization of the enzyme molecule has been indicated for several other DNA polymerases engaged in DNA repair, although little is known about their nucleic acid binding mechanisms.^{98–100}

Notice that the site size of the $(\text{pol } \beta)_5$ binding mode is rather surprising particularly in the context of the site size, $n = 9 \pm 0.6$, of the complex of the isolated 8-kDa domain with the ssDNA where also only the 8-kDa domain engages in interactions with the DNA, although the intrinsic affinities of both complexes are similar.^{43,44,49} These quantitative data on the stoichiometry and intrinsic affinity of the examined complexes indicate that the orientation of the domain in the complex of the intact enzyme with the ssDNA is different from the orientation of the isolated domain, providing the first indication of the intricate nature of the DNA binding site located on the domain. In other words, they indicate that the 8-kDa domain can engage different regions of its DNA-binding site without affecting the intrinsic affinity of the complex.⁴⁹

3.2. Macromolecular Competition Titration Method (MCT)

3.2.1. Thermodynamic Bases

So far we considered the methods of analysis of the titration curves that allow the experimenter to generate a thermodynamic binding isotherm free of any assumption about the relationship between the observed signal and the total average degree of binding or the total average binding density. However, these analyses require that ligand binding be accompanied by significant changes in the spectroscopic signal originating from the macromolecule. On the other hand, many examined systems may not and, as pointed out above, most often do not have a convenient signal to monitor the binding. For instance, the fluorescence intensity of most of naturally occurring nucleic acids is too low to be useful in conventional measurements. Although proteins usually possess strong tryptophan or tyrosine emissions or both, binding of the nucleic acid may not induce an adequate change of the protein emission to examine complex binding or kinetic mechanisms as occurred in the discussed examples (see above).

One way to overcome this problem is to use a fluorescent derivative of one of the interacting species, for example, the nucleic acid or protein, as discussed above for the DnaB,

PriA, and polymerase β interactions with the ss- and dsDNA. Yet, this may not be feasible in all cases. For instance, there is always a concern that modification introduced on a protein may affect its activity. Modifications on the short fragments of a nucleic acid may create additional, undesirable interaction areas. However, once the interactions with a modified fluorescent system have been characterized (see above), the examination of interactions with unmodified ingredients of the reaction can be performed using competition studies.¹⁰¹ In this context, interactions with unmodified polymer nucleic acids are of particular concern. For instance, while modification of homo-adenosine polymers to obtain fluorescent etheno derivative is a relatively easy reaction, obtaining a homogeneously labeled fluorescent derivative of, for example, poly(dT), is not.^{102,103} For this purpose, the macromolecular competition titration method (MCT) provides a way to quantitatively examine binding of multiple protein ligands to unmodified polymer nucleic acid lattices and, in general, to any unmodified macromolecule with multiple binding sites.^{36,43,44,47,48,49,50,101}

The approach is based on the fact that *the same thermodynamic argument leading to eqs 7–9 applies to situations where titration of a fluorescent nucleic acid with a protein ligand is performed in the presence of the second competing, nonfluorescent nucleic acid lattice.* The protein binds to two different nucleic acids present in solution, but the observed signal originates only from the fluorescent “reference” lattice. As the titration progresses, the saturation of both nucleic acid lattices increases with the increasing free protein concentration in the sample. To illustrate the general behavior of such titrations, a series of theoretical titration curves of a reference fluorescent nucleic acid with the ligand at a constant reference fluorescent nucleic acid concentration but in the presence of different concentrations of a nonfluorescent, competing nucleic acid is shown in Figure 13. The binding isotherms of the protein to both nucleic acid lattices have been generated using the combined application of the generalized McGhee–von Hippel approach and the combinatorial theory for large ligand binding to a linear, homogeneous nucleic acid, described below.^{11,12,101} For simplicity, the protein–lattice complexes for both nucleic acids have been assumed to have a site size of $n = 20$, cooperativity parameter of $\omega = 1$, and intrinsic binding constants of $K = 10^5 \text{ M}^{-1}$ and $K = 10^6 \text{ M}^{-1}$ for the fluorescent reference nucleic acid and the nonfluorescent, competing lattice, respectively; however, the analysis is independent of any particular binding model for both the reference and the unmodified nucleic acid.

Because the measured relative fluorescence increase, ΔS_{obs} , monitors exclusively ligand binding to the reference fluorescent nucleic acid, all curves span the same range of the relative fluorescence change and reach the same plateau value at selected $\Delta S_{\text{max}} = 3.5$. At higher concentrations of the competing unmodified nucleic acid, the titration curves are shifted toward higher total protein ligand concentrations, resulting from the excess of the protein required to saturate the competing nucleic acid lattice. At very high protein ligand concentrations, all curves are superimposed because the excess of the protein bound to the competing lattice becomes negligibly small when compared to the total ligand concentration, L_T .

Recall that the same value of the relative fluorescence change of the reference lattice in the presence of the protein means the same degree of reference lattice saturation, the

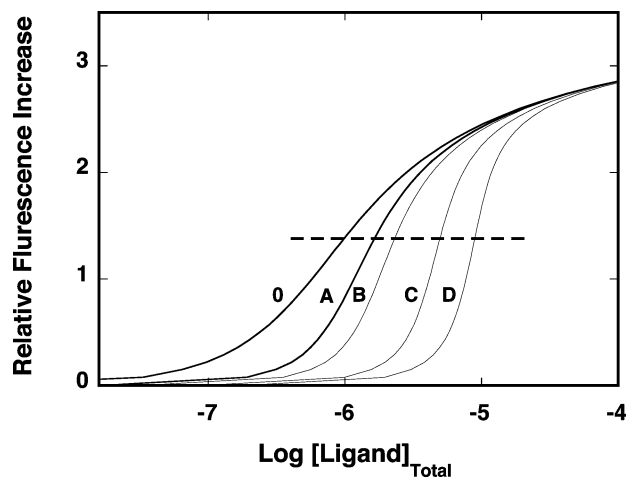


Figure 13. Theoretical fluorescence titration curves of a reference fluorescent nucleic acid with the ligand in the presence of different concentrations of a competing unmodified nucleic acid lattice. Binding of the ligand to the reference lattice is described by the generalized McGhee–von Hippel model for large ligand binding to a linear, homogeneous nucleic acid using intrinsic binding constant $K = 10^5 \text{ M}^{-1}$, cooperativity parameter $\omega = 1$, and site size $n = 20$. The maximum increase of the nucleic acid fluorescence intensity upon saturation with the ligand is $\Delta S_{\text{max}} = 3.5$. Binding of the ligand to competing, nonfluorescent nucleic acid is described by the combinatorial theory using intrinsic binding constant $K = 10^6 \text{ M}^{-1}$, cooperativity parameter $\omega = 1$, and site size $n = 20$. The selected length of the unmodified nucleic acid is 1600 nucleotides. The concentration of the competing nucleic acid (nucleotide) is (0) 0, (A) $2 \times 10^{-5} \text{ M}$, (B) $4 \times 10^{-5} \text{ M}$, (C) $1.2 \times 10^{-4} \text{ M}$, or (D) $2.4 \times 10^{-4} \text{ M}$. The concentration of the reference fluorescent nucleic acid is $2 \times 10^{-5} \text{ M}$ (nucleotide). The horizontal dashed line connects points on all titration curves characterized by the same value of the relative fluorescence increase, ΔS_{obs} . Reprinted with permission from ref 101. Copyright 1996 American Chemical Society.

same protein binding density, $(\sum \nu_i)_R$, on the reference lattice, and the same free protein ligand concentration, L_F . Therefore, in the presence of different concentrations of a competing nucleic acid and a constant concentration of the reference fluorescent lattice at the same value of ΔS_{obs} (e.g., dashed line in Figure 13), the concentration of the free protein ligand, L_F , must be the same and independent of the competing, nonfluorescent nucleic acid concentration. Because the binding density of the protein on the nonfluorescent competing nucleic acid, $(\sum \nu_i)_S$, is also a unique function of L_F , at a given value of ΔS_{obs} corresponding with the same value of L_F , the binding density, $(\sum \nu_i)_S$, must be the same and independent of the total concentration of the competing lattice, N_{TS} . Hence, one can obtain rigorous measurements of the protein total average binding density, $(\sum \nu_i)_S$, on the unmodified competing nucleic acid and the free protein ligand concentration, L_F , from titrations of samples containing constant concentrations of the reference fluorescent nucleic acid, N_{TR} , with the protein in the presence of two or more concentrations of the nonfluorescent, competing nucleic acid lattice.¹⁰¹

This can be accomplished by solving a set of mass conservation equations for the total protein ligand concentration in solution. In the presence of two different competing nucleic acid concentrations, N_{TS1} and N_{TS2} , the total protein concentrations, L_{T1} and L_{T2} , at which the same relative fluorescence change, ΔS_{obs} , is observed are defined as

$$L_{T1} = (\sum \nu_i)_R N_{TR} + (\sum \nu_i)_S N_{TS1} + L_F \quad (31a)$$

and

$$L_{T2} = (\sum \nu_i)_R N_{TR} + (\sum \nu_i)_S N_{TS2} + L_F \quad (31b)$$

Because the binding density, $(\sum \nu_i)_R$, of the protein on the reference fluorescent nucleic acid, at any value of ΔS_{obs} , can be independently determined, subtracting eq 31a from eq 31b and rearranging provides the expression for the true binding density on the competing, unmodified lattice, $(\sum \nu_i)_S$, and the free protein concentration, L_F , in terms of known quantities, $(\sum \nu_i)_R$, total protein, and nucleic acid concentrations as¹⁰¹

$$(\sum \nu_i)_S = \frac{(L_{T2} - L_{T1})}{(N_{TS2} - N_{TS1})} \quad (32a)$$

and

$$L_F = L_{Tx} - (\sum \nu_i)_S N_{TSx} - (\sum \nu_i)_R N_{TR} \quad (32b)$$

where x is 1 or 2.

Notice that although determination of L_F requires $(\sum \nu_i)_R$, the values of $(\sum \nu_i)_S$ are independent of any a priori knowledge of the nature of the binding process and the degree of binding of the protein to the reference fluorescent nucleic acid (eq 32a). Plotting the observed fluorescence change, ΔS_{obs} , as a function of $(\sum \nu_i)_S$, allows one to obtain the stoichiometry of the protein–nonfluorescent nucleic acid complex (see below). Calculations of $(\sum \nu_i)_S$ and, subsequently, L_F can be performed at any value of the observed fluorescence change, ΔS_{obs} , along the titration curves (Figure 13) generating a thermodynamic binding isotherm for protein binding to the competing unmodified nucleic acid. The approach is demonstrated in the next section using the experimental data for the binding of *E. coli* replicative helicase DnaB protein to nonfluorescent poly(dA) in the presence of the reference fluorescent etheno derivative, poly(dεA).

It should be noted that if the ligand affinities for the reference fluorescent nucleic acid and the competing, nonfluorescent lattice are different, the total average binding densities will be different for both nucleic acids at the same value of the measured relative fluorescence change, ΔS_{obs} . Figure 14 shows the theoretical dependence of the observed ΔS_{obs} upon the total average binding density of the protein on the competing unmodified lattice for two different intrinsic affinities of the protein for the competing unmodified nucleic acid. Protein binding to the reference fluorescent nucleic acid is characterized by $K = 10^5 \text{ M}^{-1}$, $\omega = 1$, and $\Delta S_{max} = 3.5$; the ligand binding to the competing lattice is characterized by $K = 10^6 \text{ M}^{-1}$, $\omega = 1$ and $K = 10^4 \text{ M}^{-1}$, $\omega = 1$ for high and low affinity cases, respectively. For simplicity, the dependence of ΔS_{obs} upon $(\sum \nu_i)_R$ for the reference lattice was selected to be strictly proportional (straight dashed line).

In the case where the macroscopic ligand affinity is higher for the competing lattice than for the reference lattice, the total average binding density, $(\sum \nu_i)_S$, of the ligand on the competing, nonfluorescent lattice is higher, when compared to the reference fluorescent lattice at any value of the observed relative fluorescence change, ΔS_{obs} , and the plot is concave down. The opposite is true in the case where the protein ligand affinity is lower for the competing lattice as compared to the reference fluorescent nucleic acid. The plot rises sharply at the initial values of the binding density and levels off at the intermediate range of $(\sum \nu_i)_S$, gradually

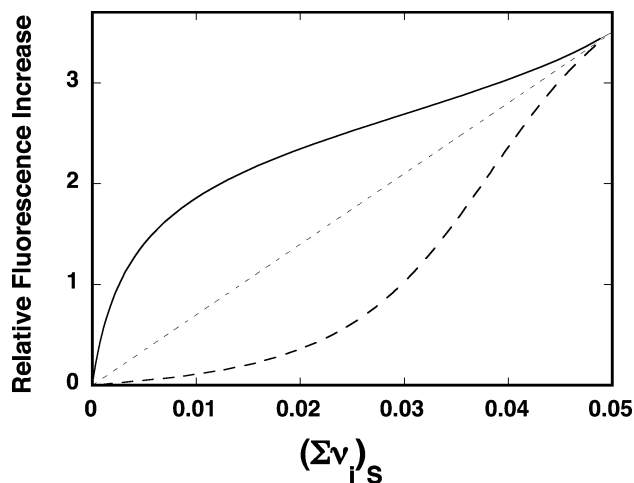


Figure 14. Computer simulation of the dependence of the relative fluorescence increase of the reference fluorescent nucleic acid, ΔF , upon the total average binding density, $(\sum \nu_i)_S$, of the ligand on the competing, nonfluorescent lattice where the competing, nonfluorescent lattice has higher (---) and lower (-·-) affinity toward the protein ligand. The binding of the protein to the reference lattice is described by the generalized McGhee–von Hippel model of large ligand binding to a linear, homogeneous nucleic acid using intrinsic binding constant $K = 10^5 \text{ M}^{-1}$, cooperativity parameter $\omega = 1$, and site size $n = 20$. The maximum increase of the nucleic acid fluorescence intensity upon saturation with the ligand is $\Delta S_{max} = 3.5$. Binding of the ligand to competing, nonfluorescent nucleic acid is described by the combinatorial theory using $\omega = 1$ and $n = 20$; $K = 10^4$ and 10^6 M^{-1} for lower and higher affinity cases, respectively. The selected length of the unmodified nucleic acid is 1600 nucleotides. Reprinted with permission from ref 101. Copyright 1996 American Chemical Society.

approaching the maximum possible value of ΔS_{obs} . This behavior results from the fact that ΔS_{obs} solely reflects the binding density on the reference lattice, $(\sum \nu_i)_R$, which, due to the higher affinity of the reference lattice for the ligand, saturates with the ligand in advance of the saturation of the competing lattice. The plots in Figure 14 indicate that to obtain the most accurate estimation of the stoichiometry of protein–competing nucleic acid complex, the affinity of the competing nonfluorescent lattice for the protein should be similar to or higher than that of the reference lattice.

3.2.2. Specific Application of the MCT Method

To illustrate the MCT method for studying the interactions between proteins and nucleic acids, we analyze the binding of the *E. coli* DnaB helicase to unmodified poly(dA). As mentioned above, binding of poly(dA) and other polydeoxy-nucleotides to the DnaB helicase does not cause any significant change in the protein fluorescence.^{36,101} On the other hand, binding of the DnaB helicase to the fluorescent etheno derivative of poly(dA), poly(dεA), induces a strong, ~ 3.5 -fold, relative fluorescence increase of the nucleic acid, which allows the precise estimation of the stoichiometry and interaction parameters of the DnaB protein–poly(dεA) complex.^{36,101} Thus, poly(dεA) can serve as a reference fluorescent nucleic acid in the MCT method.

Figure 15 shows the fluorescence titration curves of poly(dεA) with the DnaB protein in the absence and presence of two different concentrations of poly(dA). A strong shift of the titration curves toward higher total DnaB concentration, $[\text{DnaB}]_T$, in the presence of poly(dA) indicates efficient competition between the two nucleic acids for the helicase.

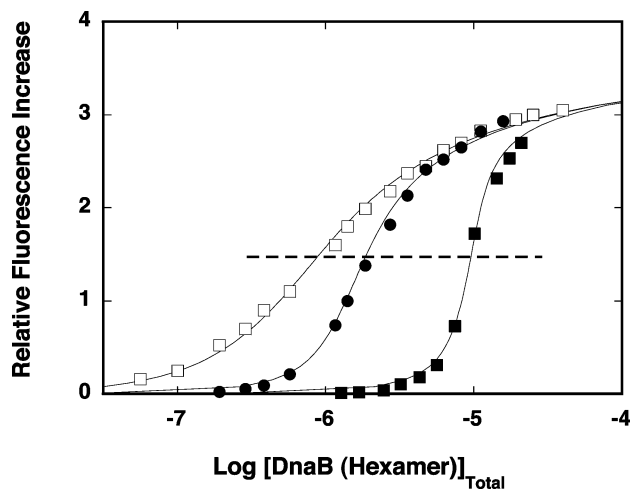


Figure 15. Fluorescence titrations ($\lambda_{\text{ex}} = 325 \text{ nm}$, $\lambda_{\text{em}} = 410 \text{ nm}$) of poly(dεA) with the DnaB helicase in 50 mM Tris/HCl (pH 8.1, 10 °C) containing 50 mM NaCl, 5 mM MgCl₂, and 1 mM AMP-PNP in the presence of different concentrations of poly(dA) (nucleotide): (□) 0; (●) $2.50 \times 10^{-5} \text{ M}$; (■) $2.22 \times 10^{-4} \text{ M}$. The concentration of the poly(dεA) is $2.0 \times 10^{-5} \text{ M}$ (nucleotide). The horizontal dashed line connects points on all titration curves characterized by the same value of the relative fluorescence increase, ΔS_{obs} . The intersection points of the dashed horizontal line with the titration curves in the presence of poly(dA) define the total DnaB concentrations, P_{T1} and P_{T2} , at which the binding density on poly(dεA), $(\sum \nu_i)_R$, the binding density on poly(dA), $(\sum \nu_i)_S$, and the free helicase concentration, L_F , are constant (details in text). The solid lines are nonlinear least-squares fits of the titration curves using the following models. Binding of the helicase to the reference nucleic acid, poly(dεA), is described by the McGhee–von Hippel model, using the independently determined intrinsic binding constant $K = 1.2 \times 10^5 \text{ M}^{-1}$, cooperativity parameter $\omega = 3$, and site size $n = 20$ (eq 19a in text). Binding the enzyme to the competing poly(dA) is described by the combinatorial theory using cooperativity parameter $\omega = 6$, site size $n = 20$, and intrinsic binding constant $K = 1.4 \times 10^6 \text{ M}^{-1}$ (eqs 33 and 34 in text). The selected length of the nucleic acid is 1600 nucleotides. Reprinted with permission from ref 101. Copyright 1996 American Chemical Society.

Recall that at the same value of the relative fluorescence increase, the total average binding density, $(\sum \nu_i)_S$, on the competing lattice, poly(dA), and the free DnaB concentration, $[\text{DnaB}]_F$, must be the same, independent of the concentration of poly(dA) (eqs 31 and 32). Therefore, from this set of titration curves, one can obtain a set of total DnaB concentrations, L_{T1} and L_{T2} , which are determined by the intersection of the horizontal line with either titration curve at which the value of the observed relative fluorescence increase is the same (e.g., horizontal line in Figure 15). Since the total concentrations of poly(dA) are known, one can calculate the true binding density, $(\sum \nu_i)_S$, of the DnaB protein on poly(dA) and the free DnaB concentration, $[\text{DnaB}]_F$, using eqs 31 and 32. This procedure is then repeated over the entire range of ΔS_{obs} , for selected intervals of ΔS_{obs} , providing $(\sum \nu_i)_S$ as a function of $[\text{DnaB}]_F$, thus, enabling the construction of a thermodynamically binding isotherm for DnaB helicase binding to poly(dA), although the signal used to monitor the binding originates exclusively from the reference fluorescent lattice, poly(dεA).

Although it is rather obvious that the site size of the DnaB protein–unmodified nucleic acid is the same as the site size of the protein complex with reference fluorescent nucleic acid, due to the same physical nature of both nucleic acids (ssDNA homopolymers), for completeness, we include the

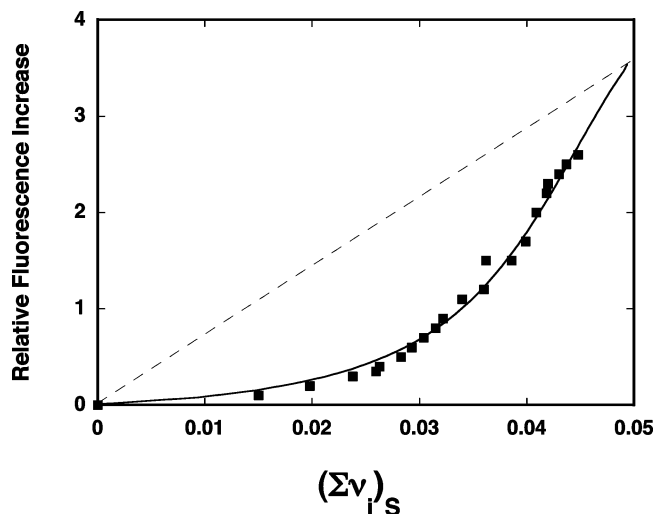


Figure 16. The dependence of the relative fluorescence increase of poly(dεA) upon the total average binding density, $(\sum \nu_i)_S$, of the DnaB helicase on competing poly(dA). The dependence of the relative fluorescence increase of poly(dεA) upon the total average binding density, $(\sum \nu_i)_R$, of the DnaB helicase on poly(dεA) in the same buffer conditions is also included (dashed, straight line). Solid line is the nonlinear least-squares fit using an approach based on the combined McGhee–von Hippel model and the combinatorial theory for binding a large ligand to two different competing homogeneous lattices for the simultaneous binding of the DnaB helicase to poly(dεA) and poly(dA) (details in text). The binding of the helicase to the reference poly(dεA), using the independently determined intrinsic binding constant $K = 1.2 \times 10^5 \text{ M}^{-1}$, cooperativity parameter $\omega = 3$, and site size $n = 20$. Binding the enzyme to the competing poly(dA) is described by the combinatorial theory using cooperativity parameter $\omega = 4.8$, site size $n = 20$, and intrinsic binding constant $K = 1.4 \times 10^6 \text{ M}^{-1}$. Reprinted with permission from ref 101. Copyright 1996 American Chemical Society.

independent determination of this quantity in our discussion.^{36,101} The estimation of the site size of the complex of the protein–nonfluorescent competing lattice can be achieved by plotting the observed relative fluorescence increase, ΔS_{obs} , as a function of the total average binding density, $(\sum \nu_i)_S$, of the DnaB protein on poly(dA). Figure 16 shows the dependence of the observed relative fluorescence change, ΔS_{obs} , upon the total average binding density, $(\sum \nu_i)_S$, of the DnaB helicase on poly(dA). For comparison, the dependence of ΔS_{obs} upon the total average binding density, $(\sum \nu_i)_R$, of the DnaB helicase on the reference lattice, poly(dεA), is also included (dashed straight line).^{36,101} The quantity $(\sum \nu_i)_S$ could be determined up to the value of ~ 0.044 , which corresponds with $\Delta S_{\text{obs}} \approx 2.6$. Extrapolation to the maximum possible value of $\Delta S_{\text{max}} = 3.6 \pm 0.3$ gives the maximum value of $(\sum \nu_i)_S = 0.05 \pm 0.005$ and the estimation of the site size of the poly(dA)–DnaB helicase complex, $n = 20 \pm 3$. As expected, this value is the same as the estimated $n = 20 \pm 3$ in identical solution conditions for the poly(dεA)–DnaB complex.^{36,101}

3.2.3. Direct Analysis of the Experimental Isotherm of Protein Ligand Binding to Two Competing Nucleic Acid Lattices

For binding to a single type of nucleic acid lattice, the generalized McGhee–von Hippel model (eqs 19a and 19b) is the simplest statistical thermodynamic description for cooperative binding of a large ligand to an infinite one-dimensional, homogeneous lattice with overlapping potential

binding sites and in one binding mode.^{11,88} In the case of protein binding to two competing nucleic acid lattices, the ligand interactions with each nucleic acid are described by an equation in the form of expression 19a, that is, the entire binding system is described by two independent isotherms. Any attempt to simultaneously use two isotherms of the type given by eq 19a is hindered by the fact that they are complex polynomial implicit functions of the total average binding density, $\sum \nu_i$, and free ligand concentration. Thus, to simulate, or fit, the competition titration curves of a large ligand binding to two competing nucleic acid lattices, complex and cumbersome numerical calculations are required.

To overcome this problem and to obtain a general method to analyze simultaneous binding of a protein ligand to two or more polymer nucleic acids without resorting to complex numerical calculations, an approach based on the combined application of the generalized McGhee–von Hippel model, as defined by eq 19a, and the combinatorial theory for large ligand binding to a linear, homogeneous lattice has been developed.^{11,12,36,49,88,101} It should be stressed that the McGhee–von Hippel model is equivalent to the combinatorial model as the length of the homogeneous lattice approaches infinity. In the approach, the binding of the protein to the reference fluorescent nucleic acid is described by the McGhee–von Hippel model. Binding of the protein to the competing nonfluorescent lattice is described by the combinatorial theory for cooperative binding of a large ligand, which covers n nucleotides, to a linear finite homogeneous nucleic acid. In the combinatorial model, the partition function of the protein ligand–nucleic acid system, Z_S , is defined by

$$Z_S = \sum_{k=0}^g \sum_{j=0}^{k-1} S_N(k,j) (K_S L_F)^k \omega_S^j \quad (33a)$$

where g is the maximum number of ligand molecules that may bind to the finite nucleic acid lattice (for the nucleic acid lattice N residues long $g = M/n$),¹² K_S is the intrinsic binding constant characterizing interactions with the unmodified lattice, ω_S is the corresponding cooperative interaction parameter, k is the number of ligand molecules bound, and j is the number of cooperative contacts between k bound ligand molecules in a particular configuration on the lattice. The combinatorial factor $S_N(k,j)$ is the number of distinct ways that k ligands bind to a lattice with j cooperative contacts and is defined by

$$S_N(k,j) = \frac{[(M - nk + 1)!(k - 1)!]}{[(M - nk + j + 1)!(k - j)!j!(k - j - 1)!]} \quad (33b)$$

The total average binding density, $(\sum \nu_i)_S$, is then obtained by using the standard statistical thermodynamic expression, $(\sum \nu_i)_S = \partial \ln Z_S / \partial \ln L_F$, as¹²

$$(\sum \nu_i)_S = \frac{\sum_{k=1}^g \sum_{j=0}^{k-1} k S_N(k,j) (K_S L_F)^k \omega_S^j}{\sum_{k=0}^g \sum_{j=0}^{k-1} S_N(k,j) (K_S L_F)^k \omega_S^j} \quad (34)$$

Expressions 33 and 34 describe the binding of a large ligand to a finite, linear homogeneous lattice. For a long enough lattice, the obtained isotherm will be, within experi-

mental accuracy, indistinguishable from the isotherm generated using the generalized eq 19a for binding the large ligand to the infinite lattice. We found that in the case of a protein like the DnaB helicase ($n = 20$), a lattice that can accommodate ≥ 40 protein molecules (800 nucleotides) represents an “infinite” lattice for any practical purpose. Contrary to the generalized McGhee–von Hippel model (eq 19a), expression 34 is an explicit function of the free ligand concentration, which allows us to calculate the binding density directly for the known K , ω , n , and L_F . Because free protein ligand concentrations can be explicitly calculated using an infinite lattice model through eq 19a, combining both equations offers a simple way of fitting the simultaneous binding of a large ligand to two (or more) competing, different linear lattices, for example, a reference fluorescent nucleic acid in the presence of a competing, unmodified nucleic acid.

This is accomplished by first applying eq 19a to the reference fluorescent nucleic acid, calculating the free ligand concentration, L_F , for given values of the parameters K , ω , and n with $(\sum \nu_i)_R$ as the independent variable. Subsequently, the obtained value of L_F is introduced into eq 34, which is used to describe the protein binding to a competing, nonfluorescent nucleic acid, and the binding density, $(\sum \nu_i)_S$, is calculated for the given K_S , ω_S , and n_S , which characterize the binding of the protein to the competing nonfluorescent lattice. The calculations are repeated for the entire range of the $(\sum \nu_i)_R$, generating the required $(\sum \nu_i)_S$ for the competing nucleic acid lattice as a function of L_F . The experimental binding isotherm, which is the observed relative fluorescence change, ΔS_{obs} , as a function of the total protein concentration, L_T , is then obtained by calculating for each value of ΔS_{obs} the total protein concentration L_T by introducing $(\sum \nu_i)_S$, $(\sum \nu_i)_R$, and L_F into eq 31a or 31b.

The solid lines in Figure 15 are nonlinear least-squares fits of the simultaneous binding of the DnaB helicase to two competing nucleic acids, using the procedure outlined above. The binding of the DnaB helicase to the fluorescent poly(dεA) has been independently determined and described by using the generalized McGhee–von Hippel (eq 19a).¹⁰¹ The binding of the protein to competing poly(dA) has been described using the combinatorial theory (eqs 33 and 34). Because the site size of the DnaB–poly(dεA) complex, $n = 20 \pm 3$, has been independently estimated (see above), there are only two parameters, the intrinsic binding constant, K_S , and the cooperativity parameter, ω_S , to be determined. It should be pointed out that the analysis of the simultaneous binding of a large ligand to competing nucleic acid lattices could also be performed by applying eqs 33 and 34 of the combinatorial theory to both the reference and the competing nucleic acids. This approach is completely equivalent to the combined application of the McGhee–von Hippel and combinatorial models described above. However, the advantage of using the combined McGhee–von Hippel and the combinatorial theories lies in the tremendously decreased computational time, particularly if a very long lattice is used.

3.2.4. Competition Titration Method Using a Single Concentration of a Nonfluorescent Nucleic Acid

The method described above allows the determination of thermodynamic binding parameters for ligand binding to a nonfluorescent nucleic acid by performing fluorescence titrations of a reference fluorescent nucleic acid with the protein at a constant reference nucleic acid concentration in

the presence of two or more concentrations of the competing nucleic acid. However, the thermodynamic binding density, $(\sum \nu_i)_S$, and the free ligand concentration, L_F , can also be obtained by using a single titration of the fluorescent nucleic acid at a single concentration of a competing nonfluorescent nucleic acid in direct reference to the titration curve of the fluorescent reference nucleic acid alone.^{49,101} This results from the fact that, independently of the presence of a competing nonfluorescent nucleic acid, the same value of the relative fluorescent change, ΔS_{obs} , reflects the same value of the ligand binding density on the fluorescent reference lattice, $(\sum \nu_i)_R$, and, in turn, the same free ligand concentration, L_F . Thus, in Figure 15, the same fluorescence increase of poly(dεA) upon binding the DnaB helicase corresponds to the same helicase binding density on poly(dεA) in the presence and absence of the competing poly(dA). Therefore, one can obtain $(\sum \nu_i)_S$ and $[DnaB]_F$ by simultaneously analyzing titration curves 1 and 2 or 1 and 3, instead of 2 and 3 in Figure 15. This approach can be particularly useful if the amount of the competing nonfluorescent nucleic acid is limited.¹⁰¹

For the analysis using a single concentration of the nonfluorescent nucleic acid, the total protein concentrations at which the same value of the relative fluorescence increase, ΔS_{obs} , of the reference fluorescent nucleic acid is observed in the absence and presence of competing nonfluorescent lattice, L_{TR} , and L_{TS_x} , respectively, are defined as

$$L_{\text{TR}} = (\sum \nu_i)_R N_{\text{TR}} + L_F \quad (35a)$$

$$L_{\text{TS}_x} = (\sum \nu_i)_R N_{\text{TR}} + (\sum n_i)_S N_{\text{TS}_x} + L_F \quad (35b)$$

where x can be 1 or 2 (Figure 15). Solving the set of eqs 35a and 35b for $(\sum \nu_i)_S$ provides

$$(\sum \nu_i)_S = \frac{(L_{\text{TS}_x} - L_{\text{TR}})}{(N_{\text{TS}_x})} \quad (36a)$$

and

$$L_F = L_{\text{TS}_x} - (\sum \nu_i)_S N_{\text{R}} - (\sum \nu_i)_R N_{\text{S}_x} \quad (36b)$$

3.2.5. Competition Fluorescence Titrations of Short Fluorescent Oligonucleotides in the Presence of a Competing Polymer Nucleic Acid

The MCT analysis does not have to be confined to a polymer reference lattice, and in some cases it can be simplified in terms of mathematics necessary to analyze the isotherms, as well as experimental procedures, by using short fragments of nucleic acid as a reference lattice, as long as both nucleic acids compete for the same binding site on the protein. A short nucleic acid fragment optimal for such analysis would be an oligomer that forms a simple 1:1 complex with the protein. The obvious choice for a short reference fluorescent nucleic acid lattice is a nucleic acid fragment long enough to exactly span the total site size of the protein–nucleic acid complex. Such an oligomer binds to the same binding site as a polymer lattice and occupies the entire total binding site. A selection of the length of the fluorescent reference fragment can be based on the initial estimation of the stoichiometry of the protein–nucleic acid complex obtained using a modified fluorescent polymer nucleic acid or a series of oligomers (see above).

In such studies, the degree of saturation of the nucleic acid oligomer with the protein ligand, $\sum \Theta_R$, and the fluorescence change, $(\Delta S_{\text{obs}})_o$, accompanying the formation of the complex are described by¹⁰¹

$$\sum \Theta_R = \frac{K_o L_F}{(1 + K_o L_F)} \quad (37a)$$

and

$$(\Delta S_{\text{obs}})_o = (\Delta S_{\text{max}})_o \left[\frac{K_o L_F}{(1 + K_o L_F)} \right] \quad (37b)$$

where $(\Delta S_{\text{max}})_o$ is the maximum observed fluorescence change of the short nucleic acid lattice upon saturation with the protein and K_o is the macroscopic binding constant of the oligomer to the protein.

In the presence of a competing polymer nucleic acid, the observed fluorescence change of the short nucleic acid fragment is thermodynamically linked with the protein total average binding density, $(\sum \nu_i)_S$, on the competing polymer through the free protein ligand concentration, L_F . However, as in the case of the reference polymer nucleic acid, the same value of $(\Delta S_{\text{obs}})_o$ (eq 37b) at different competing unmodified nucleic acid concentrations corresponds with the same degree of saturation of the short nucleic acid fragment and the same free concentration of the protein, L_F , that is, also the same $(\sum \nu_i)_S$. Therefore, performing two fluorescence titrations of a short oligomer nucleic acid with the protein at the same total concentration of the reference oligomer, O_{TR} , but in the presence of two different total concentrations of the competing nucleic acid, N_{TS_1} and N_{TS_2} , allows the determination of the total average binding density, $(\sum \nu_i)_S$, and free protein concentration, L_F , by applying the same method as described above for a polymer reference lattice. In the case of the short reference lattice, eqs 31a and 31b take the form of¹⁰¹

$$L_{\text{T1}} = (\sum \Theta_R) O_{\text{TR}} + (\sum \nu_i)_S N_{\text{TS}_1} + L_F \quad (38a)$$

$$L_{\text{T2}} = (\sum \Theta_R) O_{\text{TR}} + (\sum \nu_i)_S N_{\text{TS}_2} + L_F \quad (38b)$$

and the total average binding density, $(\sum \nu_i)_S$, is described by eq 32a. Analogously, if a single fluorescence titration of a reference oligomer with the protein in the presence of a competing polymer lattice at the concentration N_{TS_x} is used in direct reference to the fluorescence titration of the reference oligomer alone with the protein, eqs 38a and 38b become¹⁰¹

$$L_{\text{TR}} = (\sum \Theta_R) O_{\text{TR}} + L_F \quad (39a)$$

$$L_{\text{TS}_x} = (\sum \Theta_R) O_{\text{TR}} + (\sum \nu_i)_S N_{\text{TS}_x} + L_F \quad (39b)$$

and the total average binding density, $(\sum \nu_i)_S$, is described by eq 36a.

3.3. Signal Used to Monitor the Interactions Originates from the Ligand

3.3.1. Thermodynamic Bases

Analyses and examples described so far dealt with the experimental design where the binding processes were monitored by following the signal from the macromolecule

(nucleic acid), that is, “normal” titrations were performed.^{6,10,42} In the case, where some spectroscopic property (e.g., fluorescence intensity) of the ligand changes upon binding to the macromolecule and hence the signal changes monitor the apparent degree of saturation of the ligand with the macromolecule, a “reverse” titration (addition of a macromolecule to a constant ligand concentration) is generally performed. Several protein–nucleic acid systems have been analyzed by examining the changes of the fluorescence of a protein ligand titrated with a nucleic acid.^{5,9,13,14,17,23,24} In fact, a similar approach that allowed us to transform spectroscopic titration curves into thermodynamic binding isotherms for the “normal” titration method has been developed for the “reverse” titration method in studies of the binding of *E. coli* SSB protein to the single-stranded nucleic acids.⁹

Consider the general case of multiple-ligand binding to a macromolecule where there can be i states of the bound ligand with each state possessing a different molar signal, S_i . The observed signal, ΔS_{obs} , from the ligand solution, at the total concentration of the ligand, L_T , in the presence of the macromolecule at the total concentration, N_T , has contributions from the free ligand and the ligand bound to the macromolecule in any of its i possible bound states as

$$S_{\text{obs}} = S_F L_F + \sum S_i L_i \quad (40)$$

where S_F and L_F are the molar signal and concentration of the free ligand, respectively, and S_i and L_i are the molar signal and concentration of the ligand bound in state i , respectively. Expression 40 is valid when the molar signal of each species is independent of concentration (i.e., in the absence of ligand and macromolecule aggregation). The concentrations of the free and bound ligand are related to the total ligand concentration by the conservation of mass equation

$$L_T = L_F + \sum L_i \quad (41a)$$

where

$$L_i = \Theta_i N_T \quad (41b)$$

The quantity, Θ_i , is the partial degree of binding of the ligand in the i th state. Substituting eqs 41a and 41b into eq 40 and rearranging provides

$$S_{\text{obs}} - S_F L_T = N_T \sum (S_i - S_F) \Theta_i \quad (42)$$

Notice that the product $S_F L_T$ is the initial signal from the protein ligand solution before titration with the macromolecule. Dividing both sides of eq 42 by $S_F L_T$ and next multiplying by (L_T/N_T) yields

$$\left[\frac{(S_{\text{obs}} - S_F L_T)}{S_F L_T} \right] \left(\frac{L_T}{N_T} \right) = \sum \left[\frac{(S_i - S_F)}{S_F} \right] \Theta_i \quad (43)$$

This relationship can be rewritten as

$$\frac{(S_{\text{obs}} - S_F L_T)}{S_F L_T} \left(\frac{L_T}{N_T} \right) = \sum (\Delta S)_i \Theta_i \quad (44)$$

and

$$\Delta S_{\text{obs}} \left(\frac{L_T}{N_T} \right) = \sum (\Delta S)_i \Theta_i \quad (45)$$

Notice that although we used the total average degree of binding, $\sum \Theta_i$, to arrive at eq 45, an identical relationship can be obtained for the total average binding density, $\sum \nu_i$, for the ligand binding to a long homogeneous polymer lattice.^{6,9} It should be pointed out that $\Delta S_{\text{obs}} = (S_{\text{obs}} - S_F L_T)/(S_F L_T)$ is the experimentally observed fractional change in the spectroscopic signal from the ligand (with respect to the signal of the free ligand at the same total concentration, $S_F L_T$) in the presence of the ligand and the macromolecule at total concentrations L_T and N_T . The quantity $(\Delta S)_i = (S_i - S_F)/S_F$ is the molar signal change characterizing the ligand when it is bound in a state i . Expression 45 is general and independent of the spectroscopic method used to monitor the interactions.^{6,9}

The relationship expressed by eq 45 indicates that the quantity $\Delta S_{\text{obs}}(L_T/N_T)$ is equal to $\sum (\Delta S)_i \Theta_i$, the sum of the partial degrees of binding for all i states of the ligand–macromolecule system weighted by the intrinsic signal change for each bound state. The weighting factors $(\Delta S)_i$ are molecular intensive quantities, which are constant for a particular binding state i under a given set of experimental conditions (temperature, buffer, etc.). Therefore, the quantity $\sum (\Delta S)_i \Theta_i$ is constant for a given distribution of the ligand among different possible states, that is, the total average degree of binding, $\sum \Theta_i$, in the case of the binding to a set of discrete binding sites and the total binding density, $\sum \nu_i$, in the case of the protein ligand binding to a polymer nucleic acid. Therefore, at equilibrium, the values of L_F and $\sum \Theta_i$ are constant for a given value of $\Delta S_{\text{obs}}(L_T/N_T)$, independent of the macromolecule concentration, N_T . Hence, under identical solution conditions, one can obtain thermodynamically rigorous measurements of $\sum \Theta_i$ and L_F from plots of $\Delta S_{\text{obs}}(L_T/N_T)$ vs N_T for two or more titrations performed at different total ligand concentrations. This is accomplished by obtaining the set of concentrations (L_T , N_T) from each titration for which the quantity of $\Delta S_{\text{obs}}(L_T/N_T)$ is constant and solving for $\sum \Theta_i$ and L_F . The procedure is therefore analogous to the case in which a signal from the macromolecule is monitored during the “normal” titration (see above). The quantity $\Delta S_{\text{obs}}(L_T/N_T)$ is referred to as the ligand binding density function (LBDF).^{6,9,41,42}

3.3.2. Analysis of Spectroscopic Titration Curves When the Signal Used to Monitor the Interactions Originates from the Ligand

Binding of the fluorescent analogue of ADP, ϵ ADP, to the six nucleotide-binding sites of the *E. coli* DnaB protein hexamer can serve as an example of the analysis of a complex multiple-ligand binding system where the interactions are followed by monitoring the signal from the ligand.⁴¹ Thus, in these studies, the fluorescent nucleotide analogue, ϵ ADP, is the ligand, the DnaB helicase is a macromolecule, and the fluorescence of the analogue is used to monitor the binding. However, the fluorescence of the ϵ ADP is increased only by $\sim 21\%$ upon binding to the DnaB helicase.⁴¹ To increase this signal change and, in turn, to obtain a higher resolution of the titration experiments, acrylamide has been added to the solution. Acrylamide is a very efficient dynamic quencher of several biologically relevant chromophores including the etheno derivative of adenosine.^{104,105} This extra dynamic quenching process, which does not affect the

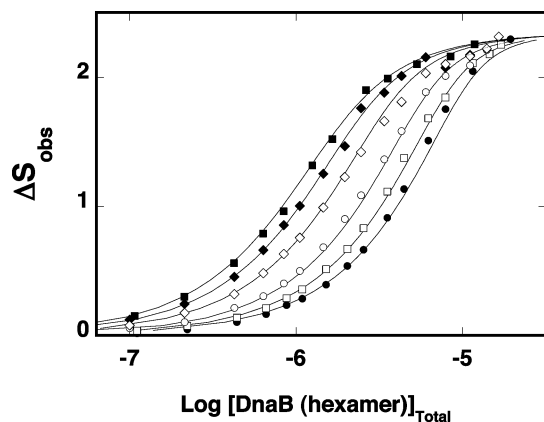


Figure 17. Fluorescence titrations (“reverse titrations”) of ϵ ADP at different concentrations of the nucleotide with the DnaB helicase in 50 mM Tris/HCl (pH 8.1, 10 °C) containing 100 mM NaCl and 100 mM acrylamide ($\lambda_{\text{ex}} = 325$ nm, $\lambda_{\text{em}} = 410$ nm). The concentrations of the nucleotide are (■) 4×10^{-6} M, (◆) 6×10^{-6} M, (◇) 1×10^{-5} M, (○) 2×10^{-5} M, (□) 3×10^{-5} M, and (●) 4×10^{-5} M. Solid lines are nonlinear least-squares fits of the experimental binding isotherms according to the hexagon model using a single set of binding parameters: $K = 4 \times 10^5 \text{ M}^{-1}$, $\sigma = 0.4$, and $\Delta S_{\text{max}} = 2.35$. Reprinted with permission from ref 41. Copyright 1997 Elsevier.

thermodynamics of the nucleotide–enzyme interactions, is much less efficient for the nucleotide bound to the DnaB protein than for the free nucleotide, leading to a much larger change in the nucleotide fluorescence upon formation of the complex with the protein, thus, increasing the resolution of the titration curves.⁴¹ The application of the differential dynamic quenching of the ligand or the macromolecule fluorescence to increase the resolution of the binding experiments is thoroughly discussed in ref 41.

A series of fluorescence titration curves of ϵ ADP with the DnaB protein at different nucleotide concentrations is shown in Figure 17. At higher nucleotide concentrations, the curves are shifted toward higher concentrations of the DnaB helicase, because more of the enzyme is necessary to saturate the increased amount of ϵ ADP in solution. All curves reach the same plateau of the relative fluorescence increase, ΔS_{obs} , at saturating concentrations of the protein. As we pointed out above, in general, the fractional change of the ligand fluorescence upon the macromolecule concentration does not necessarily strictly correspond to the fractional ligand saturation. This is never a priori known for any multiple ligand binding system. However, the estimate of the degree of binding and the free ligand concentrations can be obtained by using the LBDF approach.^{6,9,41,42}

Figure 18 shows the plot of $\Delta S_{\text{obs}}(L_T/N_T)$ as a function of the DnaB protein concentration obtained using the fluorescence titrations presented in Figure 17. For the different total concentrations of ϵ ADP at the same value of the binding density function $\Delta S_{\text{obs}}(L_T/N_T)$, the total degree of binding, $\Sigma\Theta_i$, and the free ϵ ADP concentrations must be the same, thus, allowing for their determination (eq 45). Notice that even though all of the “reverse” titrations shown in Figure 17 span the same full range of ϵ ADP fluorescence increase, they do not span the same range of the degree of binding as seen from the LBDF plot in Figure 18. Therefore, multiple titrations at different values of L_T are always required to span the full range of the total average degree of binding. This is very different from the “normal” titrations where, in favorable conditions, often only two titration curves are sufficient to obtain $\sim 90\%$ of the binding isotherm (e.g., Figure 2a,b). A

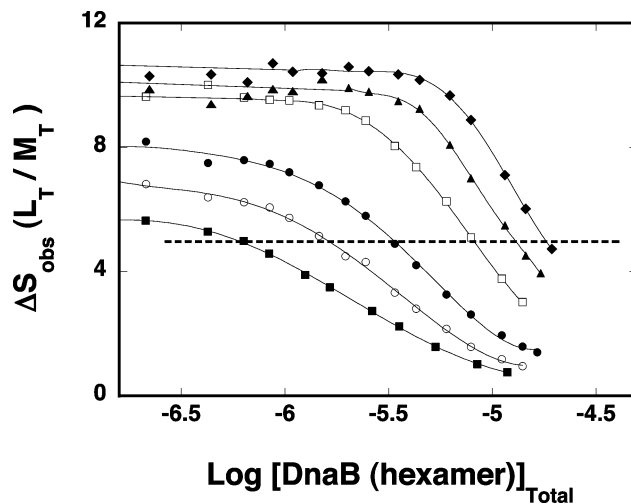


Figure 18. Dependence of the binding density function, $\Delta S_{\text{obs}}(L_T/M_T)$, on the logarithm of the total DnaB protein concentration at different concentrations of ϵ ADP: (■) 4×10^{-6} M; (○) 6×10^{-6} M; (●) 1×10^{-5} M; (□) 2×10^{-5} M; (▲) 3×10^{-5} M; (◆) 4×10^{-5} M. Solid lines separate different data sets and do not have theoretical basis. The horizontal dashed line connects points at the selected, same value of the binding density function at different ϵ ADP concentrations at which $[\epsilon\text{ADP}]_{\text{free}}$ and the degree of binding, $\Sigma\Theta_i$, of the nucleotide on the DnaB hexamer are the same (details in text). Reprinted with permission from ref 41. Copyright 1997 Elsevier.

horizontal line that intersects the LBDF curves is drawn, defining a constant value of $\Delta S_{\text{obs}}(L_T/N_T)$ (horizontal dashed line in Figure 18). The points of intersection of the horizontal line with each LBDF curve determine the set of values (N_T , L_T) for which L_F and $\Sigma\Theta_i$ are constant, as shown in Figure 18 for one constant value of the LBDF. Based on eq 1, the average degree of binding, $\Sigma\Theta_i$, and L_F can then be determined from the slope and intercept of a plot of L_T vs N_T at each constant value of $\Delta S_{\text{obs}}(L_T/N_T)$. By repetition of this procedure for a series of horizontal lines that span the range of values of $\Delta S_{\text{obs}}(L_T/N_T)$ as a function of L_F , the values of $\Sigma\Theta_i$ can be obtained and the thermodynamic binding isotherm can be determined.^{6,42}

In the construction of a series of LBDF plots, as shown in Figure 18, one should cover as wide a range of ligand concentrations as possible; however, care should be taken to avoid large changes in the total ligand concentration between two successive titrations, since this may bias the determination of L_F and $\Sigma\Theta_i$.^{9,41,42} In our experience, six to eight titrations using successive total ligand concentrations that differ by a factor of 1.5–2 will generate an accurate set of data.

The model-independent binding isotherm constructed from the full analysis of the data in Figures 17 and 18 for the binding of ϵ ADP to the DnaB protein is plotted in Figure 19. Recall that the DnaB helicase is a homohexamer that has six nucleotide binding sites.^{76–80} The plot in Figure 19 shows that in the examined concentration range of ϵ ADP the total average degree of binding, $\Sigma\Theta_i$, could be reliably determined from ~ 0.5 up to the value of ~ 4.5 , that is, from $\sim 10\%$ to $\sim 80\%$ of the entire range of the degree of binding of the DnaB hexamer–nucleotide complex.⁴¹ The solid line is a theoretical isotherm, according to the hexagon model, which describes the binding of six ligands to a short circular lattice of six discrete sites, characterized by the intrinsic binding constant, K , and cooperativity parameter, σ (see below).^{37,41,106} The obtained intrinsic binding constant

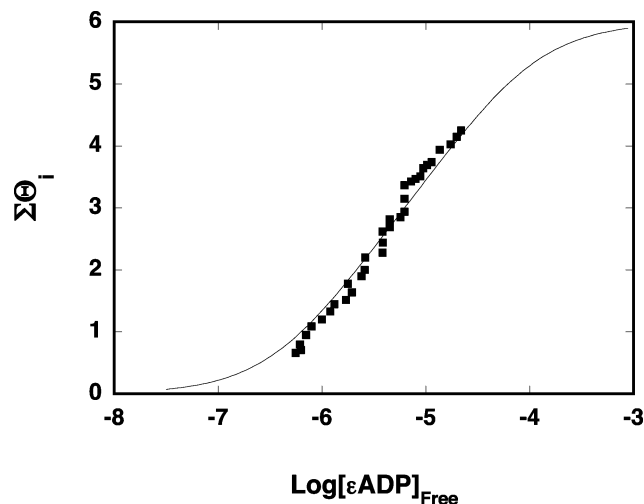


Figure 19. Thermodynamic binding isotherm for the ϵ ADP–DnaB binding system, that is, the dependence of the total average degree of binding of ϵ ADP on the DnaB helicase hexamer upon the logarithm of the free nucleotide concentrations, $[\epsilon\text{ADP}]_{\text{free}}$. The solid line is the theoretical binding isotherm according to the hexagon model using intrinsic binding constant $K = 4 \times 10^5 \text{ M}^{-1}$ and cooperativity parameter $\sigma = 0.4$. Reprinted with permission from ref 41. Copyright 1997 Elsevier.

$K = 4 \times 10^5 \text{ M}^{-1}$ and the cooperativity parameter $\sigma = 0.4$ are, within experimental accuracy, the same as the values that have been independently obtained using the quantitative fluorescence titration method in which the quenching of the protein fluorescence has been used to monitor the interactions.^{37,41}

3.3.3. Correlation between the Fractional Signal Change, ΔS_{obs} , and the Average Degree of Binding, $\Sigma\Theta_i$

Measurements of the total average degree of binding, $\Sigma\Theta_i$, as a function of the free ligand concentration enable one to determine the relationship between the observed signal change and the fraction of the bound ligand. Then, if ΔS_{obs} is found to be directly proportional to the fraction of the bound ligand (L_B/L_T), the binding parameters can be obtained with much greater ease from a titration at a single ligand concentration. The fractional fluorescence increase of ϵ ADP as a function of the fraction of ϵ ADP bound to the DnaB protein, (L_B/L_T) = $[(\Sigma\Theta_i)N_T]/L_T$, is shown in Figure 20. It is clear that in the examined range of the total average degree of binding, there is a strict linearity between the relative increase of the ϵ ADP fluorescence, ΔS_{obs} , and the fraction of the bound nucleotide, L_B/L_T . However, it is important to check the relationship between the observed fluorescence change (or any signal used to monitor the binding) and L_B/L_T over a wide range of the degree of binding, since the signal change can, in general, be dependent upon $\Sigma\Theta_i$ (see above). In the case of the DnaB protein– ϵ ADP interactions, this direct proportionality holds for $\Sigma\Theta_i$ up to ~ 5.3 .⁴¹ Therefore, in this particular case, the fractional fluorescence increase is indeed equal to the fraction of the bound nucleotide, that is, $\Delta S_{\text{obs}}/\Delta S_{\text{max}} = L_B/L_T$. Although in the considered ϵ ADP–DnaB system we determined the maximum value of the relative fluorescence change, ΔS_{max} (Figure 17), the value of ΔS_{max} was not necessary for determining the thermodynamic isotherm. For some systems, this may be an important aspect of the quantitative analyses where the accurate determination of ΔS_{max} is not possible

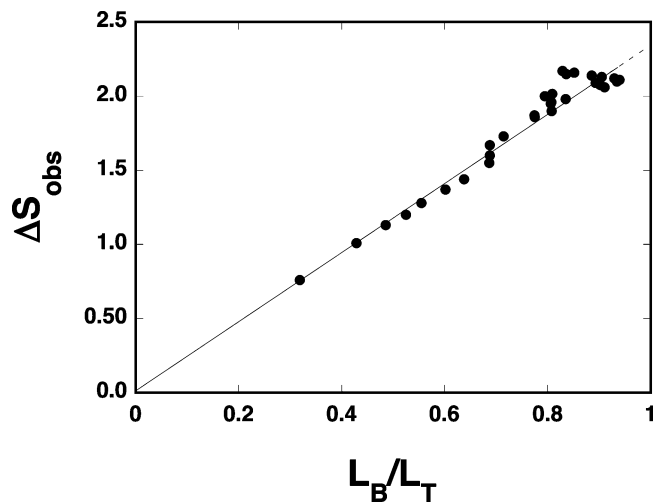


Figure 20. Dependence of the relative fluorescence increase upon the fractional saturation of the nucleotide for ϵ ADP binding to the DnaB helicase. The selected concentration of ϵ ADP is $2 \times 10^{-5} \text{ M}$. The concentration of the bound nucleotide has been calculated from $[\epsilon\text{ADP}]_B = (\Sigma\Theta_i)[\text{DnaB}]_T$. Reprinted with permission from ref 41. Copyright 1997 Elsevier.

due to the low affinity of the ligand. In fact, when ΔS_{obs} is directly proportional to L_B/L_T , one can determine the maximum extent of the observed signal, ΔS_{max} , from a linear extrapolation of a plot of ΔS_{obs} vs L_B/L_T to $L_B/L_T = 1$, which the theoretical limiting value of the ligand concentration ratio as shown in Figure 20. Short extrapolation to $L_B/L_T = 1$ provides the value of the maximum relative increase, $\Delta S_{\text{max}} = 2.35$, of the relative increase of the ϵ ADP fluorescence upon saturation with the DnaB protein in studied solution conditions in excellent agreement with the observed value of ΔS_{obs} (Figure 17).⁴¹

3.3.4. Analysis of a Binding Isotherm Using a Single Titration Curve When $\Delta S_{\text{obs}}/\Delta S_{\text{max}} = L_B/L_T$

The analysis using ligand binding density function allows one to rigorously determine a model-independent binding isotherm and to determine the relationship between the values of ΔS_{obs} and the fraction of the bound ligand, L_B/L_T . The LBDF approach is time-consuming, since six to eight titrations are required to construct a single precise binding isotherm over a wide range of binding densities, which is necessary if the relationship between ΔS_{obs} and L_B/L_T is not known a priori. However, if it is determined from the LBDF analysis that a linear relationship exists between ΔS_{obs} and L_B/L_T over a wide range of a degree of binding (as determined in the case of the ϵ ADP–DnaB protein system), then one can use this relationship to determine the average degree of binding and the free ligand concentration from a single titration curve.^{9,41} For such a simple case, eq 45 reduces to

$$\frac{\Delta S_{\text{obs}}}{\Delta S_{\text{max}}} = \frac{L_B}{L_T} \quad (46a)$$

Then

$$L_F = \left(\frac{\Delta S_{\text{max}} - \Delta S_{\text{obs}}}{\Delta S_{\text{max}}} \right) L_T \quad (46b)$$

and

$$\sum \Theta_i = \left(\frac{\Delta S_{\text{obs}}}{\Delta S_{\text{max}}} \right) \left(\frac{L_T}{N_T} \right) \quad (46c)$$

Thus, a single titration can be used to obtain $\sum \Theta_i$ or $\sum \nu_i$ as a function of L_T . However, once again, we stress that one should not simply assume that the fractional signal change is equal to the fraction of the bound ligand since, if it is not true, this could lead to significant errors. On the other hand, if direct proportionality does not exist between the signal change and the fraction of the bound ligand over a wide range of binding densities or degrees of binding, the thermodynamic binding isotherm can still be constructed without any assumptions using the LBDF analysis as depicted in Figure 19.

3.4. Empirical Function Approach

Once the thermodynamic isotherm for the ligand binding to a macromolecule is determined, it can then be fitted to extract binding parameters using a statistical thermodynamic model that is based on the physical properties of the studied system. This can be done without any considerations as to the specific values of molar spectroscopic signals originating from various complexes.^{9,34,36,37,41–44,49} The analysis and fitting of the original fluorescence titration curves is much more involved because spectroscopic parameters of all complexes enter the equations that describe the binding model. The approach discussed in this section allows the experimenter to avoid this difficult or often impossible task and fit any spectroscopic titration curves, once the binding model is formulated.^{41,106–108} Moreover, a high-resolution spectroscopic titration curve spans close to the entire range of the degree of binding for a given binding process, while a reliable determination of the total average degree of binding, $\sum \Theta_i$, is usually limited to the range between $\sim 10\%$ and 85% of the maximum stoichiometry of the complex. Thus, direct fitting the spectroscopic curve is preferable for more accurate estimate of the binding parameters.

We will consider here a case as applied to the “normal” titration approach, that is, the binding of the ligand is followed by monitoring the signal from the macromolecule. In analogy to eq 7b, the observed spectroscopic signal that is used to monitor the ligand binding is defined in a general way in terms of the spectroscopic parameters, ΔS_i , characterizing each possible i complex and the partial degree of binding, Θ_i , or binding densities as

$$\Delta S_{\text{obs}} = \left(\sum \Delta S_{\text{max}} \right)_i \Theta_i \quad (47a)$$

or, using average binding density as

$$\Delta S_{\text{obs}} = \sum (\Delta S_{\text{max}})_i \nu_i \quad (47b)$$

Thus, to fit a spectroscopic titration curve all molecular spectroscopic parameters $(\Delta S_{\text{max}})_i$ in eqs 47a and 47b must be known. In the simplest case, all $(\Delta S_{\text{max}})_i$ could be the same, that is, the plots of ΔS_{obs} as a function of the total average degree of binding, $\sum \Theta_i$, or total average binding density, $\sum \nu_i$, are strictly linear, for example, for the PriA helicase or rat pol β 8-kDa domain binding to the ssDNA 40-mer and the polymer ssDNA, respectively (Figures 7 and 8). For such systems, the spectroscopic titration curve can be directly fitted to extract binding parameters. In some

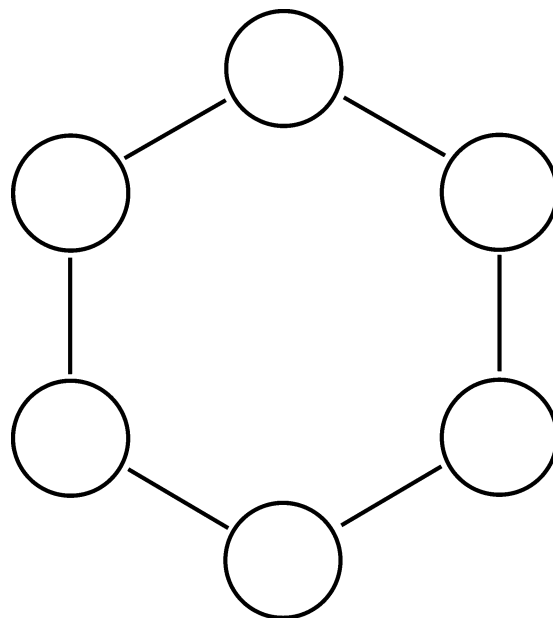


Figure 21. A schematic representation of the hexagon model of the RepA hexamer. The subunits of the hexamer are depicted as circles. The lines connecting each subunit with its two neighbors symbolize the number of possible cooperative interactions (two in the hexagon model).

situations, the behavior of the system provides information about the values of $(\Delta S_{\text{max}})_i$, for example, binding of the polymerase β to the dsDNA 10-mer (Figure 3), where the two binding phases are well separated on the ligand concentration scale, allowing an independent determination of the spectroscopic parameters characterizing each complex.⁸³ However, obviously, these situations do not apply in every case, as illustrated below for the binding of the ATP analogue TNP-ATP to the RepA hexameric helicase of the plasmid RSF1010.

The broad host nonconjugative plasmid RSF1010 confers bacterial resistance to sulfonamides and streptomycin.^{109,110} RSF1010 codes its own replicative helicase, the RepA protein. The RepA helicase unwinds the duplex DNA in the $5' \rightarrow 3'$ direction and is essential for the RSF1010 plasmid replication in bacterial cells.¹¹¹ As revealed by the crystallographic studies, the enzyme is a homohexameric helicase with a ringlike structure and a central cross-channel with the diameter of $\sim 17 \text{ \AA}$.^{112,113} The schematic ringlike structure of the RepA hexamer is depicted in Figure 21. From a statistical thermodynamic point, the protein is a short circular lattice and can be described using a hexagon model with only two interaction parameters, the intrinsic binding constant, K , and the nearest-neighbor cooperativity parameter, σ .^{37,106–108} The partition function that describes the nucleotide binding to the six binding sites of the hexamer, according to the hexagon model, is defined as

$$Z_H = 1 + 6x + 3(3 + 2\sigma)x^2 + 2(1 + 6\sigma + 3\sigma^2)x^3 + 3(3\sigma^2 + 2\sigma^3)x^4 + 6\sigma^4x^5 + \sigma^6x^6 \quad (48a)$$

and the total average degree of binding, $\sum \Theta_i$, is

$$\sum \Theta_i = [6x + 6(3 + 2\sigma)x^2 + 6(1 + 6\sigma + 3\sigma^2)x^3 + 12(3\sigma^2 + 2\sigma^3)x^4 + 30\sigma^4x^5 + 6\sigma^6x^6] / Z_H \quad (48b)$$

where $x = KL_T$, the product of the intrinsic binding constant, K , and the free nucleotide concentration, L_T .

Thus, the RepA hexamer with a given number of nucleotide molecules bound, can exist in multiple configurations, for example, the hexamer with three ligands bound has 20 possible configurations (eq 48). Although some of these configurations are physically indistinguishable, resulting from simple statistical effects of binding three molecules to the initially equivalent and independent six sites, there are configurations that differ by the number of cooperative interactions, and they may also differ by the values of the individual molecular spectroscopic parameters, $(\Delta S_{\max})_i$. For instance, the minimum analytical relationship between the observed experimental signal, ΔS_{obs} , and the individual molar spectroscopic parameters, $(\Delta S_{\max})_i$, and the interaction parameters, K and σ , is

$$\Delta S_{\text{obs}} = (6\Delta S_1 x + 6(3\Delta S_{21} + 2\sigma\Delta S_{22})x^2 + 6(\Delta S_{31} + 6\sigma\Delta S_{32} + 3\sigma^2\Delta S_{33})x^3 + 12(3\sigma^2\Delta S_{41} + 2\sigma^3\Delta S_{42})x^4 + 30\sigma^4\Delta S_5 x^5 + 6\sigma^6\Delta S_6 x^6)/Z_H \quad (49)$$

For clarity, the subscript max has been omitted at individual molar quenching constants $(\Delta S_{\max})_i$. Expression 49 contains 10 optical parameters. Thus, for example, there are three physically distinguishable configurations of the hexamer with three ligands bound that have different densities of the cooperative interactions; therefore, there are three possible different molar spectroscopic parameters, ΔS_{31} , ΔS_{32} , and ΔS_{33} , characterizing each configuration with different density of cooperative interactions. Notice that, at the saturating concentration of the ligand, the observed experimental maximum signal change, is $\Delta S_{\text{obs}} = \Delta S_6$. Nevertheless, to apply eq 49 to obtain interaction parameters K and σ from a single spectroscopic titration curve, all 10 optical constants, $(\Delta S_{\max})_i$, must be known. In practice, this is a hopeless task, even for the simplest possible set of spectroscopic parameters of the RepA hexamer–nucleotide system expressed by eq 49.

Binding of the unmodified nucleotide cofactors to the RepA hexamer is not accompanied by a change of the protein fluorescence that is adequate to perform quantitative analysis of the complex binding process. However, we found that binding of nucleotide analogues TNP-ATP and TNP-ADP to the RepA protein is accompanied by a strong quenching of the protein fluorescence, providing an excellent signal to monitor the association. Fluorescence titrations of the RepA hexamer with TNP-ADP at three different protein concentrations are shown in Figure 22a. The maximum quenching of the protein fluorescence at saturation is 0.87 ± 0.03 . The selected protein concentrations provide the separation of the binding isotherms up to the quenching value of ~ 0.83 .

The dependence of the observed fluorescence quenching, ΔF_{obs} , upon the average degree of binding, $\sum \Theta_i$, of TNP-ADP on the RepA hexamer is shown in Figure 22b. The values of $\sum \Theta_i$ are obtained by averaging the values obtained from analyses of three different possible combinations of the titration curves in Figure 22a. The separation of the binding isotherms allows us to obtain the values of $\sum \Theta_i$ up to ~ 5.1 TNP-ADP molecules per RepA hexamer. The plot in Figure 22b is clearly nonlinear. For comparison, the dashed line represents the hypothetical case when a strict proportionality between the degree of cofactor binding and the quenching of the RepA hexamer fluorescence would

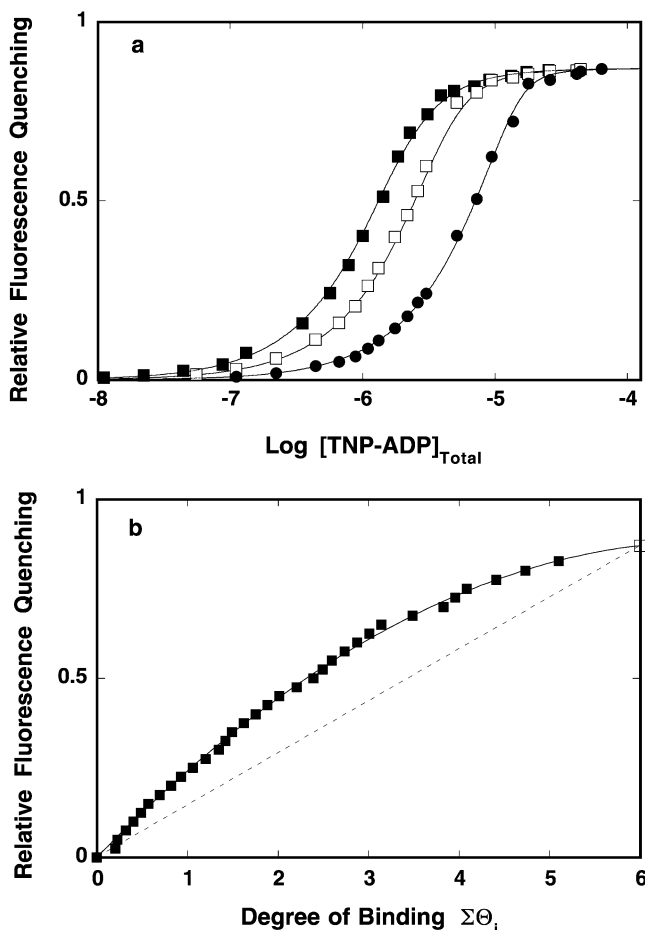


Figure 22. (a) Fluorescence titration of the RepA helicase with TNP-ADP in 50 mM Tris/HCl (pH 7.6, 10 °C) containing 10 mM NaCl and 1 mM MgCl₂ at different RepA protein concentrations: (■) 5×10^{-7} M; (□) 1×10^{-6} M; (●) 3×10^{-6} M (hexamer). The solid lines are nonlinear least-squares fits of the titration curves according to the hexagon model (eqs 48a and 48b) using a single set of binding parameters with the intrinsic binding constant $K = 8 \times 10^6 \text{ M}^{-1}$ and cooperativity parameter $\sigma = 0.36$. (b) Dependence of the relative fluorescence quenching, ΔS_{obs} , upon the average degree of binding of TNP-ADP on the RepA hexamer, $\sum \Theta_i$. The solid line is the nonlinear least-squares fit using the second-degree polynomial function defined by eq 51. The dashed line is the hypothetical dependence of ΔS_{obs} upon $\sum \Theta_i$ that assumes a strict linear relationship between the observed fluorescence quenching and the average degree of binding. The maximum value of $\Delta S_{\text{max}} = 0.87 \pm 0.03$. Reprinted with permission from ref 107. Copyright 2005 American Chemical Society.

exist. Short extrapolation to the maximum quenching, $\Delta F_{\text{max}} = 0.87 \pm 0.03$, shows that at saturation the RSF1010 RepA hexamer binds 6.0 ± 0.3 molecules of TNP-ADP.^{107,108}

It is obvious that to directly analyze the spectroscopic titration curves, using analytical eqs 48–49, one would have to know all molar fluorescence intensities of all possible RepA–nucleotide complexes. The plot in Figure 22b does not provide any information as to what the values of these parameters should be, with the exception of ΔS_6 (see above). However, the problem of finding all optical parameters can be avoided by using the following empirical function approach. This method can be used for any ligand–macromolecule systems where the determination of all optical constants in an analytical equation is practically impossible. The approach is based on introducing the representation of the observed change in the spectroscopic signal of the macromolecule, ΔS_{obs} , upon the binding of the ligand as a

function of the total average degree of binding, $\Sigma\Theta_i$, or binding density, $\Sigma\nu_i$, via an empirical function.¹⁰⁶ For this approach to succeed, high-resolution, quantitative determination of the total average degree of binding or binding density is absolutely necessary. The empirical function is usually a polynomial that accurately relates the experimentally determined dependence of the spectroscopic parameter, ΔS_{obs} , to the experimentally determined total average degree of binding, $\Sigma\Theta_i$, defined as

$$\Delta S_{\text{obs}} = \sum_{j=0}^n a_j (\Sigma\Theta_i)^j \quad (50)$$

where a_j are the fitting constants. This function is then used to generate a theoretical isotherm for a binding model and to extract intrinsic binding parameters for a particular binding model from the experimentally obtained single titration curve. This is accomplished by first calculating the value of the total average degree of binding, $\Sigma\Theta_i$, for a given free nucleotide concentration and initial estimates of the binding parameters for a given statistical thermodynamic model. Then, the obtained $\Sigma\Theta_i$ is introduced into eq 50 and the value of ΔS_{obs} corresponding to a given value of $\Sigma\Theta_i$ is obtained. These calculations are then performed for the entire titration curve.^{106–108}

In the case of TNP-ADP binding to the RepA hexamer, the plot of the observed relative fluorescence quenching, ΔS_{obs} , as a function of the total average degree of binding, $\Sigma\Theta_i$, in Figure 22b is described by a second-degree polynomial function with the coefficients $a_1 = 2.6054 \times 10^1$ and $a_2 = -1.9220 \times 10^2$, respectively, as

$$\Delta S_{\text{obs}} = 0.26054 \times 10^1 (\Sigma\Theta_i) - 0.019224 \times 10^2 (\Sigma\Theta_i)^2 \quad (51)$$

This function is then used to fit and generate theoretical titration curves of the TNP-ADP binding to the RepA helicase, using the hexagon model, and to extract intrinsic binding constant, K , and cooperativity parameter, σ , in the manner described above. The solid lines in Figure 22a are the nonlinear least-squares fits of the experimental isotherms for TNP-ADP binding to the RepA hexamer using eqs 48a, 48b, and 51 with a single set of binding parameters that provide the intrinsic binding constant $K = (8 \pm 1.5) \times 10^6 \text{ M}^{-1}$ and $\sigma = 0.36 \pm 0.05$.¹⁰⁷

Because the total average degree of binding, $\Sigma\Theta_i$, has been determined over $\sim 80\%$ of the spectroscopic titration curves of RepA with TNP-ADP, one can construct a true thermodynamic binding isotherm for the system, that is, the plot of $\Sigma\Theta_i$ as a function of the logarithm of the free nucleotide concentration. This plot is shown in Figure 23.¹⁰⁷ The solid line in Figure 23 is the nonlinear least-squares fit of the thermodynamic isotherm for TNP-ADP binding to the RepA hexamer using directly eqs 48a and 48b with only two fitting parameters, K and σ , without giving any consideration to the fluorescence changes used to obtain the isotherm. The fit provides the intrinsic binding constant $K = (7.7 \pm 1.5) \times 10^6 \text{ M}^{-1}$ and $\sigma = 0.38 \pm 0.05$. These values are in excellent agreement with the same binding parameters obtained using the empirical function method (Figure 22a).^{107,108}

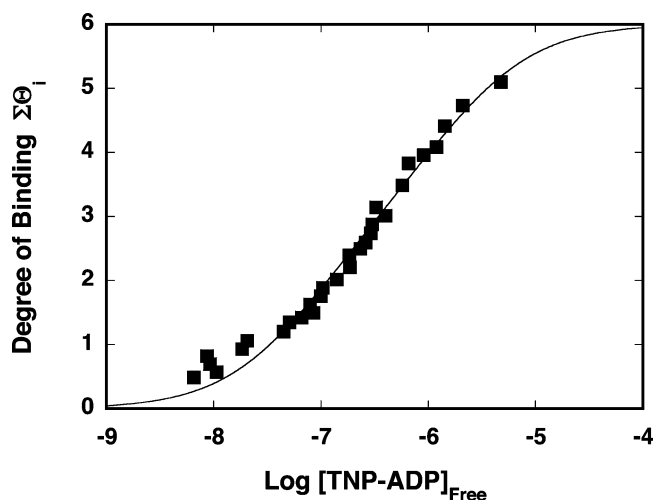


Figure 23. The total average degree of binding of TNP-ADP on the RepA hexamer as a function of the free concentration of the cofactor in 50 mM Tris/HCl (pH 7.6, 10 °C) containing 10 mM NaCl and 1 mM MgCl₂. The isotherm has been constructed using the quantitative method described in text. The solid line is a nonlinear least-squares fit according to the hexagon model (eq 48) with the intrinsic binding constant $K = 7.7 \times 10^6 \text{ M}^{-1}$ and cooperativity parameter $\sigma = 0.38$. Reprinted with permission from ref 107. Copyright 2005 American Chemical Society.

3.5. Stopped-Flow Kinetic Studies of Protein–Nucleic Acid Interactions

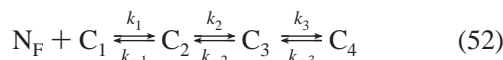
3.5.1. General Analysis of Relaxation Times and Amplitudes Using the Matrix Projection Operator Technique

Spectroscopic stopped-flow kinetic measurements of the approach to equilibrium of the protein–nucleic acid complexes provide *two independent* sets of data, the relaxation times and amplitudes, characterizing the normal modes of the observed relaxation processes.^{94–97,114–120} In fact, the characteristic behaviors of the relaxation times and the amplitudes as functions of the ligand or macromolecule concentration serve as diagnostics as to what the mechanism of the observed reaction is. It is rather unfortunate that the analysis of the amplitudes of the observed relaxation processes is often not considered in the literature on the kinetics of the protein–nucleic acid interactions, and the conclusions about the behavior of the interacting system are based exclusively on the analysis of the relaxation times only. The quantitative studies of the kinetics of the reaction, including the mechanistic details and the nature of the formed intermediates, require the examination of both the relaxation times and the amplitudes.^{94–97,114–120}

In our analyses of the stopped-flow kinetics of the protein–nucleic acid interactions, we use the matrix projection operator technique.¹²¹ As we show below, the matrix projection operator technique is extremely useful for the analysis of complex stopped-flow kinetics, particularly by providing closed-form expressions for the amplitudes of the studied reaction. This, in turn, allows the experimenter to obtain structural information about all identified intermediates.

To illustrate the approach, as an example, we consider a complex sequential reaction between a nucleic acid, N,

and protein, C, of the type



where the initial bimolecular process is followed by two isomerization reactions of the formed complex and the reaction is monitored by the changes of the fluorescence, F, of the nucleic acid. The reaction is characterized by three relaxation times and three amplitudes, that is, there are three normal modes of the reaction.¹¹⁴ The differential equations describing the time course of reaction 52, in terms of different protein species, are

$$\begin{aligned} \frac{d[\text{C}_1]}{dt} &= -k_1[\text{C}_1][\text{N}_F] + k_{-1}[\text{C}_2] \\ \frac{d[\text{C}_2]}{dt} &= k_1[\text{C}_1][\text{N}_F] - (k_{-1} + k_2)[\text{C}_2] + k_{-2}[\text{C}_3] \\ \frac{d[\text{C}_3]}{dt} &= k_2[\text{C}_2] - (k_{-2} + k_3)[\text{C}_3] + k_{-3}[\text{C}_4] \\ \frac{d[\text{C}_4]}{dt} &= k_3[\text{C}_3] - k_{-3}[\text{C}_4] \end{aligned} \quad (53)$$

To eliminate all higher order terms in differential equations, the kinetic studies are performed in pseudo-first-order conditions, that is, in a large excess of the nucleic acid, $[\text{N}_T] \gg [\text{C}_T]$, that is, $[\text{N}_T]$ is approximately constant during the reaction. In matrix notation, system 53 is then defined as

$$\begin{pmatrix} \frac{d[\text{C}_1]}{dt} \\ \frac{d[\text{C}_2]}{dt} \\ \frac{d[\text{C}_3]}{dt} \\ \frac{d[\text{C}_4]}{dt} \end{pmatrix} = \begin{pmatrix} -k_1[\text{N}_T] & k_{-1} & 0 & 0 \\ k_1[\text{N}_T] & -(k_{-1} + k_2) & k_{-2} & 0 \\ 0 & k_2 & -(k_{-2} + k_3) & k_3 \\ 0 & 0 & k_3 & -k_{-3} \end{pmatrix} \begin{pmatrix} [\text{C}_1] \\ [\text{C}_2] \\ [\text{C}_3] \\ [\text{C}_4] \end{pmatrix} \quad (54)$$

and

$$\dot{\mathbf{C}} = \mathbf{M}\mathbf{C} \quad (55)$$

where $\dot{\mathbf{C}}$ is a vector of the time derivatives, \mathbf{M} is the coefficient matrix, and \mathbf{C} is a vector of concentrations. In standard matrix approach, the solution of the system 55 is

$$\mathbf{C} = \exp(\mathbf{M}t)\mathbf{C}_0 \quad (56)$$

and

$$\mathbf{C} = \mathbf{V} \begin{pmatrix} \exp(\lambda_0 t) & 0 & 0 & 0 \\ 0 & \exp(\lambda_1 t) & 0 & 0 \\ 0 & 0 & \exp(\lambda_2 t) & 0 \\ 0 & 0 & 0 & \exp(\lambda_3 t) \end{pmatrix} \mathbf{V}^{-1} \mathbf{C}_0 \quad (57)$$

where $\lambda_0, \lambda_1, \lambda_2,$ and λ_3 are eigenvalues of matrix \mathbf{M} , \mathbf{V} is a matrix whose columns are the eigenvectors of matrix \mathbf{M} , and \mathbf{C}_0 is the vector of the initial concentrations of the different protein species. In the considered sequential reaction 52, \mathbf{C}_0 is a column vector $([\text{C}_T], 0, 0, 0)$ where $[\text{C}_T]$ is the total concentration of the protein. The form of the vector \mathbf{C}_0 reflects the fact that at $t = 0$ the concentration of the free protein is equal to its total concentration, while the concentrations of all other species are zero. To solve system 55, that is, to find relaxation times and amplitudes of the reaction, first one has to obtain the eigenvalues of matrix \mathbf{M} , then the corresponding eigenvectors. For a multistep mechanism, like the one considered here, this can be achieved only through cumbersome numerical analyses, particularly for the eigenvectors.

However, instead of finding eigenvectors corresponding to the each eigenvalue, λ_i , of matrix \mathbf{M} , we expanded the matrix $\exp(\mathbf{M}t)$ using its eigenvalues, $\exp(\lambda_i t)$, and corresponding projection operators, \mathbf{Q}_i , as^{121,122}

$$\exp(\mathbf{M}t) = \sum_{i=0}^3 \mathbf{Q}_i \exp(\lambda_i t) \quad (58)$$

The projection operators, \mathbf{Q}_i , can easily be defined by Sylvester's theorem¹²² using the original coefficient matrix \mathbf{M} and its eigenvalues, λ_i . In general, a projection operator, \mathbf{Q}_i , corresponding to an eigenvalue, λ_i , is¹²²

$$\mathbf{Q}_i = \frac{\prod_{j \neq i}^n (\mathbf{M} - \lambda_j \mathbf{I})}{\prod_{j \neq i}^n (\lambda_i - \lambda_j)} \quad (59)$$

where n is the number of eigenvalues and \mathbf{I} is the identity matrix of the same size as \mathbf{M} . In the considered reaction, there are four eigenvalues $\lambda_0, \lambda_1, \lambda_2,$ and λ_3 ; however, $\lambda_0 = 0$ because of the mass conservation in the reaction system. Therefore, using eq 59, one obtains¹²²

$$\mathbf{Q}_0 = \frac{(\mathbf{M} - \lambda_1 \mathbf{I})(\mathbf{M} - \lambda_2 \mathbf{I})(\mathbf{M} - \lambda_3 \mathbf{I})}{\lambda_1 \lambda_2 \lambda_3} \quad (60a)$$

$$\mathbf{Q}_1 = \frac{\mathbf{M}(\mathbf{M} - \lambda_2 \mathbf{I})(\mathbf{M} - \lambda_3 \mathbf{I})}{\lambda_1(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)} \quad (60b)$$

$$\mathbf{Q}_2 = \frac{\mathbf{M}(\mathbf{M} - \lambda_1 \mathbf{I})(\mathbf{M} - \lambda_3 \mathbf{I})}{\lambda_2(\lambda_2 - \lambda_1)(\lambda_2 - \lambda_3)} \quad (60c)$$

$$\mathbf{Q}_3 = \frac{\mathbf{M}(\mathbf{M} - \lambda_1 \mathbf{I})(\mathbf{M} - \lambda_2 \mathbf{I})}{\lambda_3(\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2)} \quad (60d)$$

The solution of the system of the differential eqs 55, expressed in terms of matrix projection operators, is then

$$\mathbf{C} = \mathbf{Q}_0 \mathbf{C}_0 + \mathbf{Q}_1 \mathbf{C}_0 \exp(\lambda_1 t) + \mathbf{Q}_2 \mathbf{C}_0 \exp(\lambda_2 t) + \mathbf{Q}_3 \mathbf{C}_0 \exp(\lambda_3 t) \quad (61)$$

where \mathbf{Q}_i are defined by eqs 60a–d.

Therefore, using projection operators the numerical analysis of the complex multistep reaction is reduced to finding only the eigenvalues of the original coefficient matrix \mathbf{M} .

Inspection of eqs 60a–d and their comparison with eq 61 show that projection operators, \mathbf{Q}_i , are matrices of the same size as the size of the original coefficient matrix \mathbf{M} . Also, the products $\mathbf{Q}_i\mathbf{C}_0$ are column vectors, \mathbf{P}_i , which are the projections of \mathbf{C}_0 on each eigenvector of the matrix \mathbf{M} . Notice that \mathbf{P}_i are obtained without determining the eigenvectors of \mathbf{M} . Thus,

$$\mathbf{C} = \mathbf{P}_0 + \mathbf{P}_1 \exp(\lambda_1 t) + \mathbf{P}_2 \exp(\lambda_2 t) + \mathbf{P}_3 \exp(\lambda_3 t) \quad (62a)$$

and

$$\begin{pmatrix} [C]_1 \\ [C]_2 \\ [C]_3 \\ [C]_4 \end{pmatrix} = \begin{pmatrix} P_{01} \\ P_{02} \\ P_{03} \\ P_{04} \end{pmatrix} + \begin{pmatrix} P_{11} \\ P_{12} \\ P_{13} \\ P_{14} \end{pmatrix} \exp(\lambda_1 t) + \begin{pmatrix} P_{21} \\ P_{22} \\ P_{23} \\ P_{24} \end{pmatrix} \exp(\lambda_2 t) + \begin{pmatrix} P_{31} \\ P_{32} \\ P_{33} \\ P_{34} \end{pmatrix} \exp(\lambda_3 t) \quad (62b)$$

where P_{ij} is the j th element of the projection of the vector of the initial concentrations \mathbf{C}_0 on the eigenvector corresponding to the i th eigenvalue of matrix \mathbf{M} . In stopped-flow experiments, the concentrations of all protein species change from the concentrations at $t = 0$ to the equilibrium concentrations at $t = \infty$, defined by the elements of vector \mathbf{P}_0 . It should be pointed out that each element of P_{ij} in eq 62b is an algebraic expression in term of eigenvalues, λ_i , rate constants of the system, and total ligand and macromolecule concentrations, defined by the products $\mathbf{Q}_i\mathbf{C}_0$.

There are three normal modes of the reaction and three amplitudes, A_1, A_2 , and A_3 , corresponding to relaxation times $\tau_1 = -1/\lambda_1$, $\tau_2 = -1/\lambda_2$, and $\tau_3 = -1/\lambda_3$. In spectroscopic stopped-flow experiments, concentrations of the reactants and products are indirectly monitored through some spectroscopic parameter (e.g., fluorescence) characterizing interacting species. In general, each intermediate will have different fluorescence properties. Thus, there are four molar fluorescence intensities, F_1, F_2, F_3 , and F_4 , characterizing N_F, C_2, C_3 , and C_4 states of the nucleic acid, free and in the complex with the protein. The concentrations of all nucleic acid species, at any time of the reaction, follow the mass conservation relationship

$$N_T = N_F + C_2 + C_3 + C_4 \quad (63)$$

where C_2, C_3 , and C_4 are defined by eq 62b.

The fluorescence of the system at time t of the reaction, $F(t)$, is defined by

$$F(t) = F_1 N_F + F_2 C_2 + F_3 C_3 + F_4 C_4 \quad (64)$$

Introducing eqs 62b and 63 into eq 64, one obtains

$$F(t) = F_1 N_T + \begin{pmatrix} F_2 P_{02} + F_3 P_{03} + F_4 P_{04} \\ F_2 P_{12} + F_3 P_{13} + F_4 P_{14} \\ F_2 P_{22} + F_3 P_{23} + F_4 P_{24} \\ F_2 P_{32} + F_3 P_{33} + F_4 P_{34} \end{pmatrix}^T \begin{pmatrix} 1 \\ \exp(\lambda_1 t) \\ \exp(\lambda_2 t) \\ \exp(\lambda_3 t) \end{pmatrix} \quad (65)$$

where index T indicates the transpose matrix. The observed total amplitude, A_T , of the stopped-flow trace is the sum of individual amplitudes of all normal modes

$$A_T = A_1 + A_2 + A_3 \quad (66a)$$

Experimentally, the total amplitude, A_T is described by

$$A_T = F(0) - F(\infty) \quad (66b)$$

where $F(0)$ and $F(\infty)$ are the observed fluorescence intensities, $F(t)$, of the system at $t = 0$ and $t = \infty$, respectively. Introducing the mass conservation relationship defined by eq 63 into eq 65, $t = 0$ for $F(0)$ and $t = \infty$ for $F(\infty)$, one obtains

$$F(0) = F_1 N_T + (P_{02} + P_{12} + P_{22} + P_{32} P_{03} + P_{13} + P_{23} + P_{33} P_{04} + P_{14} + P_{24} + P_{34}) \begin{pmatrix} F_2 - F_1 \\ F_3 - F_1 \\ F_4 - F_1 \end{pmatrix} \quad (67)$$

and

$$F(\infty) = F_1 N_T + (P_{02} P_{03} P_{04}) \begin{pmatrix} F_2 - F_1 \\ F_3 - F_1 \\ F_4 - F_1 \end{pmatrix} \quad (68)$$

which define the total amplitude (eq 66b) as

$$A_T = (P_{12} + P_{13} + P_{14} P_{22} + P_{23} + P_{24} P_{32} + P_{33} + P_{34}) \begin{pmatrix} F_2 - F_1 \\ F_3 - F_1 \\ F_4 - F_1 \end{pmatrix} \quad (69)$$

The individual amplitudes A_1, A_2 , and A_3 for each normal mode are then

$$A_1 = (P_{12} P_{13} P_{14}) \begin{pmatrix} F_2 - F_1 \\ F_3 - F_1 \\ F_4 - F_1 \end{pmatrix} \quad (70)$$

$$A_2 = (P_{22} P_{23} P_{24}) \begin{pmatrix} F_2 - F_1 \\ F_3 - F_1 \\ F_4 - F_1 \end{pmatrix} \quad (71)$$

$$A_3 = (P_{32} P_{33} P_{34}) \begin{pmatrix} F_2 - F_1 \\ F_3 - F_1 \\ F_4 - F_1 \end{pmatrix} \quad (72)$$

Expressions 69–72 are closed-form, explicit relationships for the total and individual amplitudes for the three-step reaction mechanism described by eq 52. Thus, once the matrix operators are formulated in terms of the original matrix of coefficients \mathbf{M} , the total and individual amplitudes of the reaction system can be easily defined. Extension of the analysis to more complex reaction systems is straightforward.^{116,118}

The relationships derived above also provide an important intuitive insight into the effect of different values of spectroscopic properties of the intermediates of the reaction on observed amplitudes. For instance, even if all intermediates have the same fluorescence properties ($F_2 = F_3 = F_4$), the amplitudes of all normal modes will be observed, although there are no additional fluorescence changes in all transitions following C_2 . Expressions 69–72 show that the progressive fluorescence changes in subsequent transitions impact less on the individual amplitudes than the difference between the fluorescence of the free nucleic acid (ligand) and a given intermediate. For some kinetic systems, it may not be possible to detect all present normal modes of the reaction. However, this would result from combined effect of rate constants, relaxation times, and spectroscopic changes and not from a similar or identical spectroscopic signal change accompanying the formation of subsequent intermediates, as sometime assumed. The obtained expressions for the individual amplitudes of the kinetic steps makes it possible to extract the spectroscopic properties of each intermediate of the reaction, thus providing information about the structure of the intermediate unavailable by other methods. Notice that this can be facilitated by setting the fluorescence of the free ligand, $F_1 = 1$. Then, all remaining molar fluorescence intensities, F_2 , F_3 , and F_4 , are uniquely determined relative to F_1 by eqs 69–72. Because the quantum yield for the free ligand can be independently obtained, if required, the true quantum yields for all other intermediates can also be determined.

Examination of the relaxation times of the studied kinetics as a function of the ligand concentration constitutes the first, although not exclusive (see above), fundamental step in establishing the mechanism of the complex reaction and determining the rate constants of particular elementary processes.¹¹⁴ The reciprocal relaxation times for the three-step sequential reaction, described by eq 52, as a function of the free ligand concentration are shown in Figure 24. Relaxation times have been obtained by direct numerical determination of the eigenvalues, λ_1 , λ_2 , and λ_3 , of the matrix \mathbf{M} at a given free ligand concentration, $[C_1]$, using the identities of $1/\tau_1 = -\lambda_1$, $1/\tau_2 = -\lambda_2$, and $1/\tau_3 = -\lambda_3$. The selected rate constants are $k_1 = 1 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$, $k_{-1} = 0.05 \text{ s}^{-1}$, $k_2 = 0.1 \text{ s}^{-1}$, $k_{-2} = 0.05 \text{ s}^{-1}$, $k_3 = 0.01 \text{ s}^{-1}$, and $k_{-3} = 0.005 \text{ s}^{-1}$. Because of the large differences between the values of the selected rate constants among the elementary steps, the relaxation times differ significantly at any concentration of the nucleic acid, that is, the normal modes of the reaction are close to the “uncoupled” ones.¹¹⁴ In such a situation, it could be possible to obtain approximate formulas for each of the relaxation times; however, the numerical approach applied here avoids such approximations. The largest relaxation time has typical characteristics of the bimolecular binding process with $1/\tau_1$ increasing linearly with the ligand concentration in the high ligand concentration range.^{114,116} On the other hand, at low ligand concentrations, there is clearly a nonlinear region. This region, observed experimentally (see below), is only evident because no approximate expression is used for this relaxation time. Both $1/\tau_2$ and $1/\tau_3$ for the considered sequential three-step reaction show hyperbolic dependence upon nucleic acid concentration reaching the plateau values at high nucleic acid concentration. As a result, τ_2 and τ_3 become independent of the ligand concentration in the high concentration range. There are also two features, often unnoticed in experimental practice, of

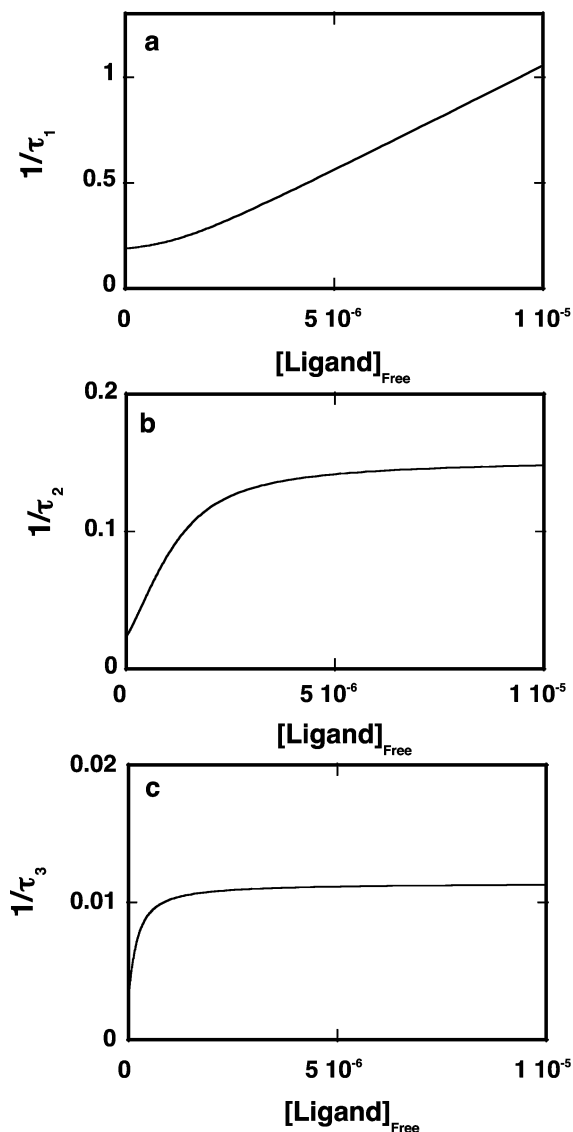


Figure 24. Computer simulation of the dependence of reciprocal relaxation times for the three-step sequential mechanism of the ligand binding to a macromolecule upon free ligand concentration: (a) $1/\tau_1$; (b) $1/\tau_2$; (c) $1/\tau_3$. Relaxation times have been obtained by numerically determining the eigenvalues of the coefficient matrix, \mathbf{M} ($\lambda_1, \lambda_2, \lambda_3$), and using identities $1/\tau_1 = -\lambda_1$, $1/\tau_2 = -\lambda_2$, and $1/\tau_3 = -\lambda_3$. The simulations have been performed using rate constants $k_1 = 1 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$, $k_{-1} = 0.05 \text{ s}^{-1}$, $k_2 = 0.1 \text{ s}^{-1}$, $k_{-2} = 0.05 \text{ s}^{-1}$, $k_3 = 0.01 \text{ s}^{-1}$ and $k_{-3} = 0.005 \text{ s}^{-1}$. The selected total ligand, N_T , and macromolecule, C_T , concentrations are $1 \times 10^{-5} \text{ M}$ and $1 \times 10^{-8} \text{ M}$, respectively. Reprinted with permission from ref 116. Copyright 2000 Elsevier.

the relaxation time plots clearly seen in the computer simulations in Figure 24. First, for the selected values of the rate constants, $1/\tau_3$ reaches the plateau at significantly lower ligand concentrations than $1/\tau_2$. Such behavior results from the fact that for the selected values of the rate constants, each partial step of the reaction contributes favorably to the total free energy of binding, ΔG° . In other words, this feature can serve as a diagnostic of favorable free energy changes in the subsequent steps. Second, the hyperbolic dependence of $1/\tau_3$ upon ligand concentration is often unnoticed when the pseudo-first-order conditions are applied, placing the experimental data already in the plateau.

Analysis of the amplitudes of the spectroscopic relaxation processes provides an independent test of the kinetic mechanisms selected on the basis of the behavior of the relaxation

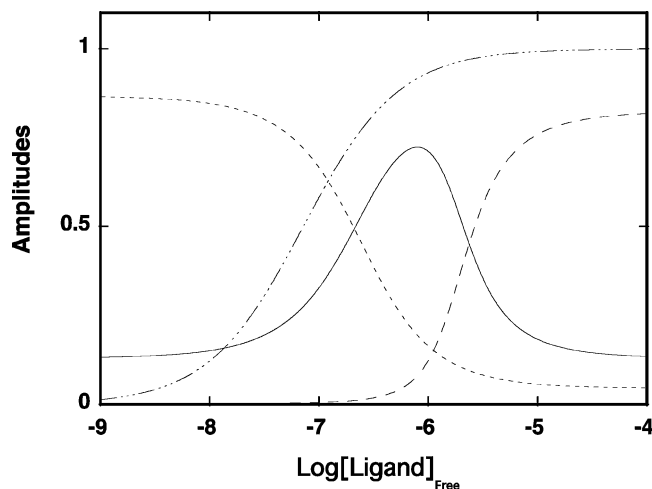


Figure 25. Computer simulation of the dependence of individual, A_1 , A_2 , and A_3 , and total, A_T , relaxation amplitudes for the three-step sequential mechanism of a ligand binding to a macromolecule upon the logarithm of the free ligand concentration: A_1 (---); A_2 (—); A_3 (- - -); A_T (- ···). The relative fluorescence intensities, F_2 , F_3 , and F_4 , characterizing corresponding intermediates, C_2 , C_3 , and C_4 , are 3, 3.5, and 3.5, respectively. The fluorescence of the free ligand, F_N , is taken as 1. The individual amplitudes are expressed as fractions of the total amplitude A_T , while the total amplitude has been normalized to 1 at saturating ligand concentrations. The simulations have been performed using closed-form expressions defined by eqs 69–72 with the rate constants $k_1 = 1 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$, $k_{-1} = 0.05 \text{ s}^{-1}$, $k_2 = 0.1 \text{ s}^{-1}$, $k_{-2} = 0.05 \text{ s}^{-1}$, $k_3 = 0.01 \text{ s}^{-1}$, and $k_{-3} = 0.005 \text{ s}^{-1}$. Reprinted with permission from ref 116. Copyright 2000 Academic Press.

times. Moreover, it offers a unique opportunity to obtain information about the structure of the reaction intermediates and the physical nature of the elementary steps. The dependence of individual amplitudes, A_1 , A_2 , and A_3 , expressed as fractions ($A_i/\sum A_i$) of the total amplitude ($A_T = \sum A_i$) upon the ligand concentration is shown in Figure 25. The total amplitude, A_T , normalized to 1 at high ligand concentrations is also included. The computer simulations have been performed using eqs 69–72. The selected molar fluorescence intensities are $F_1 = 1$, $F_2 = 3$, $F_3 = 3.5$, and $F_4 = 3.5$; the rate constants are the same as those in Figure 24. It is clear that for the selected values of the rate constants, at low ligand concentrations only the amplitudes of the second, A_2 , and third, A_3 , normal modes of the reaction contribute significantly to the observed A_T , although the major fluorescence change, as compared to the fluorescence of the free nucleic acid, accompanies the formation of C_2 . Such behavior is the result of the low efficiency of the C_2 complex formation at low nucleic acid concentration, while the formed complex still relaxes with the second and third normal mode. At high ligand concentrations, the amplitude of the first normal mode, A_1 , dominates the relaxation process. The computer simulations in Figure 25 show that all three amplitudes of the present relaxation modes are detectable. This is despite the fact that there is no additional fluorescence change in the transition from C_3 to C_4 intermediate. As mentioned above, this is evident from eqs 69–72, which show that individual amplitude for a given normal mode of the reaction is mainly affected by the difference between the spectroscopic properties of intermediates and the free nucleic acid (see above). Computer simulations shown in Figure 25 were performed with given values of the relative fluorescence intensities for all intermediates. In experimental studies of a kinetic system, this process is

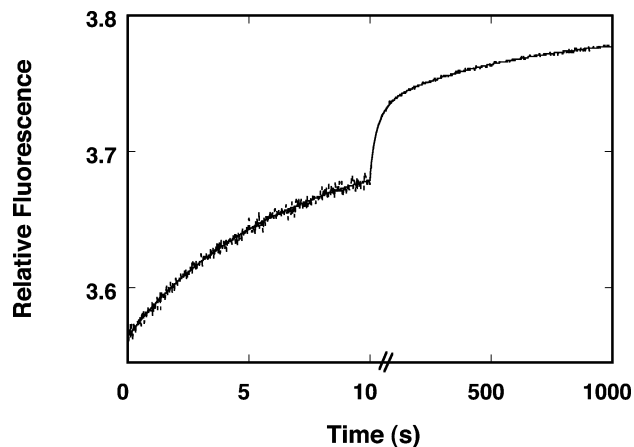


Figure 26. The fluorescence stopped-flow kinetic trace ($\lambda_{\text{ex}} = 320 \text{ nm}$, $\lambda_{\text{em}} > 400 \text{ nm}$), recorded in two time bases, 10 and 1000 s, after mixing the DnaB helicase with the 20-mer $d\epsilon A(p\epsilon A)_{19}$ in 50 mM Tris/HCl (pH 8.1, 10 °C), containing 100 mM NaCl, 5 mM MgCl_2 , and 1 mM AMP-PNP. The final concentrations of the helicase and the 20-mer are $1.5 \times 10^{-7} \text{ M}$ (hexamer) and $5 \times 10^{-6} \text{ M}$ (oligomer), respectively. The solid line is the three-exponential, nonlinear least-squares fit of the experimental curve. Reprinted with permission from ref 116. Copyright 2000 Elsevier.

reversed, that is, from the dependence of the amplitudes of the system upon ligand (or macromolecule) concentrations, one can determine the spectroscopic parameters characterizing all intermediates, as we discuss below for the case of the DnaB helicase association with the ssDNA oligomers.

3.5.2. Protein–Nucleic Acid System with Slow Bimolecular Step. Kinetics of the ssDNA 20-mer Binding to the DnaB Helicase.

3.5.2.1. Relaxation Times. The site size of the DnaB hexamer in the complex with the ssDNA is 20 ± 3 nucleotides per hexamer.^{35,36,101} The DnaB hexamer binds a single 20-mer molecule, and the oligomer encompasses the entire total binding site of the helicase. As discussed above, binding of the fluorescent etheno derivative of $dA(pA)_{19}$, $d\epsilon A(p\epsilon A)_{19}$, is accompanied by a strong, ~ 3 -fold, increase of the nucleic acid fluorescence, providing an excellent signal to monitor the kinetics of the helicase–ssDNA complex formation (Figure 1a).¹¹⁶ The stopped-flow experiments have been performed under pseudo-first-order conditions by mixing the DnaB helicase with a large excess of the ssDNA 20-mer.

The stopped-flow kinetic trace of the $d\epsilon A(p\epsilon A)_{19}$ fluorescence after mixing $5 \times 10^{-6} \text{ M}$ oligomer with $1.5 \times 10^{-7} \text{ M}$ (hexamer) DnaB helicase (final concentrations) is shown in Figure 26. The curve is shown in two time bases, 10 and 1000 s. The observed kinetics is complex, clearly showing the presence of multiple steps. The solid line in Figure 26 is a nonlinear least-squares fit of the spectroscopic kinetic curve using a three-exponential function. The two-exponential function provides a much less than adequate description of the experimentally observed kinetics (data not shown). Thus, the three-exponential fit is necessary to represent the observed experimental curve. A higher number of exponents do not significantly improve the statistics of the fit. Therefore, the association of the ssDNA 20-mer with the total binding site of the DnaB helicase is a complex process that includes at least three steps. Notice that the observed kinetics of the ssDNA 20-mer binding to the DnaB helicase is relatively

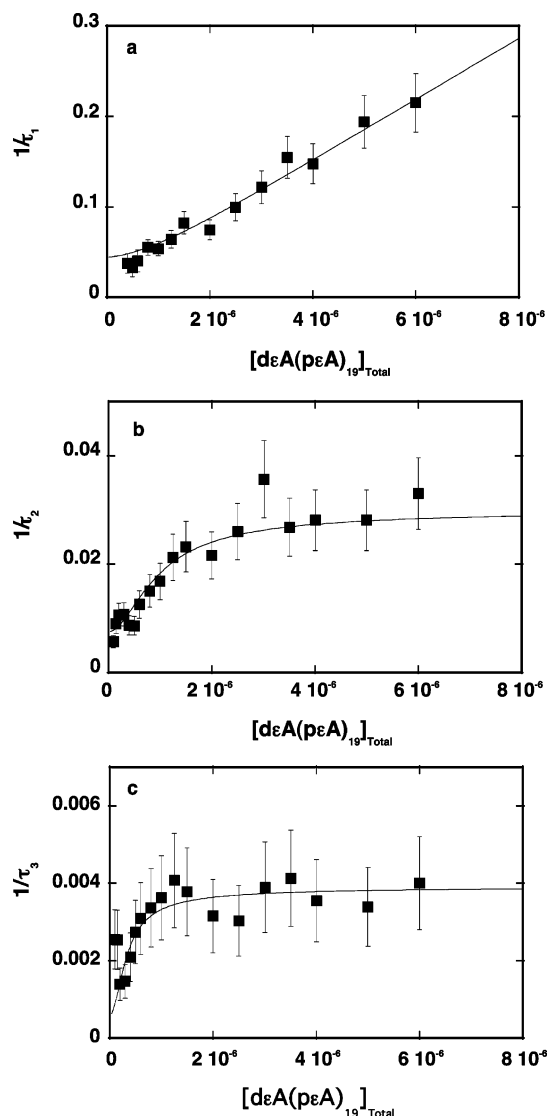
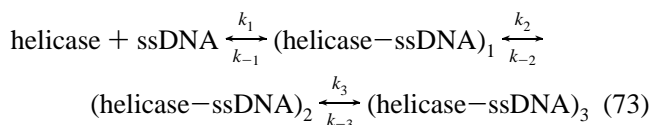


Figure 27. The dependence of the reciprocal of the relaxation times for the binding of the 20-mer $d\epsilon A(p\epsilon A)_{19}$ to the DnaB helicase in 50 mM Tris/HCl (pH 8.1, 10 °C) containing 100 mM NaCl, 5 mM MgCl₂, and 1 mM AMP-PNP upon the total concentration of $d\epsilon A(p\epsilon A)_{19}$: (a) $1/\tau_1$; (b) $1/\tau_2$; (c) $1/\tau_3$. The solid lines are nonlinear least-squares fits according to the three-step sequential mechanism with the rate constants $k_1 = 3.4 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$, $k_{-1} = 0.018 \text{ s}^{-1}$, $k_2 = 0.021 \text{ s}^{-1}$, $k_{-2} = 0.01 \text{ s}^{-1}$, $k_3 = 0.004 \text{ s}^{-1}$, and $k_{-3} = 0.0012 \text{ s}^{-1}$ (details in text). Reprinted with permission from ref 116. Copyright 2000 Elsevier.

slow. The entire kinetic process takes ~ 17 min to reach equilibrium.

The reciprocal relaxation times, $1/\tau_1$, $1/\tau_2$, and $1/\tau_3$, characterizing the three relaxation steps as a function of the total $d\epsilon A(p\epsilon A)_{19}$ concentration are shown in Figure 27. The largest reciprocal time, $1/\tau_1$, increases with $[d\epsilon A(p\epsilon A)_{19}]$, and the dependence becomes linear at high 20-mer concentrations, although there is a nonlinear phase at the low $[d\epsilon A(p\epsilon A)_{19}]$ concentration. Such behavior is typical for the relaxation time characterizing the bimolecular binding step (see Figure 24a). On the other hand, both $1/\tau_2$ and $1/\tau_3$ show hyperbolic dependence upon $[d\epsilon A(p\epsilon A)_{19}]$ and reach plateaus at high ssDNA concentrations. The minimum mechanism that can account for the observed dependence of the relaxation times upon the $d\epsilon A(p\epsilon A)_{19}$ concentration is a three-step, sequential binding process, in which bimolecular association is followed by two isomerization

steps as described by



The solid lines in Figure 27a,b,c are nonlinear least-squares fits of the relaxation times according to mechanism 73. First, the analysis was performed by the numerical nonlinear least-squares fitting of the individual relaxation times; then, the values of the rate constants were refined by global fitting, which simultaneously includes all relaxation times. The obtained rate constants for the considered mechanism are¹¹⁶ $k_1 = (3.4 \pm 0.6) \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$, $k_{-1} = 0.018 \pm 0.005 \text{ s}^{-1}$, $k_2 = 0.021 \pm 0.005 \text{ s}^{-1}$, $k_{-2} = 0.01 \pm 0.003 \text{ s}^{-1}$, $k_3 = 0.004 \pm 0.001 \text{ s}^{-1}$, and $k_{-3} = 0.0012 \pm 0.0005 \text{ s}^{-1}$. The partial equilibrium constants for each step in the mechanism are $K_1 = k_1/k_{-1}$, $K_2 = k_2/k_{-2}$, and $K_3 = k_3/k_{-3}$. Introducing the values of the rate constants provides $K_1 = (1.9 \pm 0.6) \times 10^6 \text{ M}^{-1}$; $K_2 = 2.1 \pm 1$; and $K_3 = 3.3 \pm 1$. Thus, the first step has a predominant contribution to the free energy of ssDNA binding, although the next two steps also increase the affinity. The overall binding constant, K_{20} , is related to the partial equilibrium steps by

$$K_{20} = K_1(1 + K_2 + K_2K_3) \quad (74)$$

The value of $K_{20} = (3 \pm 1) \times 10^7 \text{ M}^{-1}$ has previously been independently obtained in the same solution conditions by the equilibrium fluorescence titration method.¹¹⁶ Introducing the values of equilibrium constants for partial equilibrium steps into eq 74 gives $K_{20} = (1.9 \pm 0.7) \times 10^7 \text{ M}^{-1}$. Within experimental accuracy, this value of the overall binding constant is in excellent agreement with the K_{20} determined by equilibrium titrations. However, often, the data are not as extensive as those in Figure 27, although the behavior of the relaxation times and amplitudes provide a clear indication about the mechanism. In such situations, one can use the independently determined overall equilibrium constant to eliminate one of the kinetic parameters^{94–97,123} (see below).

3.5.2.2. Individual Amplitudes of Relaxation Steps in the ssDNA Binding to the DnaB Hexamer. The dependence of the individual amplitudes, A_1 , A_2 , and A_3 , of each of the three relaxation steps upon the ssDNA 20-mer $d\epsilon A(p\epsilon A)_{19}$ concentration is shown in Figure 28. The individual amplitudes are expressed as fractions of the total amplitude, A_T . At a low DNA concentration, only the amplitudes A_2 and A_3 of the second and third relaxation steps have a detectable contribution to the A_T . The amplitude A_2 goes through a maximum, while A_3 steadily decreases with the $d\epsilon A(p\epsilon A)_{19}$ concentration. As the concentration of the 20-mer increases, the amplitude of the first step, A_1 (bimolecular step), increases and becomes a dominant relaxation effect in the observed kinetic process.

First, such behavior of the individual amplitudes is in full agreement with the proposed mechanism (eq 73) deduced from the relaxation time analysis (Figure 27). Second, the amplitude analysis also allows us to determine the relative molar fluorescence intensities characterizing each intermediate of the reaction, that is, to assess the conformational state of the protein–nucleic acid complex in each intermediate.^{94–97,116,123} In such analyses, one utilizes the fact that the maximum fractional increase of the nucleic acid fluorescence is known, $\Delta F_{\text{max}} = 3.1$, from the equilibrium

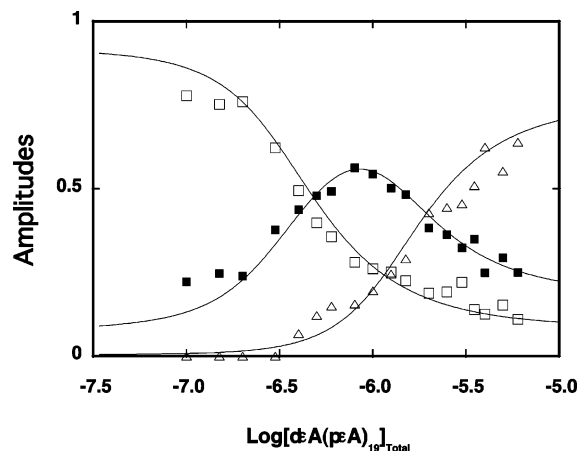


Figure 28. The dependence of the individual relaxation amplitudes for the binding of the 20-mer $d\epsilon A(p\epsilon A)_{19}$ to the DnaB helicase upon the logarithm of the total concentration of $d\epsilon A(p\epsilon A)_{19}$: A_1 (Δ); A_2 (\blacksquare); A_3 (\square). The solid lines are nonlinear least-squares fits according to the three-step sequential mechanism with the relative fluorescence intensities $F_1 = 1$, $F_2 = 3.3$, $F_3 = 4.1$, and $F_4 = 4.1$. The maximum fluorescence increase of the nucleic acid is taken from the equilibrium fluorescence titration in the same solution conditions as $\Delta F_{\max} = 3.1$. Reprinted with permission from ref 116. Copyright 2000 Elsevier.

titrations.¹¹⁶ Moreover, ΔF_{\max} can be analytically expressed as

$$\Delta F_{\max} = \frac{\Delta F_2}{1 + K_2 + K_2 K_3} + \frac{K_2 \Delta F_3}{1 + K_2 + K_2 K_3} + \frac{K_2 K_3 \Delta F_4}{1 + K_2 + K_2 K_3} \quad (75)$$

where $\Delta F_2 = (F_2 - F_1)/F_1$, $\Delta F_3 = (F_3 - F_1)/F_1$, and $\Delta F_4 = (F_4 - F_1)/F_1$ are fractional fluorescence intensities of each intermediate in the formation of the complex relative to the molar fluorescence intensity of the free DNA oligomer, F_1 . Contrary to the ΔF_i 's, the fluorescence parameters, F_2 , F_3 , and F_4 , are relative molar fluorescence intensities, but not fractional intensities, with respect to the free nucleic acid fluorescence. Expression 75 provides an additional relationship among the fitted spectroscopic parameters, thus decreasing the number of independent variables with the value of ΔF_{\max} playing a role of a scaling factor. The solid lines in Figure 28 are nonlinear least-squares fits of the experimentally determined fractional individual amplitudes of the reaction using eqs 70–72. The applied fitting procedure was similar to the one used for the relaxation times described above. Nonlinear least-squares fitting was performed with the individual amplitudes using the same rate constants as those obtained from the examination of the relaxation times or allowing the rate constants to float between $\pm 15\%$ of the values determined in the relaxation time analysis. Both approaches provide similar values of the relative fluorescence intensities. Finally, global fitting with the simultaneous analysis of all three individual amplitudes refined the obtained parameters. The fluorescence of the free $d\epsilon A(p\epsilon A)_{19}$ was taken as $F_1 = 1$. The results indicate that the largest fluorescence change as compared to the free 20-mer occurs in the first binding step, that is, in the formation of the (helicase–ssDNA)₁ (see above). Upon formation of this complex, the fluorescence of the 20-mer increases by factor 3.3 ($F_2 = 3.3 \pm 0.4$) as compared to the free DNA. Conformational transition to (helicase–ssDNA)₂ induces

only an $\sim 30\%$ additional increase of the 20-mer fluorescence over F_1 ($F_3 = 4.1 \pm 0.4$), while the transition to the (helicase–ssDNA)₃ is not accompanied by an additional fluorescence increase over F_3 ($F_4 = 4.1 \pm 0.4$).¹¹⁶

3.5.2.3. Effect of the Protein Concentration on the Measured Relaxation Times of the ssDNA 20-mer–DnaB Protein Association. The fact that the DnaB protein is a hexamer introduces another aspect of the interactions that is not present when the interacting protein is a monomer. Thus, the stability of the hexamer is an important factor in the examination of the kinetics of the helicase–ssDNA complex formation. Any dissociation of the hexamer into lower oligomers would obscure the kinetic processes and the possibility of quantitatively interpreting it, unless the kinetics of such dissociation and its effect on the dynamics of the ssDNA binding is also examined. On the other hand, dissociation of the hexamer could be a part of the binding mechanism. Figure 29a,b,c shows the determined relaxation times as a function of the DnaB (hexamer) concentration over an order of magnitude change of the protein concentration. The data clearly show that, within experimental accuracy, all three relaxation times are independent of the helicase concentration, indicating that the observed kinetics are not affected by the protein–protein interactions. Also, the lack of a protein concentration on the observed kinetic process indicates that the DnaB protein hexamer does not dissociate prior to binding the ssDNA, but rather the entry of the ssDNA into the cross-channel of the hexamer occurs through a local opening of the hexamer.¹¹⁶

3.5.2.4. Some Molecular Aspects of the DnaB Protein–ssDNA Interactions. It should be noticed that the value of the bimolecular association rate constant, $k_1 = (3.4 \pm 0.6) \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$, is dramatically lower than expected for the diffusion-controlled reaction.^{124,125} The intrinsic association rate constant could be even lower because the six subunits of the DnaB hexamer are chemically identical, that is, there could be six entry sites for the ssDNA into the cross channel of the hexamer. Thus, the determined association rate constant may contain a statistical factor as high as 6. In chemical bimolecular reactions in solution, formation of a collision/encounter complex, a process that is controlled by diffusion of the reactants,^{124,125} precedes formation of the product. Theoretical values of the maximum rate constant for the diffusion-controlled association can be estimated using the Smoluchowski equation¹²⁶

$$k_D = \frac{4\pi N_A (D_P + D_D)(r_P + r_D)}{1000} \quad (76)$$

where N_A is Avogadro's number, D_P and D_D are diffusion coefficients of the protein and the nucleic acid, respectively, and r_P and r_D are their interaction radii. The diffusion coefficient of the DnaB helicase hexamer, $D_P = (2.8 \pm 0.3) \times 10^{-7} \text{ cm}^2/\text{s}$, has been determined using the dynamic light scattering technique.¹¹⁶ The diffusion coefficient of the ssDNA 20-mer can be estimated from the Svedberg equation¹²⁷

$$D_D = \frac{sRT}{M_{20}(1 - \bar{v}\rho)} \quad (77)$$

where s and M_{20} are the sedimentation coefficient and molecular weight of the 20-mer, R is the gas constant, T is the temperature (kelvin), \bar{v} is the ssDNA specific volume, and ρ is the solvent density. Using the analytical ultra-

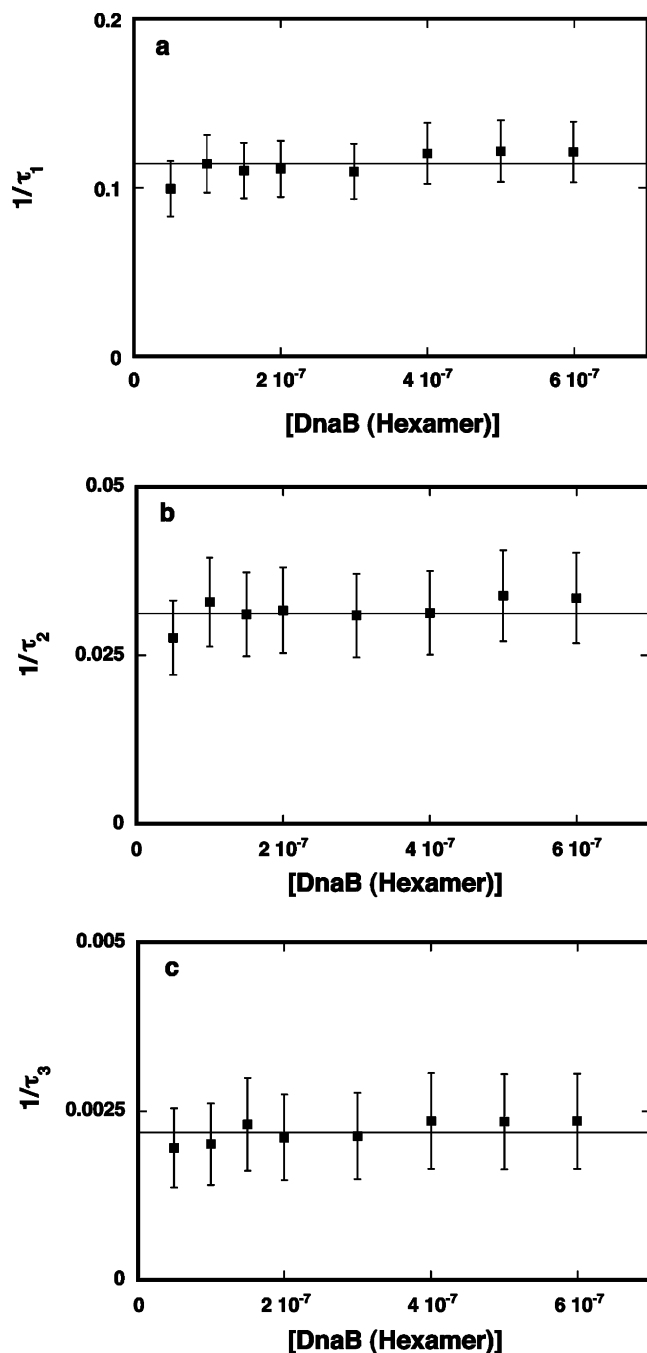
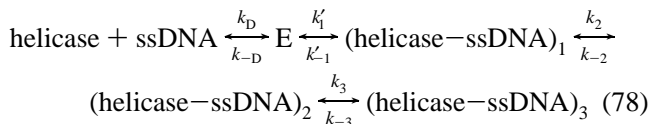


Figure 29. The dependence of the reciprocal relaxation times of the $\epsilon\text{A}(\epsilon\text{A})_{19}$ –DnaB helicase system upon the enzyme concentration (hexamer) in 50 mM Tris/HCl (pH 8.1, 10 °C) containing 100 mM NaCl, 5 mM MgCl_2 , and 1 mM AMP-PNP: (a) $1/\tau_1$; (b) $1/\tau_2$; (c) $1/\tau_3$. Reprinted with permission from ref 116. Copyright 2000 Elsevier.

centrifugation technique, we determined the sedimentation coefficient of the 20-mer, $s_{20,w} = 1.4 \pm 0.12$ S. Using this value of $s_{20,w}$ and $M_{20} \approx 6400$ g/mol, $\bar{v} = 0.505$ mL/g, and $r = 1$ g/mL, one obtains¹¹⁶ $D_D = 1.1 \times 10^{-6}$ cm²/s. Correcting this value to our buffer conditions, one obtains $D_D = 6.1 \times 10^{-7}$ cm²/s. The interacting radii have been taken as equal to the approximate size of the 20-mer $r_D = r_p \approx 68$ Å. With these values, the diffusion-controlled association rate constant is $k_D \approx 8.9 \times 10^9$ M⁻¹ s⁻¹. Smaller interaction radii or orientation factors could lower this value to $k_D \approx 10^7$ – 10^8 M⁻¹ s⁻¹. Despite the approximate nature of these estimates, the determined bimolecular rate constant k_1 is

~ 3 – 6 orders of magnitude lower than the k_D , predicted by the diffusion-controlled collision, indicating that the bimolecular association step contains an additional conformational transition of the helicase–ssDNA complex. This conclusion is also evident from the amplitude analysis, which shows that the largest increase of the nucleic acid fluorescence in the complex with the DnaB helicase (most probably largest conformational change) occurs in the formation of the (helicase–ssDNA)₁ (see below). Such dramatic conformational changes cannot take place in a collision complex because then it would not be a collision complex.^{124,125} Therefore, the reaction mechanism should be enlarged by an extra step following the collision complex, E, as described by



where k_D and k_{-D} are rate constants for the formation and dissociation of the collision complex and k'_1 and k_{-1} are the rate constants for the transition from the collision complex to the (helicase–ssDNA)₁. The equilibrium constant for the first step is then $K_D = k_D/k_{-D}$. Because the formation of E is a very fast process and E equilibrates before any significant transition to the (helicase–ssDNA)₁ takes place, the observed apparent bimolecular rate constant is $k_1 = K_D k'_1$. Analysis of the kinetic traces indicates that there is no amplitude lost in the dead time of the instrument (~ 1.4 ms).¹¹⁶ The lack of any amplitude corresponding to the formation of the collision complex results from the fact that the process is very fast and there is no conformational transition accompanying its formation.

The estimate of the range of the values of k'_1 can be obtained as follows. Notice that the dependence of the reciprocal relaxation time for the bimolecular process, $1/\tau_1$, is a linear function of the nucleic acid concentration at the highest concentrations examined ($\sim 6 \times 10^{-6}$ M). This shows that K_D is much lower¹¹⁶ than $\sim 2 \times 10^5$ M⁻¹. Taking a conservative value $K_D \approx 10^4$ M⁻¹ gives $k'_1 \approx 3.4$ s⁻¹. Thus, the data indicate that the collision complex E undergoes a transition to the (helicase–ssDNA)₁, which has a forward rate constant of ~ 2 – 3 orders of magnitude larger than the forward rate constants for the subsequent formation of the (helicase–ssDNA)₂ and (helicase–ssDNA)₃.

The amplitude analysis indicates that the largest fluorescence increase accompanies the formation of the (helicase–ssDNA)₁ complex. Notice that at the excitation wavelength applied ($\lambda_{\text{ex}} = 320$ nm), predominantly, the etheno-adenosine is excited. Thus, the observed fluorescence increase results from a nucleic acid quantum yield increase in the complex with the helicase, not through the energy transfer processes. Moreover, tryptophans of the DnaB protein are located far enough away from the ssDNA-binding site to eliminate any efficient energy transfer.¹²⁸ The fluorescence of ϵA is dramatically quenched (8–10-fold) in the etheno oligomers as compared to the free ϵAMP .^{102,103,129,130} Stacking interactions among neighboring ϵA bases are similar to stacking interactions in unmodified adenosine polymers.¹³⁰ The quenching of the ϵA fluorescence has been modeled as a dynamic phenomenon in which the motion of the ϵA leads to quenching via intramolecular collision. Fluorescence of the etheno derivative depends little on solvent conditions.^{102,103,131} In other words, changes of the fluorescence

of the etheno derivative ssDNA oligomers are induced predominantly through conformational changes of the nucleic acid. This is a particularly useful property of the etheno derivatives in any studies of the protein–nucleic acid complexes. The observed strong fluorescence increase of the etheno derivative ssDNA oligomers upon binding to the DnaB helicase indicates significantly restricted mobility and separation of the nucleic acid bases in the complex with the enzyme.¹¹⁶ The largest fluorescence increase observed in the formation of the (helicase–ssDNA)₁ strongly suggests that these dramatic changes of the nucleic acid conformation occur in the formation of this complex and are preserved in the (helicase–ssDNA)₂ and (helicase–ssDNA)₃.

As we pointed out, a large conformational change accompanying the formation of the (helicase–ssDNA)₁ argues against the possibility that this complex is a result of a simple collision or an encounter. Conformational transitions of the single-stranded nucleic acid oligomers, particularly base stacking, are very fast processes, which occur in the range of microseconds.^{132,133} The fact that the transition from the collision complex E to the (helicase–ssDNA)₁ is characterized by a rate constant in the range of $\sim 3 \text{ s}^{-1}$ indicates that this is not exclusively the change of the nucleic acid structure upon association, but rather a conformational transition of the enzyme–ssDNA complex. In other words, the observed dynamics of the formation of the (helicase–ssDNA)₁ is an intrinsic property of the DnaB helicase in response to the nucleic acid binding and not simply an adjustment of the bound ssDNA to the structure of the enzyme binding site.¹¹⁶

3.5.3. Protein–Nucleic Acid System with Fast Bimolecular Step with Undetectable Amplitude. Kinetics of PriA Helicase–ssDNA Interactions

3.5.3.1. Kinetics of the PriA Helicase Binding to the ssDNA 20-mer.

Analyses of the DnaB protein binding to the ssDNA indicate that the binding is characterized by a slow bimolecular step indicating the presence of some additional first-order step. Such complex behavior reflects, in part, the complex structure of the enzyme built as a ringlike hexamer of six identical subunits with the ssDNA-binding site located inside the cross-channel of the hexamer⁴⁴ (see above). A different character of the bimolecular step is observed in the case of the monomeric *E. coli* PriA helicase interactions with ssDNA.¹²³ Recall that thermodynamic studies indicate that the total site size of the PriA–ssDNA complex, that is, the maximum number of nucleotides occluded by the PriA helicase in the complex, is 20 ± 3 residues per protein monomer, although the proper ssDNA-binding site occludes ~ 8 nucleotides.^{47,48} To address the mechanism of the PriA helicase binding to the ssDNA within the total site size of the formed complex, the kinetic stopped-flow experiments were first performed with the ssDNA 20-mer, dεA(pεA)₁₉.

The fluorescence stopped-flow kinetic trace of the ssDNA 20-mer dεA(pεA)₁₉ after mixing $3 \times 10^{-7} \text{ M}$ nucleic acid with $7.3 \times 10^{-6} \text{ M}$ PriA (final concentrations) is shown in Figure 30a. To increase the resolution, the plot is shown in logarithmic scale with respect to time. The initial horizontal part of the trace corresponds to the steady-state fluorescence intensity of the sample, recorded for $\sim 2 \text{ ms}$, before the flow stops.¹²³ The solid line in Figure 30a is a nonlinear, least-squares fit of the experimental curve using a two-exponential function. The included single-exponential function does not provide an adequate description of the observed kinetics

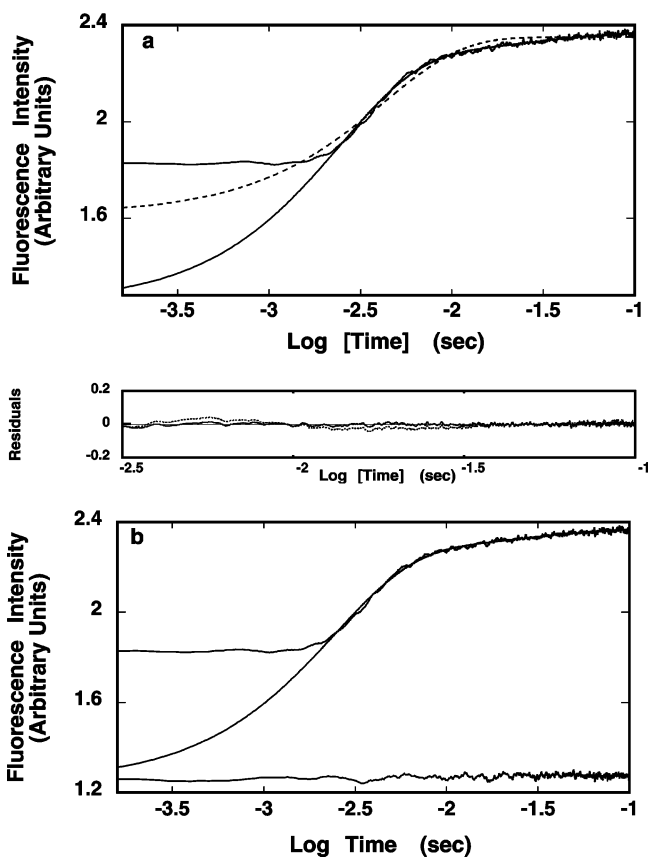


Figure 30. (a) The fluorescence stopped-flow kinetic trace after mixing PriA helicase with the ssDNA 20-mer dεA(pεA)₁₉ in 10 mM sodium cacodylate/HCl (pH 7.0, 10 °C) containing 100 mM NaCl ($\lambda_{\text{ex}} = 325 \text{ nm}$, $\lambda_{\text{em}} > 400 \text{ nm}$). The final concentrations of the helicase and the 20-mer are 7.3×10^{-6} and $3 \times 10^{-7} \text{ M}$, respectively. The solid line is the two-exponential, nonlinear least-squares fit of the experimental curve. The dashed line is the nonlinear least-squares fit using the single-exponential function. The horizontal, initial part of the trace is the steady-state value of the fluorescence of the sample recorded 2 ms before the flow stopped. (b) The same fluorescence stopped-flow trace as in panel a, together with the zero line trace (lower trace), which is obtained after mixing the nucleic acid at the same concentration as that used with the protein but only with the buffer. The solid line is the same two-exponential, nonlinear least-squares fit of the experimental curve as that shown in panel a. Reprinted with permission from ref 123. Copyright 2003 American Chemical Society.

(dashed line). Using a larger number of exponents in the fitting function (eq 1) does not improve the statistics of the fit.¹²³

The stopped-flow kinetic trace, together with the trace corresponding to the ssDNA oligomer alone, at the same concentration of the nucleic acid as used with the protein but only mixed with the background buffer (zero line), is shown in Figure 30b. It should be stressed that comparison between the determined relaxation amplitudes and the total amplitude of the kinetic trace at several enzyme concentrations is crucial in establishing that the resolved relaxation processes account for the total observed signal.^{94–97,123} The two-exponential fit provides an excellent description of the observed kinetic process, yielding the sum of amplitudes, which is the same as the observed total amplitude of the overall relaxation process. Because this behavior is observed at all studied enzyme concentrations, that is, no signal change is lost in the instrumental dead time, the simplest interpretation would be that the enzyme binding to the ssDNA is a two-step process.¹¹⁴ However, the behavior of the relaxation

times and amplitudes shows that the binding is characterized by a more complex mechanism.¹²³

The dependence of the reciprocal relaxation times, $1/\tau_1$ and $1/\tau_2$, upon the total concentration of PriA is shown in Figure 31a,b. The functional dependence of $1/\tau_1$ upon PriA concentration shows a typical nonlinear hyperbolic dependence upon PriA concentration in the examined enzyme concentration range. Thus, the shortest observed relaxation time does not describe the bimolecular reaction, where a strictly linear dependence upon the enzyme concentration is expected.^{94–97,114,116,123} The nonlinear character of the plot in Figure 31a indicates that τ_1 characterizes an intramolecular transition. The values of $1/\tau_2$ are independent of the helicase concentration, clearly indicating that this relaxation time characterizes another intramolecular transition of the protein–ssDNA complex. Therefore, the simplest mechanism that can describe the observed dependence of the relaxation times upon the PriA concentration is a sequential reaction in which the PriA helicase binds the ssDNA in a very fast bimolecular step, followed by two first-order transitions of the formed protein–ssDNA complex, as described by

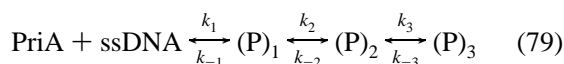


Figure 31c shows the dependence of the normalized, individual amplitudes, A_1 and A_2 , of the two observed relaxation steps upon the logarithm of PriA concentration. The amplitude of the first relaxation step, A_1 , dominates the relaxation process over the entire examined range of PriA concentration. Also, its values slightly increase with increasing concentrations of the helicase. Values of the amplitude A_2 are significantly lower than those of A_1 and slightly decrease with the increase of the enzyme concentration. As mentioned above, the amplitude of the bimolecular step is undetectable, that is, its values must be below $\sim 1\%$ of the total signal, in the examined protein concentration range. The observed behavior of the resolved individual amplitudes as functions of the PriA concentration is in excellent agreement with the proposed mechanism.¹²³

The analysis of the relaxation data in Figure 31a,b,c is initiated by numerical nonlinear least-squares fitting of the individual relaxation times. Because the rate of the bimolecular step is very fast and the amplitude of this fast normal mode is undetectable, the bimolecular step, in such case, is beyond the resolution of the stopped-flow technique (see below). On the other hand, because of the fast rate, the bimolecular step equilibrates before the transition to the next intermediate takes place, allowing us to use the overall partial equilibrium constant, $K_1 = k_1/k_{-1}$, as a fitting parameter. Notice that the overall rate constant k_1 differs from k_1 by a statistical factor resulting from the presence of potential binding sites on the DNA.¹²³ The analyses are facilitated by the fact that we also know the value of the overall macroscopic binding constant, K_N , for the enzyme binding to the ssDNA 20-mer. The value of K_{20} is $(4.9 \pm 0.6) \times 10^5 \text{ M}^{-1}$ and has been independently obtained in the same solution conditions by the equilibrium fluorescence titration method.^{47,48} The macroscopic binding constant, K_{20} , is related to the overall bimolecular partial equilibrium constant K_1 and partial equilibrium constants characterizing the intramolecular transitions by

$$K_{20} = K_1(1 + K_2 + K_2K_3) \quad (80)$$

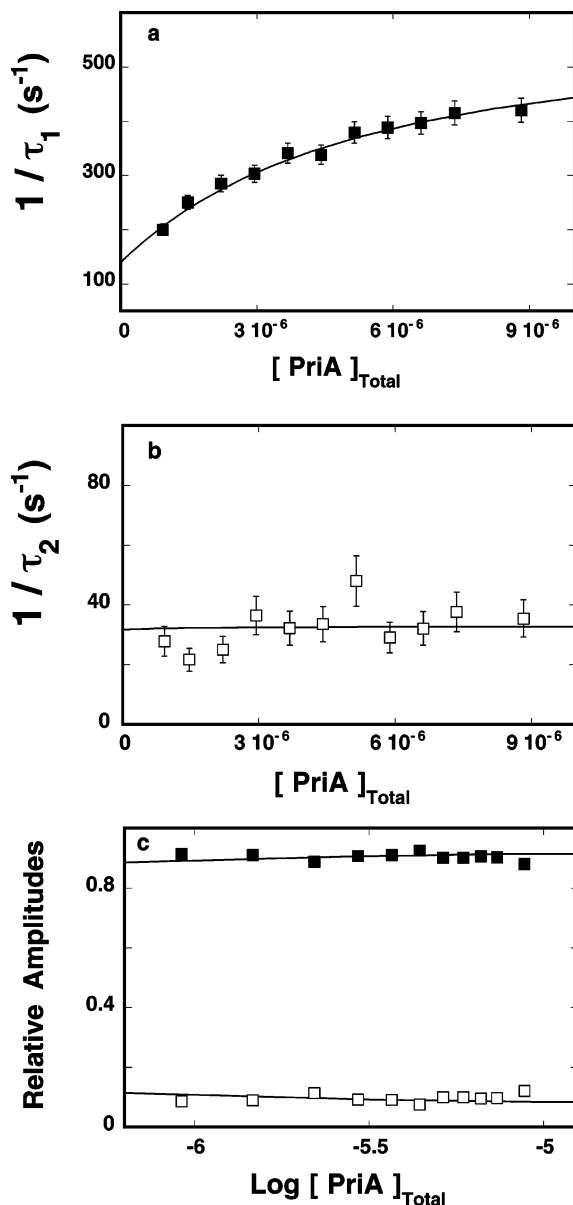


Figure 31. (a) The dependence of the reciprocal relaxation time $1/\tau_1$ for the binding of the PriA helicase to the ssDNA 20-mer $\epsilon\epsilon\text{A}(\text{p}\epsilon\text{A})_{19}$ in 10 mM sodium cacodylate/HCl (pH 7.0, 10 °C) containing 100 mM NaCl upon the total concentration of the enzyme. The concentration of the nucleic acid is $3 \times 10^{-7} \text{ M}$ (final concentration). The solid line is the nonlinear least-squares fit according to the three-step sequential mechanism with the overall partial equilibrium constant $K_1 = 2 \times 10^5 \text{ M}^{-1}$ and the rate constants $k_2 = 230 \text{ s}^{-1}$, $k_{-2} = 140 \text{ s}^{-1}$, $k_3 = 1 \text{ s}^{-1}$, and $k_{-3} = 32 \text{ s}^{-1}$ (details in text). (b) The dependence of the reciprocal relaxation time $1/\tau_2$ for the binding of the PriA helicase to the ssDNA 20-mer $\epsilon\epsilon\text{A}(\text{p}\epsilon\text{A})_{19}$ upon the total concentration of the enzyme. The solid line is the nonlinear least-squares fit according to the three-step sequential mechanism with the same overall partial equilibrium constant, K_1 , and rate constants, k_2 , k_{-2} , k_3 , and k_{-3} , as those in panel a. (c) The dependence of the individual relaxation amplitudes, A_1 and A_2 , for the binding of the PriA helicase to the 20-mer $\epsilon\epsilon\text{A}(\text{p}\epsilon\text{A})_{19}$ in buffer C (pH 7.0, 10 °C) containing 100 mM NaCl upon the total concentration of the enzyme: A_1 (■), A_2 (□). The solid lines are nonlinear least-squares fits according to the three-step sequential mechanism with the relative fluorescence intensities $F_2 = 1.02$, $F_3 = 2.75$, and $F_4 = 7.0$. The maximum fluorescence increase of the nucleic acid is taken from the equilibrium fluorescence titration in the same solution conditions as $\Delta F_{\text{max}} = 2.2$. The rate constants are the same as those in panel a. Reprinted with permission from ref 123. Copyright 2003 American Chemical Society.

where $K_2 = k_2/k_{-2}$ and $K_3 = k_3/k_{-3}$ (see below). The above relationship reduces the number of independent parameters in fitting the relaxation times to four. Subsequently, the obtained rate constants are used as starting values in the fitting of the individual amplitudes and extract relative molar fluorescence parameters, that is, to assess the conformational state of the helicase–nucleic acid complex in each intermediate. This is accomplished using the matrix projection operator technique^{116,123} (see above). This part of the analysis uses the value of the maximum relative increase of the ssDNA fluorescence accompanying the complex formation, $\Delta F_{\max} = 2.2 \pm 0.1$, which is known from independent equilibrium fluorescence titrations.^{47,48} The ΔF_{\max} parameter can be analytically expressed by an expression analogous to eq 75. The refinement of the values of rate constants and molar fluorescence parameters is accomplished by global fitting of all relaxation times and amplitudes. The solid lines in Figure 31a,b,c are nonlinear least-squares fits of the relaxation times and amplitudes, according to the proposed mechanism, using a single set of rate and spectroscopic parameters.

Comparison between the value of the overall partial equilibrium constant, $K_1 = (2 \pm 0.5) \times 10^5 \text{ M}^{-1}$, with the overall equilibrium constant, $K_{20} = (4.9 \pm 0.6) \times 10^5 \text{ M}^{-1}$ indicates that the fast bimolecular step provides the major part of the free energy of binding, ΔG° , of the enzyme to the ssDNA. Nevertheless, very low, if any, fluorescence change of the nucleic acid in this step strongly suggests the lack of any significant conformational changes of the nucleic acid structure accompanying the formation of the (P)₁ intermediate (see below). The transition to the second intermediate, (P)₂, is also a fast process with the forward rate constant $k_2 = 230 \pm 40 \text{ s}^{-1}$. However, in contrast to (P)₁, there is a large molar fluorescence increase ($F_3 = 2.75 \pm 0.15$) accompanying the formation of (P)₂, an indication of a large conformational change of the DNA, as compared to the free nucleic acid.¹²³ Nevertheless, the value of $k_{-2} = 140 \pm 30 \text{ s}^{-1}$ indicates that the enzyme can quickly return to the (P)₁ intermediate. On the other hand, the transition to (P)₃ is much slower with the forward rate constant k_3 being approximately more than 2 orders of magnitude lower than k_2 . Also, the (P)₂ ↔ (P)₃ transition is accompanied by a dramatically large increase of the nucleic acid fluorescence.¹²³ The obtained rate constants for the second and third step provide partial equilibrium constants $K_2 = 1.6 \pm 0.6$ and $K_3 \approx 0.031$. Thus, only the second step contributes an additional favorable contribution to the ΔG° , while the (P)₂ ↔ (P)₃ transition is energetically unfavorable.

3.5.3.2. Dependence of the Kinetics of PriA–ssDNA Interactions upon the Length of the ssDNA Substrate. Thermodynamic studies of PriA interactions with ssDNA showed that although the total site size of the PriA–ssDNA complex is 20 ± 3 nucleotides, the proper DNA-binding site of the enzyme occludes only 8 ± 1 nucleotides.^{47,48} To obtain further insight into the dynamics of the monomeric helicase–ssDNA interactions, stopped-flow kinetic studies have been performed with a series of ssDNA oligomers of different lengths. The shortest oligomer, dεA(pεA)₇, contains eight residues, that is, it corresponds to the determined maximum size of the proper ssDNA-binding site. The longest oligomer, dεA(pεA)₂₃, is three times longer than the proper ssDNA-binding site but can still accept only one PriA molecule.^{47,48}

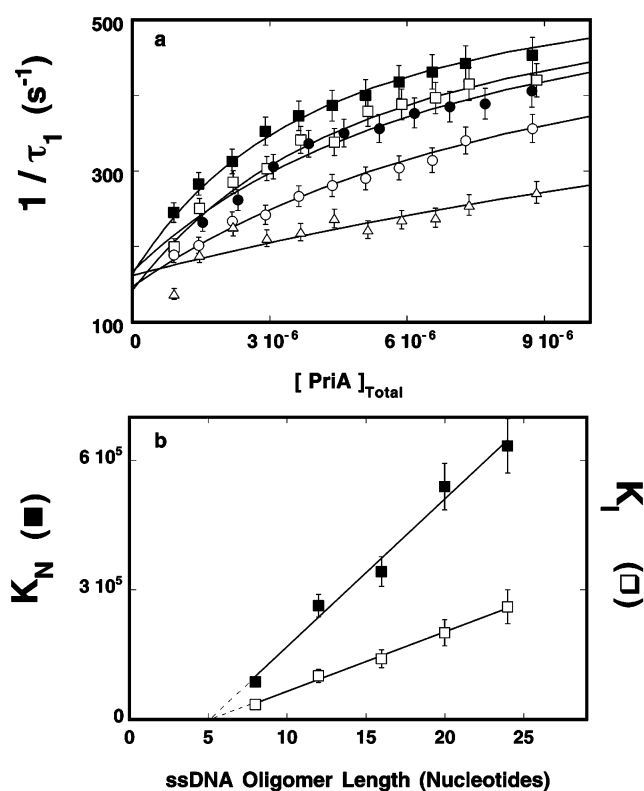


Figure 32. (a) The dependence of the reciprocal relaxation time, $1/\tau_1$, for the binding of the PriA helicase to the ssDNA oligomers differing in the number of nucleotides in 10 mM sodium cacodylate (pH 7.0, 10 °C) containing 100 mM NaCl upon the total concentration of the enzyme: (□) 8-mer, dεA(pεA)₇; (○) 12-mer, dεA(pεA)₁₁; (●) 16-mer, dεA(pεA)₁₅; (◻) 20-mer, dεA(pεA)₁₉; (■) 24-mer, dεA(pεA)₂₃. The concentration of the nucleic acids is 3×10^{-7} M (oligomer). The solid lines are nonlinear least-squares fits according to the three-step sequential mechanism. (b) The dependence of the overall equilibrium constant, K_N (■), and overall partial equilibrium constant, K_1 (□), characterizing the bimolecular step for PriA binding to ssDNA oligomers with different numbers of nucleotides upon the length of the ssDNA oligomer (nucleotides). The solid lines are the linear least-squares fits of the plots (details in text). The dashed lines are extrapolations of the plots to zero value of the corresponding equilibrium constant. Reprinted with permission from ref 123. Copyright 2003 American Chemical Society.

For all examined ssDNAs, the experimental kinetic traces required a two-exponential fit and the amplitudes of the two resolved relaxation processes account for the total amplitude of the kinetic traces. The reciprocal relaxation time, $1/\tau_1$, for the association of PriA with various ssDNA oligomers as a function of the total PriA concentration is shown in Figure 32a. There are two key aspects of these data. First, the kinetic mechanism is independent of the length of the ssDNA. Therefore, the association of PriA with all examined ssDNA oligomers is described by the same three-step mechanism (see above). With increasing length of the ssDNA oligomer, the values of $1/\tau_1$ show a more and more pronounced hyperbolic dependence upon PriA concentration and higher plateau at saturating concentrations of the helicase (Figure 32a). Second, the plots intercept the reciprocal relaxation time axis at a very similar point, indicating that the values of k_{-2} are very close for all oligomers.¹²³ Contrary to $1/\tau_1$, the values of $1/\tau_2$ exhibit little dependence upon the length of ssDNA oligomers, that is, the (P)₂ ↔ (P)₃ transition is only slightly affected by the length of the ssDNA.

The analyses of relaxation times and amplitudes indicate that the increasing hyperbolic dependence of $1/\tau_1$ results from the increasing values of the overall partial equilibrium constant, K_1 , characterizing the bimolecular step in the enzyme binding to the longer ssDNA oligomers,¹²³ and not from the increased values of the rate constant, k_2 . In fact, both k_2 and k_{-2} are, within the experimental accuracy, unaffected by the length of the nucleic acid.¹²³ The same is true for k_3 and k_{-3} . Thus, the entire effect of the different length of the ssDNA oligomers on the dynamics of the enzyme interactions with the ssDNA is confined to the bimolecular step of the reaction. Figure 32b shows the overall equilibrium constant K_N and the overall partial equilibrium constant K_1 for PriA binding to ssDNA oligomers with a different number of nucleotides as functions of the ssDNA oligomer length. A very characteristic feature of these plots is that they are strictly linear, although their slopes are different. Moreover, extrapolations of the plots to zero value of the corresponding equilibrium constant intercept the DNA length axis at very similar points (Figure 31b). Such strictly linear behavior of both K_N and K_1 as a function of the length of ssDNA oligomers can quantitatively be understood in terms of the existence of several potential binding sites on the ssDNA oligomers resulting from the fact that the size of the proper ssDNA-binding site of the PriA helicase, n , is significantly smaller than the total site size of the enzyme–ssDNA complex, 20 ± 3 .^{47,48}

In terms of the potential binding sites and partial equilibrium constants, the overall binding constant, K_N , for the helicase binding to the ssDNA oligomer containing N nucleotides is analytically defined as

$$K_N = K_1(1 + K_2 + K_2K_3) \quad (81a)$$

$$K_N = (N - n + 1)K_1(1 + K_2 + K_2K_3) \quad (81b)$$

and

$$K_N = NK_1(1 + K_2 + K_2K_3) - (n - 1)K_1(1 + K_2 + K_2K_3) \quad (81c)$$

where N is the total length of the ssDNA oligomer in nucleotides and K_1 is the partial equilibrium constant characterizing the bimolecular step, that is, $K_1 = k_1/k_{-1}$. Thus, it is evident that K_N must be a linear function of N with the slope $\partial K_N/\partial N = K_1(1 + K_2 + K_2K_3)$ (Figure 31b). Moreover, K_N is equal to 0 for $N = n - 1$, that is, no binding will be observed for the ssDNA oligomer shorter by one residue than the size of the proper ssDNA-binding site. Thus, the plot of K_N as a function of the nucleic acid length will intercept the N axis at the value of $N = n - 1$. Analogously, the overall partial equilibrium constant, K_1 , characterizing the bimolecular step is defined in terms of potential binding sites and partial equilibrium constant, K_1 , as

$$K_1 = (N - n + 1)K_1 \quad (82a)$$

and

$$K_1 = NK_1 - (n - 1)K_1 \quad (82b)$$

Thus, the plot of K_1 as a function of N is linear with respect to N with the slope $\partial K_1/\partial N = NK_1$. The value of $\partial K_1/\partial N$ is lower than $\partial K_N/\partial N$ by a factor of $(1 + K_2 + K_2K_3)$. The plot intercepts the nucleic acid axis at $N = n - 1$.

Extrapolations of both plots in Figure 31b provide $n = 6.3 \pm 1$ as compared to the maximum value of $n = 8 \pm 1$ obtained in independent thermodynamic analyses^{47,48} (see above).

Unlike the previously discussed DnaB protein, the bimolecular step in the mechanism of the PriA helicase is very fast, beyond the resolution stopped-flow measurements, and independent of the length of the ssDNA oligomer. The amplitude analyses show that there is very little, if any, change of the nucleic acid fluorescence accompanying the formation of the $(P)_1$ intermediate in the proposed mechanism (see above). As discussed above, interpretation of the fluorescence changes of etheno-ssDNA oligomers is facilitated by the fact that dramatic quenching of the nucleic acid fluorescence is well understood in terms of intramolecular collisions due to the motion of ϵA separated by a close distance in the DNA.^{129,130} Thus, the extent of the fluorescence increase in the protein–nucleic acid complexes reflects the conformational changes of the nucleic acid (immobilization and separation of bases) that limits these quenching processes.^{129,130} The lack of detectable fluorescence change in the $(P)_1$ intermediate indicates that the ssDNA conformation in this intermediate is very similar to the conformation of the free DNA, that is, unaffected by the protein binding. This is an expected behavior for a collision complex.^{124,125} However, very fast protein/nucleic acid conformational change occurring in the formation of $(P)_1$ cannot be completely excluded on the basis of stopped-flow measurements alone.

The step following the bimolecular association is also fast and provides very modest additional stabilization of the protein–ssDNA complex. Nevertheless, the values of k_2 and k_{-2} indicate that the lifetime of the $(P)_2$ intermediate is in the range of several milliseconds. Moreover, amplitude analyses show that, contrary to the $(P)_1$ intermediate, the formation of $(P)_2$ is accompanied by a strong nucleic acid fluorescence increase, indicating large changes in the DNA structure in the complex.¹²³ These data indicate that in the second step the adjustment of the ssDNA conformation to the structure of the binding site occurs, that is, it is the recognition step in the binding process. Although the subsequent transition to the $(P)_3$ intermediate is also accompanied by a strong conformational change of the nucleic acid, it is energetically very unfavorable and, as a result, contributes very little to the total population of the helicase–DNA complexes. Kinetics of the nucleotide hydrolysis or of the dsDNA unwinding by the PriA helicase is unknown. However, the favorable free energy change of the $(P)_1 \leftrightarrow (P)_2$ transition, adjustment of the DNA structure to the structure of the DNA-binding site, and the lifetime of the complex in the range of milliseconds suggest that the $(P)_2$ and not the $(P)_3$ intermediate may play an important role in the catalytic activities of the enzyme.¹²³

Recall, the proper DNA-binding site of the PriA helicase is located in the central part of the helicase molecule and occludes only 8 ± 1 residues.^{47,48} In other words, the protein matrix extends over approximately six residues on both sides of the binding site without engaging in interactions with the DNA (Figure 6). Therefore, if the protein matrix outside the proper ssDNA-binding site does not engage in interactions with the DNA, for any ssDNA oligomer longer than ~ 8 nucleotides, the overall binding constant K_N and the overall partial equilibrium constant K_1 characterizing the bimolecular step must contain a factor $N - n + 1$ resulting from the

existence of $N - n + 1$ potential binding sites (eqs 81a, 81b, and 81c). On the other hand, the partial equilibrium constants K_2 and K_3 should be independent of the length of the ssDNA oligomers. Both thermodynamic and kinetic analyses clearly show that the dependence of K_N and K_1 upon the length of ssDNA oligomers strictly follows the predicted linear dependence upon N (Figure 32b), while partial equilibrium constants characterizing intramolecular steps are, within experimental accuracy, independent of N . Such behavior could only be observed if the proper ssDNA-binding site of the PriA helicase is exclusively involved in interactions with the ssDNA.

The plot of K_N and K_1 intercept the oligomer length axis at $n = 6.3 \pm 1$ indicating that the ssDNA-binding site of the helicase encompasses ~ 2 nucleotides less than that determined in thermodynamic analyses (see above).^{47,48} This difference should not be surprising. The shortest length of the ssDNA oligomer that can be accessed in direct thermodynamic studies corresponds to the shortest oligomer whose binding can still be detected, that is, eight nucleotides. Combined application of the thermodynamic and kinetic studies provides a much more accurate estimate of the site size of the proper ssDNA-binding site. The linear dependence of only K_N upon N does not establish the existence of a purely statistical effect in the observed increase of the overall binding constant. On the other hand, the linear dependence of both K_N and K_1 and the common interception point on the ssDNA oligomer length axis (Figure 32b) constitutes very strong evidence that the observed increase of both equilibrium binding constants with the length of the ssDNA oligomer results from the statistical effect of the potential binding sites. It also indicates that a ssDNA patch of approximately six nucleotides in length is absolutely necessary for the enzyme to form a stable complex with the ssDNA and confirms the lack of any "end effect" in the PriA binding to the ssDNA, previously determined in thermodynamic studies.^{47,48}

In this context, it is not surprising that dynamics of the enzyme binding to various ssDNA oligomers in subsequent steps, $(P)_1 \leftrightarrow (P)_2$ and $(P)_2 \leftrightarrow (P)_3$, is independent of the length of the oligomers. It simply means that, as long as the stretch of approximately six nucleotides of the ssDNA is available and a stable $(P)_1$ intermediate is formed, it undergoes similar transitions, little dependent on the surrounding nucleic acid. This is because the protein matrix outside the ssDNA-binding site does not engage in interactions with the ssDNA. However, amplitude analyses indicate that molecular fluorescence intensities characterizing intermediates $(P)_2$ and $(P)_3$ are different among different ssDNA oligomers.¹²³ Recall that the molecular fluorescence intensities characterize the relative fluorescence change of the entire ssDNA oligomer with respect to the same free DNA. Because the enzyme associates with the ssDNA using only its proper ssDNA-binding site, these data indicate that the conformational transitions of the nucleic acid generated at the binding site–ssDNA interface extend to the rest of the bound nucleic acid molecule, although part of the ssDNA is not involved in direct interactions with the helicase.

3.5.3.3. Functional Implication of the Kinetic Data for the PriA Helicase Activities. It is remarkable that although the DNA binding site encompasses only 6.3 ± 1 nucleotides within ~ 20 nucleotides of the total site size of the PriA–ssDNA complex, the surrounding protein matrix does not enter in thermodynamically and kinetically detectable inter-



Figure 33. Schematic model of the PriA helicase–ssDNA complex. The proper ssDNA-binding site of the enzyme engages in interactions with only $n = 6 \pm 1$ nucleotides. The site is located on a separated structural domain of the enzyme, most probably the helicase domain, which protrudes from the remaining protein matrix. Only the DNA-binding site engages in strong interactions with the nucleic acid. The domain containing the DNA-binding site is placed in the central part of the helicase molecule. The protein matrix protrudes symmetrically on both sides of the strong ssDNA-binding site without engaging in interactions with the nucleic acid. As a result, the total site size that the PriA helicase–ssDNA complex occludes is $n = 20 \pm 3$ nucleotides. Reprinted with permission from ref 123. Copyright 2003 American Chemical Society.

actions with the ssDNA in the examined nucleic acid and protein concentration ranges.^{47,48,123} The length dependence of the kinetics of the PriA binding to the ssDNA indicates that even in the first intermediate, $(P)_1$, between the protein and the DNA, only the DNA-binding site but not the surrounding protein matrix makes the first and only contact with the nucleic acid. Moreover, the lack of any length effect on following intermediates $(P)_2$ and $(P)_3$ clearly indicates that the subsequent conformational transitions of the helicase–DNA complex do not engage the protein matrix outside the proper ssDNA-binding site. This behavior can be understood in the context of the structural model of the PriA helicase, as schematically depicted in Figure 33. To achieve the experimentally observed independence from the rest of the enzyme molecule, the DNA-binding site must protrude from the remaining protein matrix. Although the three-dimensional structure of the PriA helicase is still unknown, the obtained data strongly suggest that the ssDNA-binding site is placed on a well-defined structural domain of the protein, most probably the helicase domain, located in the central part of the enzyme molecule.^{47,48,123}

In chromosomal DNA replication, the role of the PriA helicase seems to be predominantly related to the initiation of the restarting of DNA replication after the replication fork stalls at the damaged DNA sites.^{134,135} The helicase activity of the protein would allow the unwinding of the duplex conformation of the lagging strand of the fork, preparing it for the binding of the DnaB helicase and assembling the preprimosome complex. Analyses of the enzyme activity on different synthetic DNA substrates strongly suggest that the enzyme requires short ssDNA gaps to initiate its helicase activity.¹³⁴ In other words, the PriA helicase is able to efficiently search and recognize very short ssDNA gaps in the presence of an overwhelmingly large excess of the dsDNA conformation. In fact, the optimal length of the recognized ssDNA gap is approximately five nucleotides. It is clear that the data obtained in this work indicating that the ssDNA-binding site occludes approximately six nucleotides corroborate very well with these findings. Thus, the helicase active site is built to efficiently search small ssDNA patches, that is, it protrudes from the rest of the protein molecule. Such structure of the active site would be an evolutionary adaptation of the PriA helicase to perform

specific ssDNA gap searching and recognition. Nevertheless, although the DNA-binding capability of the protein matrix that occludes the remaining nucleotides of the total site size is not detectable, it may play an important role in orienting the enzyme within the specific DNA substrate structures, for example, arrested replication fork structures. In such situations, when the local concentration of the DNA becomes very high, even low affinity binding sites may begin to play a role in the binding. As we pointed out above, these inherently weak interactions with more complex DNA structures may be under ATP/ADP binding/hydrolysis control.

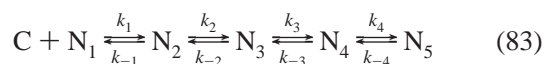
The very fast bimolecular step and the fast following recognition step make PriA binding to the ssDNA very different from the DnaB hexameric helicase but similar to other well-known fast protein–nucleic acid recognition reactions, including aminoacyl-tRNA synthetases, ribosomal proteins, and mammalian DNA repair polymerase β .^{94–97,136,137} A common feature of these systems is that the enzyme/protein recognizes a specific nucleic acid substrate (sequence, specific nucleic acid structure, or both) among many substrates or within the context of several nonfunctional binding sites. Fast association and dissociation reactions in the first binding step and fast recognition step provide the protein with a means to quickly examine the structure of the encountered nucleic acid. As mentioned above, the PriA helicase is involved in restarting the DNA replication when the replication fork encounters damaged DNA.^{134,135} Thus, very fast binding and recognition steps would allow the enzyme to efficiently recognize the specific ssDNA gap on the lagging strand and initiate the assembling of the preprimosome complex.

3.5.4. Protein–Nucleic Acid System with Fast Bimolecular Step Having Detectable Amplitude. Kinetics of the Human Pol β Binding to the ssDNA

Human polymerase β is the analogue of the rat pol β in human cells.^{85–87} The enzymes differ by 14 amino acids in the primary structures and share the same three-dimensional structure, which includes the presence of the 8-kDa and 31-kDa structural and functional domains. Moreover, they possess similar, although not identical, thermodynamic and kinetic properties of interactions with nucleic substrates.^{94–97} Quantitative thermodynamic data have shown that both human and rat pol β bind the ssDNA in two binding modes that differ in the number of occluded nucleotides, the (pol β)₁₆ and (pol β)₅ binding modes.^{43,44} Both binding modes differ in affinities and abilities to induce conformational changes in the nucleic acid. The intrinsic affinity of the enzyme in the (pol β)₁₆ binding mode is approximately an order of magnitude higher than the affinity in the (pol β)₅ binding mode. However, the enzyme induces much more profound structural changes in the ssDNA when bound in the (pol β)₅ binding mode, indicating strong base–base separation and immobilization in the complex. In the (pol β)₁₆ binding mode, both the 8-kDa and the 31-kDa domains of the enzyme are involved in interactions with the ssDNA, while in the (pol β)₅ binding mode, the 8-kDa domain is predominately engaged in interactions with the DNA.^{43,44,94–97} Kinetic studies of human pol β interactions with the ssDNA in the (pol β)₁₆ and (pol β)₅ binding modes using the fluorescence stopped-flow technique and discussed in the next section indicate that both binding modes differ significantly in the energetics and dynamics of the formed

intermediates of the binding process.^{94,95} Moreover, the discussed analysis allows the experimenter to access the very fast bimolecular step that is usually left unresolved in standard stopped-flow experiments.

3.5.4.1. Kinetics of Human Pol β Binding to the ssDNA 20-mer in the (Pol β)₁₆ Binding Mode. Stopped-flow kinetic data indicate that the human pol β binds the ssDNA in both the (pol β)₁₆ and (pol β)₅ binding modes including the bimolecular step leading to the formation of the initial complex, which subsequently undergoes three first-order conformational transitions.⁹⁵ For heuristic purposes, we first consider the theoretical behavior of the relaxation times and the amplitudes for such a complex mechanism. This is a sequential reaction between the polymerase (ligand), C, and the nucleic acid (macromolecule), N₁, and described by



The reaction is monitored by the fluorescence change of the macromolecule, that is, the nucleic acid (see below). There are four normal modes of the reaction in the considered mechanism, characterized by four relaxation times and amplitudes.

The reciprocal relaxation times for the reaction described by eq 83 as a function of the free protein ligand concentration are shown in Figure 34. Relaxation times have been obtained by direct numerical determination of the eigenvalues, λ_1 , λ_2 , λ_3 , and λ_4 , of the coefficient matrix at a given free protein concentration, C_1 , using the identities of $1/\tau_1 = -\lambda_1$, $1/\tau_2 = -\lambda_2$, $1/\tau_3 = -\lambda_3$, and $1/\tau_4 = -\lambda_4$. The selected rate constants are $k_1 = 3 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$, $k_{-1} = 300 \text{ s}^{-1}$, $k_2 = 600 \text{ s}^{-1}$, $k_{-2} = 250 \text{ s}^{-1}$, $k_3 = 50 \text{ s}^{-1}$, and $k_{-3} = 10 \text{ s}^{-1}$, $k_4 = 10 \text{ s}^{-1}$, and $k_{-4} = 3 \text{ s}^{-1}$. Thus the first step is very fast, close to the diffusion controlled one; the second step is also fast and occurs in the few millisecond time range. The remaining two steps are significantly slower than the first two steps.⁹⁵ The three relaxation times, $1/\tau_2$, $1/\tau_3$, and $1/\tau_4$, for the considered sequential four-step reaction show characteristic hyperbolic dependence upon ligand concentration reaching the plateau values at higher ligand concentrations. Thus, in the high ligand concentration range, the values of $1/\tau_2$, $1/\tau_3$, and $1/\tau_4$ become independent of the ligand concentration. It is also evident that, for the selected values of the rate constants, $1/\tau_3$ and $1/\tau_4$ reach their plateaus at a lower ligand concentration than $1/\tau_2$ as a result of the favorable contributions of the subsequent steps to the total free energy of binding (see above). On the other hand, the largest reciprocal of relaxation time, $1/\tau_1$, shows typical behavior for the relaxation time characterizing the bimolecular binding process.^{114,116}

Notice that with the selected rate constants, the value of $1/\tau_1$ is already above 1000 s^{-1} and is strongly increasing even in the low ligand concentration range. Such a high value of $1/\tau_1$ makes its accurate determination practically impossible in a standard stopped-flow experiment with a typical dead time of ~ 1 – 2 ms . If the first fast relaxation process has a detectable amplitude, it will be hidden in the unresolved fast change of the observed signal from the time $t = 0$ to the steady-state value of the signal before the flow stops. However, if the amplitude of the fast process can be recovered from the data and the simultaneous analysis of all amplitudes and relaxation times is performed over a large range of the enzyme or nucleic acid concentration or both, then the values of rate constants for the bimolecular step can be obtained with adequate accuracy (see below). This is

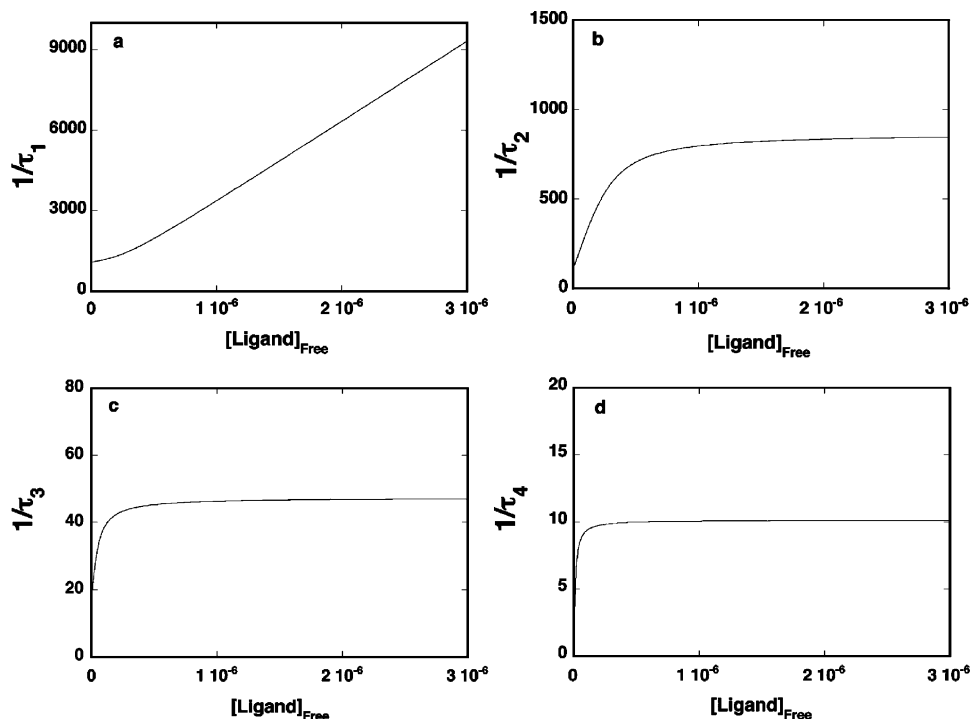


Figure 34. Computer simulation of the dependence of reciprocal relaxation times for the four-step sequential mechanism of ligand binding to a macromolecule defined by eq 83 upon the free ligand concentration: (a) $1/\tau_1$; (b) $1/\tau_2$; (c) $1/\tau_3$; (d) $1/\tau_4$. Relaxation times have been obtained by numerically determining the eigenvalues of the matrix of coefficient \mathbf{M} ($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) and using identities $1/\tau_1 = -\lambda_1$, $1/\tau_2 = -\lambda_2$, $1/\tau_3 = -\lambda_3$ and $1/\tau_4 = -\lambda_4$. The simulations have been performed using rate constants $k_1 = 3 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$, $k_{-1} = 300 \text{ s}^{-1}$, $k_2 = 600 \text{ s}^{-1}$, $k_{-2} = 250 \text{ s}^{-1}$, $k_3 = 50 \text{ s}^{-1}$, $k_{-3} = 10 \text{ s}^{-1}$, $k_4 = 10 \text{ s}^{-1}$, and $k_{-4} = 3 \text{ s}^{-1}$. Reprinted with permission from ref 95. Copyright 2001 American Chemical Society.

because the amplitudes, like relaxation times, constitute an independent set of data and are very sensitive functions of spectroscopic parameters and rate constants of the reaction.^{95,114}

The effect of the rate constants characterizing the bimolecular step on the dependence of individual amplitudes, A_1, A_2, A_3 , and A_4 , of the reaction upon the ligand concentration is shown in Figure 35. The amplitudes are normalized, that is, they are expressed as fractions ($A_i/\sum A_i$) of the total amplitude $A_T = \sum A_i$. The computer simulations were performed for different sets of rate constants for the bimolecular step but the same set of spectroscopic parameters characterizing each intermediate. Also, although the values of the rate constants are changed, the equilibrium constant for the first step, $K_1 = k_1/k_{-1}$, is held constant, thus preserving the same free energy change accompanying the first step of the reaction. The selected molar fluorescence intensities of intermediates with respect to the fluorescence intensity of the free ligand $F_1 = 1$ are $F_2 = 1.3$, $F_3 = 2$, $F_4 = 2$, and $F_5 = 2$.

The plots show several very characteristic features of the behavior of the examined system. At low protein ligand concentrations mainly the amplitude A_4 of the fourth normal mode of the reaction contributes to the observed kinetic trace at any value of the rate constants k_1 and k_{-1} . Such behavior is not a result of a dominant fluorescence intensity of N_5 , as often suggested, but a low efficiency of the formation of N_2, N_3 , and N_4 . The amplitude of the third normal mode, A_3 , is always bell-shaped; however, the value of its maximum decreases with increased values of k_1 and k_{-1} . Notice that if the experiments are performed at higher ligand concentrations, only one side of the bell-shaped curve of A_3 will be recorded and the amplitude will appear as continuously decaying with the increasing ligand concentration. Amplitude

A_2 shows behavior similar to that of A_3 , but the values of k_1 and k_{-1} affect not only the value and the location of the A_2 maximum but also the amplitude shape.

With the selected values of the relative molar fluorescence intensities for the intermediates, the first relaxation mode of the reaction is characterized by detectable amplitude. The amplitude of the first normal mode, A_1 , has very low values at lower ligand concentrations, independent of the values of k_1 and k_{-1} . However, the A_1 contribution increases with the increased values of k_1 and k_{-1} . Moreover, at higher ligand concentration, A_1 begins to contribute substantially to the relaxation process. For the selected values of F_2 , lower than other relative molar fluorescence intensities, A_1 assumes both negative and positive values and exhibits the maximum depending on the values of k_1 and k_{-1} . Also, the shape and the location of the A_1 minimum depend on k_1 and k_{-1} . Notice that although the value of equilibrium constant K_1 is unchanged, all amplitudes are shifted toward high ligand concentrations at lower values of k_1 and k_{-1} .

The effect of the spectroscopic parameter F_2 characterizing the first intermediate N_2 on the behavior of all amplitudes is shown in Figure 36. Whether or not F_2 is the major fluorescence intensity, the amplitude A_4 always dominates the observed relaxation process at low ligand concentration. The seemingly small change of F_2 substantially affects the remaining amplitudes including the shape of both A_2 and A_3 . Moreover, in the case where F_2 is the major fluorescence intensity (Figure 36c), A_3 and A_2 become negative in the high ligand concentration range. As expected, F_2 has the most dramatic effect on A_1 . When F_2 is lower than the molar fluorescence intensities of the remaining intermediates, A_1 is negative over a large range of the ligand concentration and shows a well-pronounced minimum (Figure 36a). When the value of F_2 is the same or larger than those of F_3, F_4 ,

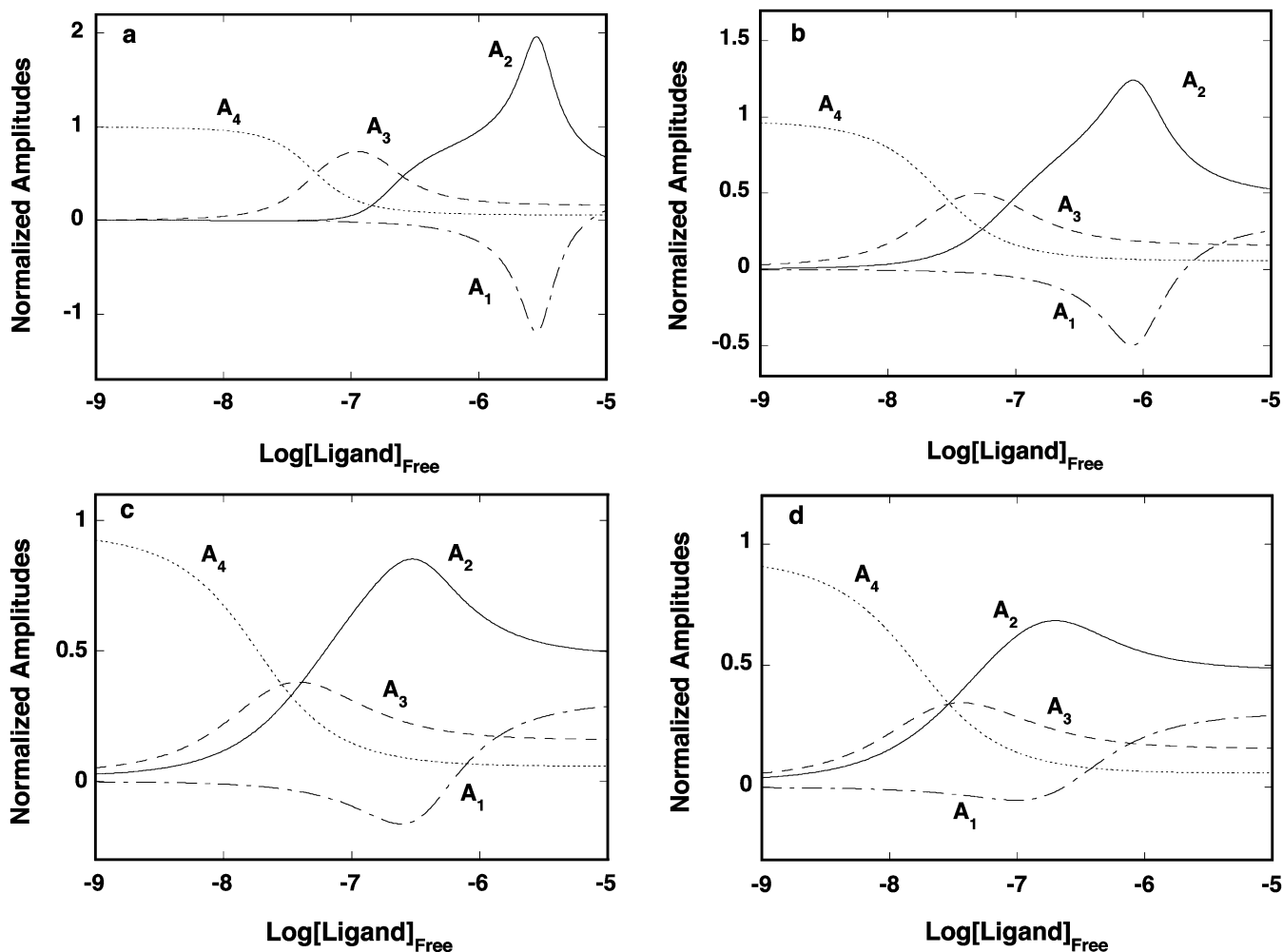


Figure 35. Computer simulation of the dependence of individual, A_1 , A_2 , A_3 , and A_4 , relaxation amplitudes for the four-step sequential mechanism of a ligand binding to a macromolecule defined by eq 83 upon the logarithm of the free ligand concentration for the same set of relative molar fluorescence intensities but with different values of rate constants k_1 and k_{-1} characterizing the bimolecular binding step: (a) $k_1 = 3 \times 10^8 \text{ M}^{-1} \text{ s}^{-1}$, $k_{-1} = 30 \text{ s}^{-1}$; (b) $k_1 = 1 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$, $k_{-1} = 100 \text{ s}^{-1}$; (c) $k_1 = 3 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$, $k_{-1} = 300 \text{ s}^{-1}$; (d) $k_1 = 6 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$, $k_{-1} = 600 \text{ s}^{-1}$; A_1 (---), A_2 (—), A_3 (---), A_4 (⋯). The relative fluorescence intensities F_2 , F_3 , F_4 , and F_5 corresponding to intermediates N_2 , N_3 , N_4 , and N_5 are 1.3, 2, 2, and 2, respectively. The fluorescence of the free macromolecule, N_1 , is taken as $F_1 = 1$. The individual amplitudes are expressed as fractions of the total amplitude, A_T . The simulations have been performed with the same set of remaining rate constants: $k_2 = 600 \text{ s}^{-1}$, $k_{-2} = 250 \text{ s}^{-1}$, $k_3 = 50 \text{ s}^{-1}$, $k_{-3} = 10 \text{ s}^{-1}$, $k_4 = 10 \text{ s}^{-1}$, and $k_{-4} = 3 \text{ s}^{-1}$. Reprinted with permission from ref 95. Copyright 2001 American Chemical Society.

and F_5 , the amplitude A_1 is always positive and has a typical sigmoidal shape. It becomes a dominant relaxation effect only at higher ligand concentration. Computer simulations included in Figures 35 and 36 show that if individual amplitude can be determined from the experimental data over a large range of the protein or nucleic acid concentration, the analysis of the individual amplitudes of the reaction should allow the extraction the kinetic and spectroscopic properties of each step and intermediate of the reaction. This includes the fast bimolecular step, even if the relaxation time for this step is not directly resolved in the experiment but its amplitude is.

A fundamental difficulty in examining the kinetics of the protein–DNA interactions for a protein like pol β is that it forms two different binding modes in the complex with the nucleic acid. To examine the kinetics of the protein binding in a particular binding mode, it is imperative to apply conditions where exclusively a particular binding mode exists.^{94,95} Human pol β forms the high-affinity (pol β)₁₆ binding mode in an excess of ssDNA and low protein concentration.⁴⁴ The transition to the low site size (pol β)₅ binding mode occurs only at high enzyme concentrations,

as a result of the increased binding density (degree of binding) of the protein on the nucleic acid lattice or as a result of the limited access to the DNA, as in the case of the short ssDNA oligomers. The difference between the affinities of the (pol β)₁₆ and (pol β)₅ binding modes and the resulting separation of both modes on the protein concentration scale allow us to study the kinetics of the (pol β)₁₆ binding mode formation independently of the formation of the (pol β)₅ binding mode. Thus, equilibrium titration data are absolutely necessary to provide the concentration range of the enzyme that can be used to exclusively examine the formation of the (pol β)₁₆ binding mode.

The kinetics of the formation of the (pol β)₁₆ binding mode by the human pol β has been addressed in the experiments with the ssDNA 20-mer. This oligomer is long enough to interact with the total DNA-binding site of the polymerase, which occludes 16 ± 2 nucleotides.⁴⁴ Recall that fluorescence changes of fluorescent etheno derivatives of the ssDNA accompanying the binding of the human pol β to the nucleic acid were previously used to examine the thermodynamics of the interactions (see above).⁴⁴ However, the observed

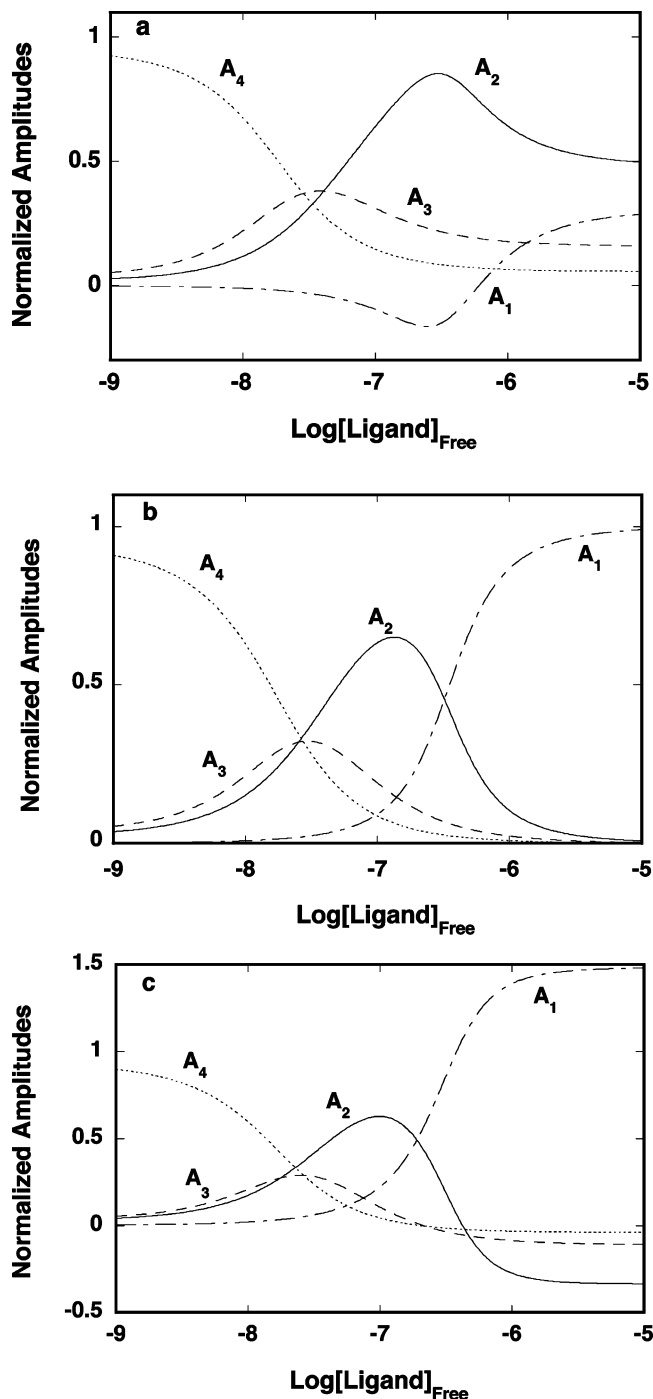


Figure 36. Computer simulation of the dependence of individual, A_1 , A_2 , A_3 , and A_4 , relaxation amplitudes for the four-step sequential mechanism of a ligand binding to a macromolecule defined by eq 83 upon the logarithm of the free ligand concentration for the same set of rate constants but different relative molar fluorescence intensity F_2 corresponding to the intermediate N_2 : (a) $F_2 = 1.3$; (b) $F_2 = 2$; (c) $F_2 = 2.5$; A_1 (---), A_2 (—), A_3 (---), A_4 (···). The fluorescence of the free macromolecule, N_1 , is taken as $F_1 = 1$. The individual amplitudes are expressed as fractions of the total amplitude, A_T . The simulations have been performed with the constant set of rate constants $k_1 = 3 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$, $k_{-1} = 300 \text{ s}^{-1}$, $k_2 = 600 \text{ s}^{-1}$, $k_{-2} = 250 \text{ s}^{-1}$, $k_3 = 50 \text{ s}^{-1}$, $k_{-3} = 10 \text{ s}^{-1}$, $k_4 = 10 \text{ s}^{-1}$, and $k_{-4} = 3 \text{ s}^{-1}$ and with the constant set of values of $F_3 = 2$, $F_4 = 2$, and $F_5 = 2$. Reprinted with permission from ref 95. Copyright 2001 American Chemical Society.

signal is not adequate to quantitatively examine the very complex kinetics of the reaction, partly due to the protein fluorescence contribution to the observed signal. On the other

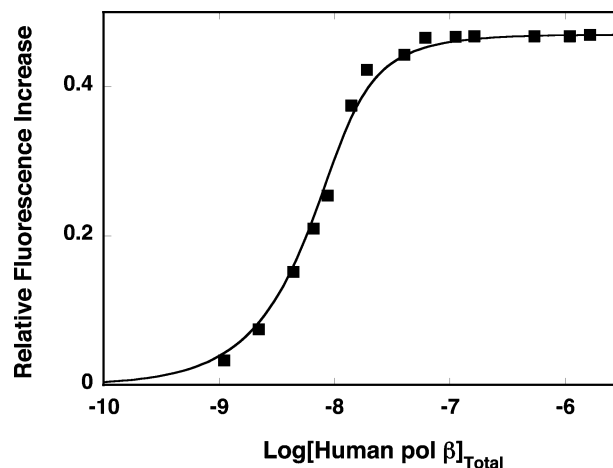


Figure 37. Fluorescence titrations of $\text{dT}(\text{pT})_8\text{-pFlu-(pT)}_{10}$ ($\lambda_{\text{ex}} = 485 \text{ nm}$, $\lambda_{\text{em}} = 520 \text{ nm}$) with human $\text{pol } \beta$ in buffer C (pH 7.0, 10°C) containing 50 mM NaCl . The solid lines are computer fits of the binding isotherms to a single-site binding isotherm $\Delta F = \Delta F_{\text{max}}(K_{20}[\text{enzyme}]_{\text{free}})/(1 + K_{20}[\text{enzyme}]_{\text{free}})$. The concentration of $\text{dT}(\text{pT})_8\text{-pFlu-(pT)}_{10}$ is $1 \times 10^{-8} \text{ M}$ (oligomer) (details in text). Reprinted with permission from ref 95. Copyright 2001 American Chemical Society.

hand, binding of the enzyme to a ssDNA 20-mer containing the fluorescein residue, for example, $\text{dT}(\text{pT})_8\text{-pFlu-(pT)}_{10}$, is accompanied by a significant fluorescence increase of the DNA, providing an excellent signal to monitor the kinetics of the polymerase–ssDNA complex formation. The fluorescein residue has a high quantum yield that allows us to perform experiments at a very low nucleic acid concentration. Any contribution of the protein fluorescence to the observed traces is eliminated by the excitation at 485 nm . Moreover, the mechanism of binding is independent of the location of the fluorescein in the ssDNA oligomer.⁹⁵

Fluorescence titration of $\text{dT}(\text{pT})_8\text{-pFlu-(pT)}_{10}$ with human $\text{pol } \beta$ is shown in Figure 37. The titrations have been performed in the limited protein concentration range ($< 1.5 \times 10^{-6} \text{ M}$) to avoid the formation of the $(\text{pol } \beta)_5$ binding mode. As we discussed above, this is possible because of the large difference in affinities between the two ssDNA-binding modes.^{43,44} Thus, only the formation of the $(\text{pol } \beta)_{16}$ binding mode is observed. The solid line in Figure 37 is a computer fit of the experimental isotherm to a single-site binding isotherm, $\Delta F = \Delta F_{\text{max}}(K_{20}[\text{enzyme}]_{\text{free}})/(1 + K_{20}[\text{enzyme}]_{\text{free}})$, where ΔF_{max} and K_{20} are the maximum fractional fluorescence increase at saturation and the macroscopic binding constant of the human $\text{pol } \beta$ to the 20-mer in the $(\text{pol } \beta)_{16}$ binding mode, respectively. The theoretical curve provides an excellent description of the experimental data confirming the presence of a single binding mode of the protein on the nucleic acid. Thus, the formation of the $(\text{pol } \beta)_{16}$ binding mode on $\text{dT}(\text{pT})_8\text{-pFlu-(pT)}_{10}$ is characterized by $\Delta F_{\text{max}} = 0.47 \pm 0.05$ and $K_{20} = (5 \pm 1.2) \times 10^8 \text{ M}^{-1}$.

The kinetic experiments have been performed under pseudo-first-order conditions by mixing the 20-mer with an excess of the human $\text{pol } \beta$. The stopped-flow kinetic trace of the $\text{dT}(\text{pT})_8\text{-pFlu-(pT)}_{10}$ fluorescence after mixing $1 \times 10^{-7} \text{ M}$ oligomer with $8 \times 10^{-7} \text{ M}$ human $\text{pol } \beta$ (final concentrations) is shown in Figure 38a. To increase the resolution, the plot is shown in logarithmic scale with respect to the time. There is an initial $\sim 2 \text{ ms}$ horizontal part of the trace (not considered in the fitting procedures) that corresponds to the steady-state fluorescence intensity recorded for

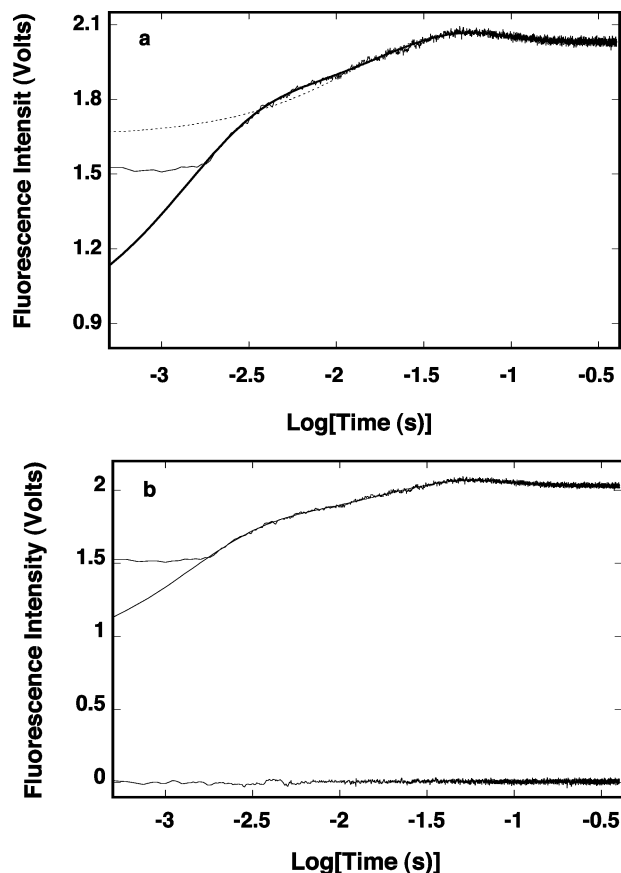


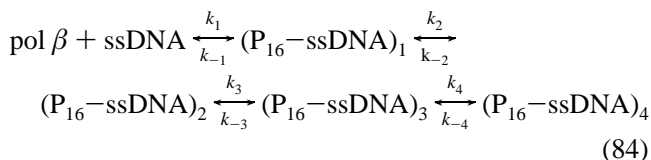
Figure 38. (a) Fluorescence stopped-flow kinetic trace after mixing human pol β with the 20-mer dT(pT)₈-pFlu-(pT)₁₀ in 10 mM sodium cacodylate/HCl (pH 7.0, 10 °C) containing 50 mM NaCl ($\lambda_{\text{ex}} = 485$ nm, $\lambda_{\text{em}} > 495$ nm). The final concentrations of the polymerase and the 20-mer are 8×10^{-7} and 1×10^{-7} M (oligomer), respectively. The experimental kinetic trace is shown in logarithmic scale with respect to time. The horizontal part of the trace is the steady-state value of the fluorescence of the sample recorded 2 ms before the flow stopped. The solid line is the three-exponential, nonlinear least-squares fit of the experimental curve. The dashed line is the nonlinear least-squares fit using the two-exponential function. (b) The same fluorescence stopped-flow trace as that in panel a together with the zero line trace (lower trace), which is obtained after mixing the nucleic acid only with the buffer. The solid line is the same three-exponential, nonlinear least-squares fit of the experimental curve as shown in panel a. Reprinted with permission from ref 95. Copyright 2001 American Chemical Society.

2 ms in the instrument before the flow stops. The observed kinetics is complex, and even a visual inspection shows the presence of multiple steps. A slower relaxation process follows a fast initial process, both characterized by positive amplitudes. However, there is also a third process that is characterized by a negative fluorescence change. The solid line in Figure 38a is a nonlinear least-squares fit of the experimental curve using a three-exponential function, which provides an excellent description of the kinetic trace. It is evident that the two-exponential fit (dashed line) is not sufficient to describe the observed kinetics.

The same kinetic trace depicted in Figure 38a, together with the trace corresponding to the nucleic acid alone (zero line) at the same concentration as that used with the protein but mixed only with the buffer, is shown in Figure 38b. The striking feature shown by these data is that although the three-exponential fit provides an excellent description of the recorded trace, it does not account for the observed total

fluorescence increase resulting from the complex formation. This is very different from the binding of both the DnaB and PriA helicases to the ssDNA as discussed above. The difference between the fluorescence intensity at the end point of the kinetic trace and the zero line recorded for the nucleic acid alone is the total amplitude of the reaction, A_T . Thus, the data indicate that there is at least one additional fast step preceding the observed trace, characterized by the relaxation time τ_1 , which is too short to be determined in the stopped-flow experiment. Therefore, the association of human pol β with the ssDNA 20-mer in the (pol β)₁₆ binding mode is a process that includes at least four normal modes.

The reciprocal relaxation times, $1/\tau_2$, $1/\tau_3$, and $1/\tau_4$, extracted from the experimental trace as a function of the total human pol β concentration are shown in Figure 39. The value of $1/\tau_2$ shows hyperbolic dependence upon human pol β concentration, while both $1/\tau_3$ and $1/\tau_4$ are, within experimental accuracy, independent of enzyme concentration. Such behavior indicates that all three relaxation times characterize the intramolecular transitions of the complex (Figure 34).^{95,114,116} Therefore, the minimum mechanism that can account for the observed dependence of the relaxation times upon the human pol β concentration and the presence of the third unresolved process is a four-step, sequential binding reaction in which bimolecular association is followed by three isomerization steps, as described by



Although one cannot determine the relaxation time, τ_1 , for the fast normal mode, the amplitude of this mode, A_1 , can be obtained from the known amplitudes of the second, third, and fourth normal modes, A_2 , A_3 , and A_4 , and the total amplitude of the reaction, A_T , as

$$A_1 = A_T - A_2 - A_3 - A_4 \quad (85)$$

The dependence of the normalized individual amplitudes, A_1 , A_2 , A_3 and A_4 , of the four relaxation steps upon the human pol β concentration is shown in Figure 40. In the examined enzyme concentration ranges, all four amplitudes contribute to the A_T , even at the lowest enzyme concentration. The positive amplitude A_2 goes through a maximum, while the positive A_3 steadily decreases with the human pol β concentration. Amplitude A_4 has a negative value over the entire polymerase concentration range. On the other hand, A_1 initially has a negative value and shows a minimum at intermediate enzyme concentrations. As the concentration of the polymerase increases, A_1 assumes positive values and becomes a dominant relaxation effect. Such behavior of the individual amplitudes is in full agreement with the proposed mechanism (eq 84).

Because only three relaxation times are available from the experiment, the determination of rate constants of particular steps and molar fluorescence parameters characterizing each intermediate requires the simultaneous analyses of both the relaxation times and the amplitudes. We applied the following strategy to obtain all rate and spectroscopic parameters of the system. We utilized the fact that we know the value of the macroscopic binding constant, $K_{20} = (5 \pm 1.2) \times 10^8$ M⁻¹, independently obtained in the same solution conditions

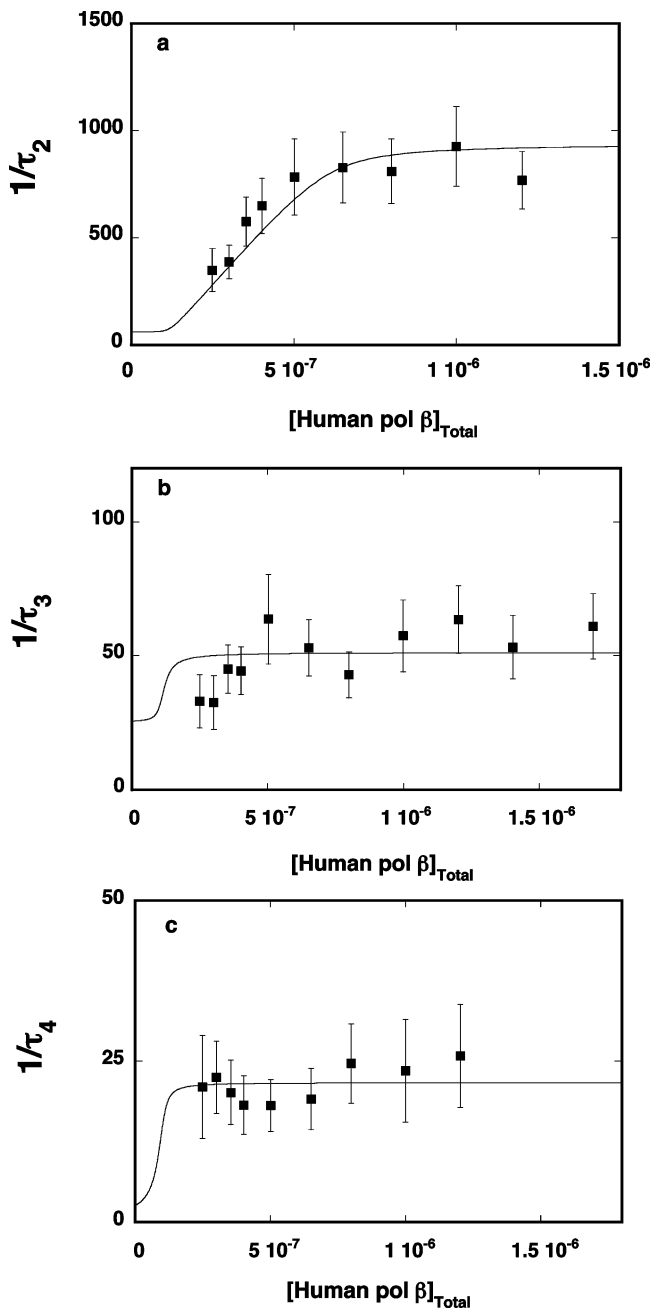


Figure 39. The dependence of the reciprocal relaxation times for the binding of human pol β to the 20-mer dT(pT)₈-pFlu-(pT)₁₀ in the (pol β)₁₆ binding mode in sodium cacodylate/HCl (pH 7.0, 10 °C) containing 50 mM NaCl upon the total concentration of the enzyme: (a) $1/\tau_2$; (b) $1/\tau_3$; (c) $1/\tau_4$. The solid lines are nonlinear least-squares fits according to the four-step sequential mechanism with the rate constants $k_1 = 1.8 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$, $k_{-1} = 40 \text{ s}^{-1}$, $k_2 = 570 \text{ s}^{-1}$, $k_{-2} = 350 \text{ s}^{-1}$, $k_3 = 50 \text{ s}^{-1}$, $k_{-3} = 15 \text{ s}^{-1}$, $k_4 = 10 \text{ s}^{-1}$, and $k_{-4} = 18 \text{ s}^{-1}$. Reprinted with permission from ref 95. Copyright 2001 American Chemical Society.

by the equilibrium fluorescence titration method (Figure 37). The macroscopic binding constant K_{20} is related to the equilibrium constants characterizing partial equilibrium steps by

$$K_{20} = K_1(1 + K_2 + K_2K_3 + K_2K_3K_4) \quad (86)$$

where $K_1 = k_1/k_{-1}$, $K_2 = k_2/k_{-2}$, $K_3 = k_3/k_{-3}$, and $K_4 = k_4/k_{-4}$. This relationship reduces the number of the independent parameters to five in the numerical fitting of the three

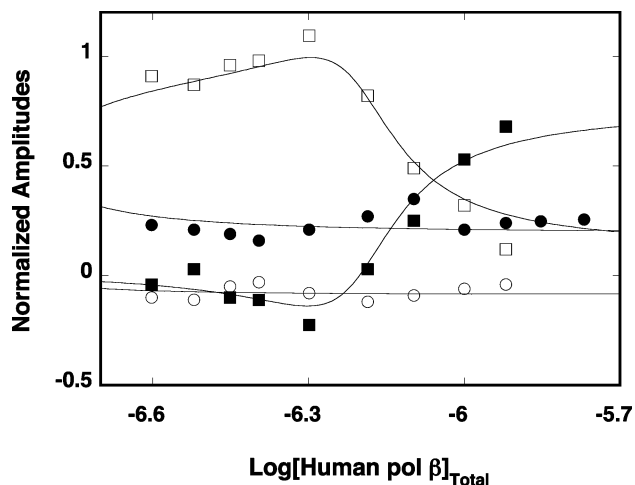


Figure 40. The dependence of the individual relaxation amplitudes, A_1 , A_2 , A_3 , and A_4 , for the binding of human pol β to the 20-mer dT(pT)₈-pFlu-(pT)₁₀ in sodium cacodylate/HCl (pH 7.0, 10 °C) containing 50 mM NaCl upon the logarithm of the total enzyme concentration: A_1 (■); A_2 (□); A_3 (●); A_4 (○). The solid lines are nonlinear least-squares fits according to the four-step sequential mechanism (see text) with the relative fluorescence intensities $F_2 = 1.344$, $F_3 = 1.456$, $F_4 = 1.531$, and $F_5 = 1.375$. The maximum nucleic acid fluorescence increase is taken from the equilibrium fluorescence titration in the same solution conditions as $\Delta F_{\text{max}} = 0.47$. The rate constants are the same as those in Figure 39. Reprinted with permission from ref 95. Copyright 2001 American Chemical Society.

individual relaxation times. Because the unresolved step is very fast, the fit is first performed with the starting value of the rate constant k_1 close to the diffusion-controlled limit, for example, $\sim 5 \times 10^9 \text{ M}^{-1}$. Subsequently, various values of k_1 and other rate constants have been tested in these analyses.

The obtained rate constants were then used in the fitting of the four individual amplitudes that include the relative molar fluorescence intensities, using a matrix projection operator technique.¹¹⁶ This determination is facilitated by the fact that the maximum, fractional increase of the nucleic acid fluorescence, $\Delta F_{\text{max}} = 0.47 \pm 0.05$, is known from the equilibrium titrations (Figure 37). Moreover, ΔF_{max} can be analytically expressed as⁹⁵

$$\Delta F_{\text{max}} = \frac{\Delta F_2}{Z} + \frac{K_2 \Delta F_3}{Z} + \frac{K_2 K_3 \Delta F_4}{Z} + \frac{K_2 K_3 K_4 \Delta F_5}{Z} \quad (87)$$

where $Z = 1 + K_2 + K_2K_3 + K_2K_3K_4$, $\Delta F_2 = (F_2 - F_1)/F_1$, $\Delta F_3 = (F_3 - F_1)/F_1$, $\Delta F_4 = (F_4 - F_1)/F_1$, and $\Delta F_5 = (F_5 - F_1)/F_1$ are fractional fluorescence intensities of each intermediate in the binding of human pol β to the ssDNA 20-mer in the (pol β)₁₆ binding mode relative to the molar fluorescence intensity of the free nucleic acid, F_1 . Expression 87 provides an additional relationship between the fluorescence parameters with the value of ΔF_{max} playing the role of a scaling factor.^{95,116} In the final step of the analysis, global fitting that simultaneously includes all relaxation times and amplitudes refines the values of the rate constants and relative molar fluorescence parameters. The solid lines in Figures 39 and 40 are theoretical plots of the relaxation times and amplitudes as a function of the total polymerase β concentration according to the above mechanism using a single set of the parameters obtained from the nonlinear least-squares fits.

Introducing the obtained values of the rate constants (Figure 40) into the partial equilibrium constants for each

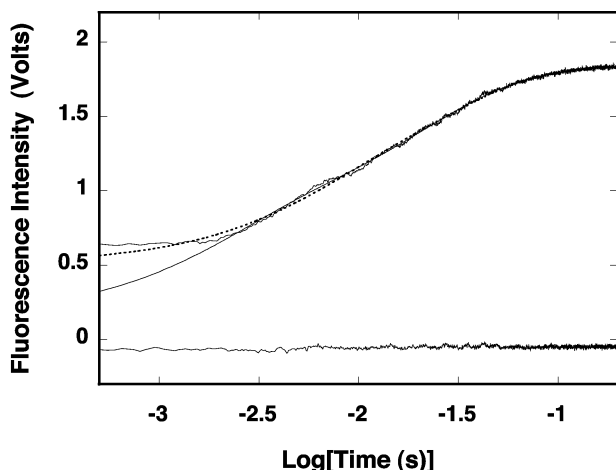


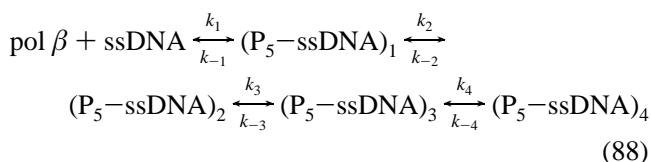
Figure 41. Fluorescence stopped-flow kinetic trace after mixing human pol β with the 10-mer, dT(pT)₃-pFlu-(pT)₅ in sodium cacodylate/HCl (pH 7.0, 10 °C), containing 50 mM NaCl ($\lambda_{\text{ex}} = 485$ nm, $\lambda_{\text{em}} > 495$ nm). The final concentrations of the polymerase and the 10-mer are 8×10^{-7} M and 1×10^{-7} M (oligomer), respectively. The experimental kinetic trace is shown in logarithmic scale with respect to time. The initial, horizontal part of the trace is the steady-state value of the fluorescence of the sample recorded for 2 ms before the flow stopped. The solid line is the three-exponential, nonlinear least-squares fit of the experimental curve. The dashed line is the nonlinear least-squares fit using the two-exponential function. The lower horizontal trace is the zero line, which is obtained after only mixing the nucleic acid with the buffer. Reprinted with permission from ref 95. Copyright 2001 American Chemical Society.

step provides $K_1 = (4.6 \pm 1.6) \times 10^7 \text{ M}^{-1}$, $K_2 = 1.6 \pm 0.5$, $K_3 = 3.3 \pm 0.7$, and $K_4 = 0.6 \pm 0.3$. Thus, the first step has a dominant contribution to the free energy, ΔG° , of the ssDNA binding. The next two steps also increase the affinity, while the fourth step has a negative contribution to ΔG° . The data indicate that the first intermediate, $(\text{P}_{16}\text{-ssDNA})_1$, is characterized by the largest relative molar fluorescence intensity ($F_2 = 1.35 \pm 0.05$) as compared to the free 20-mer. The conformational transition to $(\text{P}_{16}\text{-ssDNA})_2$ induces only an $\sim 7\%$ additional increase of the 20-mer fluorescence over F_2 ($F_3 = 1.46 \pm 0.05$), the transition to the $(\text{P}_{16}\text{-ssDNA})_3$ is accompanied by an additional increase of the ssDNA fluorescence by $\sim 7\%$ as compared to F_3 ($F_4 = 1.53 \pm 0.05$), but transition to the $(\text{P}_{16}\text{-ssDNA})_4$ is accompanied by a nucleic acid fluorescence decrease as compared to F_4 ($F_5 = 1.38 \pm 0.05$).

3.5.4.2. Kinetics of Human Pol β Binding to the ssDNA 10-mer in the $(\text{pol } \beta)_5$ Binding Mode. In the $(\text{pol } \beta)_5$ binding mode, the polymerase associates with only 5 ± 2 nucleotides, exclusively engaging its 8-kDa domain in the interactions with the DNA.^{43,44} The kinetics of human pol β binding to the ssDNA in the $(\text{pol } \beta)_5$ binding mode has been studied with the ssDNA 10-mer dT(pT)₃-pFlu-(pT)₅. Thermodynamic studies showed that with the ssDNA oligomer 10 nucleotides long, the enzyme can only form the $(\text{pol } \beta)_5$ binding mode.^{43,44} The stopped-flow kinetic trace of the dT(pT)₃-pFlu-(pT)₅ fluorescence after mixing 1×10^{-7} M (oligomer) 10-mer with 8×10^{-7} M human pol β (final concentrations) is shown in Figure 41. Also, the trace corresponding to the nucleic acid alone (zero line) at the same concentration as that used with the protein but mixed only with the buffer is included. The solid line is the three-exponential fit of the trace. The two-exponential fit (dashed line) does not provide an adequate description of the

experimental trace. As observed in the case of the formation of the $(\text{pol } \beta)_{16}$ binding mode, although the three-exponential fit provides an excellent description of the observed kinetics, it does not account for the entire fluorescence increase that results from the complex formation. The data indicate that there is a fast step preceding the observed trace characterized by a relaxation time too short to be extracted in the stopped-flow experiment. In other words, the formation of the $(\text{pol } \beta)_5$ binding mode must also include at least four steps.

The reciprocal relaxation times, $1/\tau_2$, $1/\tau_3$, and $1/\tau_4$, characterizing the resolvable three relaxation processes as a function of the total human pol β concentration are shown in Figure 42. All three relaxation times show little dependence upon enzyme concentration, indicating that they characterize isomerizations of the formed complex (Figure 35).¹¹⁶ The simplest mechanism of the formation of the $(\text{pol } \beta)_5$ binding mode that can account for such behavior of the relaxation times is a sequential four-step reaction in which the bimolecular association is followed by three conformational transitions of the formed complex, analogous to eq 84, as



Similar to the kinetics of the $(\text{pol } \beta)_{16}$ binding mode formation, the relaxation time τ_1 is not available from the experiment; however, the amplitude, A_1 , of this mode can be obtained from the known amplitudes of the second, third, and fourth normal modes, A_2 , A_3 , and A_4 , and the total amplitude of the reaction, as defined by eq 85. The dependence of the individual amplitudes, A_1 , A_2 , A_3 , and A_4 , of the four normal modes of the reaction upon the human pol β concentration is shown in Figure 43. The individual amplitudes are expressed as fractions of the total amplitude. With the exception of A_1 , all individual amplitudes significantly contribute to the total amplitude over the entire examined protein concentration range. While A_2 , A_3 , and A_4 are positive, A_1 assumes negative values at low protein concentration and rises to positive values in the high enzyme concentration range. Such behavior is in full agreement with the proposed mechanism (eq 88) and already indicates that the first intermediate has relative molar fluorescence intensity significantly lower than the corresponding parameters characterizing other intermediates (Figures 35 and 36).

The same strategy, relying on the simultaneous analyses of the three relaxation times and the four amplitudes, as discussed above for the $(\text{pol } \beta)_{16}$ binding mode formation has been used to determine all rate constants of the particular steps and molar fluorescence parameters characterizing each intermediate in the $(\text{pol } \beta)_5$ binding mode formation. The macroscopic binding constant, $K_{10} = (9.4 \pm 1.5) \times 10^6 \text{ M}^{-1}$, and the maximum fractional increase of the nucleic acid fluorescence, $\Delta F_{\text{max}} = 0.22 \pm 0.02$, for the human pol β molecule binding to dT(pT)₃-Flu-(pT)₅ in the $(\text{pol } \beta)_5$ binding mode have been independently obtained in the same solution conditions by the equilibrium fluorescence titration method.⁹⁵ Both quantities are analytically related to the partial equilibrium constants and fractional fluorescence changes of each intermediate by expressions analogous to eqs 84 and 85, respectively. The solid lines in Figures 42 and 43 are nonlinear least-squares fits of the relaxation times and

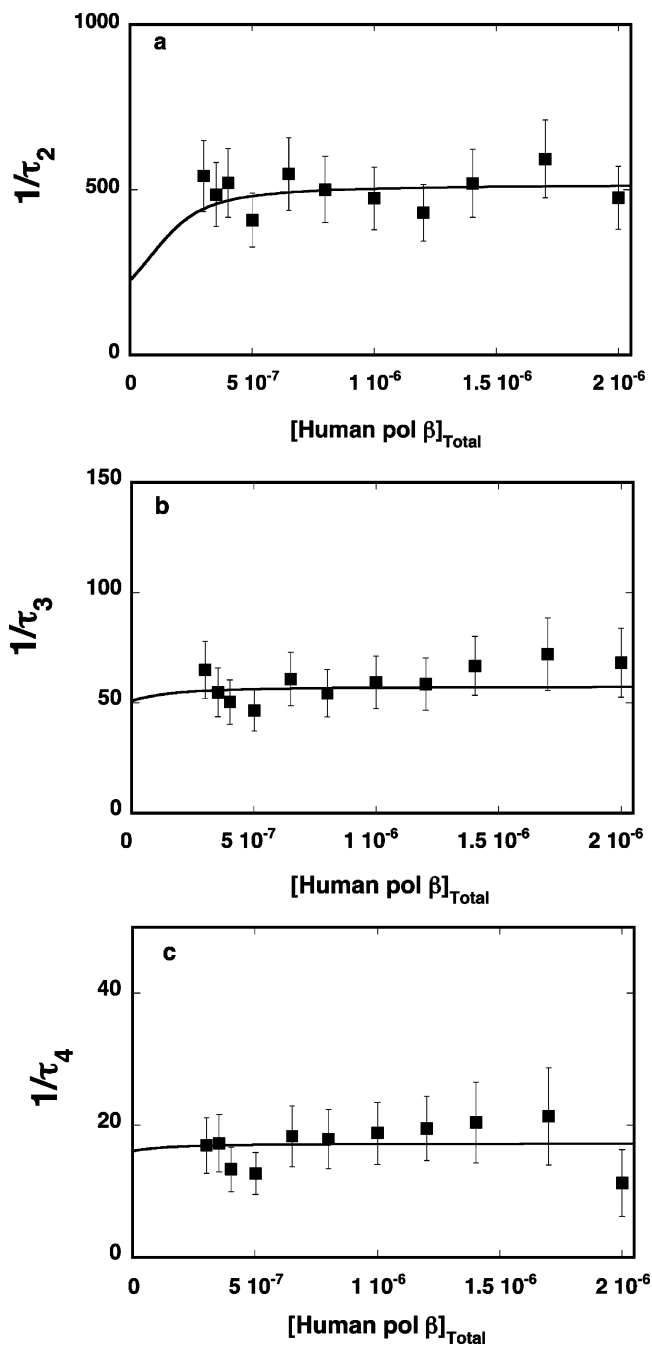


Figure 42. The dependence of the reciprocal relaxation times for the binding of human pol β to the 10-mer dT(pT)₃-pFlu-(pT)₅ in sodium cacodylate/HCl (pH 7.0, 10 °C) containing 50 mM NaCl upon the total concentration of the enzyme: (a) $1/\tau_2$; (b) $1/\tau_3$; (c) $1/\tau_4$. The solid lines are nonlinear least-squares fits according to the four-step sequential mechanism with the rate constants $k_1 = 2 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$, $k_{-1} = 300 \text{ s}^{-1}$, $k_2 = 90 \text{ s}^{-1}$, $k_{-2} = 400 \text{ s}^{-1}$, $k_3 = 30 \text{ s}^{-1}$, $k_{-3} = 48 \text{ s}^{-1}$, $k_4 = 5 \text{ s}^{-1}$, and $k_{-4} = 19 \text{ s}^{-1}$. Reprinted with permission from ref 95. Copyright 2001 American Chemical Society.

amplitudes according to the mechanism defined by eq 88 using a single set of parameters.

The bimolecular step in the formation of the (pol β)₅ binding mode is similar to the analogous step determined for the (pol β)₁₆ binding mode; however, the dissociation rate constant k_{-1} is larger by a factor of ~ 8 than the corresponding value obtained for the large site size binding mode (Figure 43). On the other hand, transition to the second intermediate, (P₅-ssDNA)₂, is characterized by a signifi-

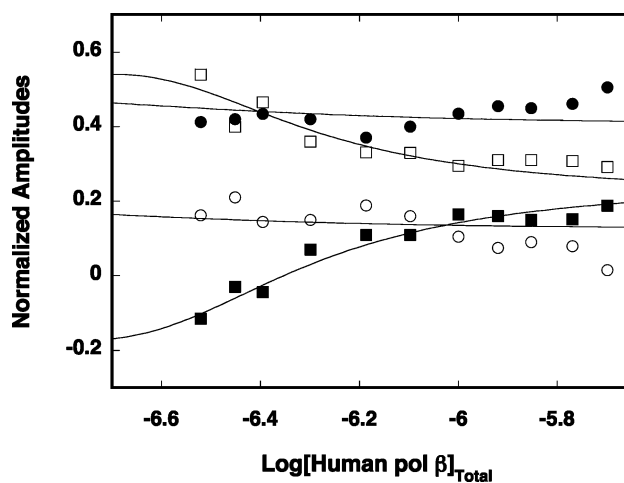


Figure 43. The dependence of the individual relaxation amplitudes, A_1 , A_2 , A_3 , and A_4 , for the binding of human pol β to the 10-mer dT(pT)₃-pFlu-(pT)₅ in sodium cacodylate/HCl (pH 7.0, 10 °C) containing 50 mM NaCl upon the logarithm of the total concentration of the enzyme: A_1 (■); A_2 (□); A_3 (●); A_4 (○). The solid lines are nonlinear least-squares fits according to the four-step sequential mechanism with the relative molar fluorescence intensities $F_2 = 1.024$, $F_3 = 1.17$, $F_4 = 1.482$, and $F_5 = 1.22$. The maximum fluorescence increase of the nucleic acid is taken from the equilibrium fluorescence titration in the same solution conditions as $\Delta F_{\text{max}} = 0.22$. The rate constants are the same as those in Figure 41. Reprinted with permission from ref 95. Copyright 2001 American Chemical Society.

cantly lower (factor of ~ 6) rate constant than the analogous step in the formation of the (pol β)₁₆ binding mode (Figure 42). Introducing the values of the rate constants to the equilibrium constants for each step provides $K_1 = (6.7 \pm 2) \times 10^6 \text{ M}^{-1}$, $K_2 = 0.23 \pm 0.09$, $K_3 = 0.63 \pm 0.3$, and $K_4 = 0.26 \pm 0.1$. It is evident that, similar to the (pol β)₁₆ binding mode, the first step has a predominant contribution to the free energy of the ssDNA binding, ΔG° , in the formation of the (pol β)₅ binding mode. However, contrary to the results obtained with the ssDNA 20-mer, where only the last step contributes negatively to ΔG° , all remaining steps in the formation of the (pol β)₅ binding mode contribute negatively to the total free energy of binding.⁹⁵

There are also significant differences between the two ssDNA-binding modes in the structures of the intermediates, as indicated by their relative molar fluorescence intensities. The molar fluorescence intensity of the first intermediate in the (pol β)₅ binding mode is only $\sim 3\%$ larger than the fluorescence of the free nucleic acid, as compared to the $\sim 35\%$ difference in the case of the (pol β)₁₆ binding mode, indicating that a much less pronounced conformational nucleic acid change accompanies the formation of the first intermediate, (P₅-ssDNA)₁, in the (pol β)₅ binding mode. Significant conformational changes occur in the transition to (P₅-ssDNA)₂ and particularly to (P₅-ssDNA)₃ which is characterized by the relative molar fluorescence intensity $F_4 = 1.48 \pm 0.05$, comparable to the intermediate (P₁₆-ssDNA)₃ in the (pol β)₁₆ binding mode.⁹⁵

3.5.4.3. Some Mechanistic Implications for pol β Binding to the ssDNA. The Binding-Initiating Role of the 8-kDa Domain and the Complex-Stabilizing Role of the 31-kDa Domain. The formation of the (pol β)₅ binding mode is described by the same sequential mechanism as that observed for the (pol β)₁₆ binding mode (eqs 84 and 88). The fact that in the (pol β)₅ binding mode exclusively the 8-kDa domain is involved in interactions with the nucleic

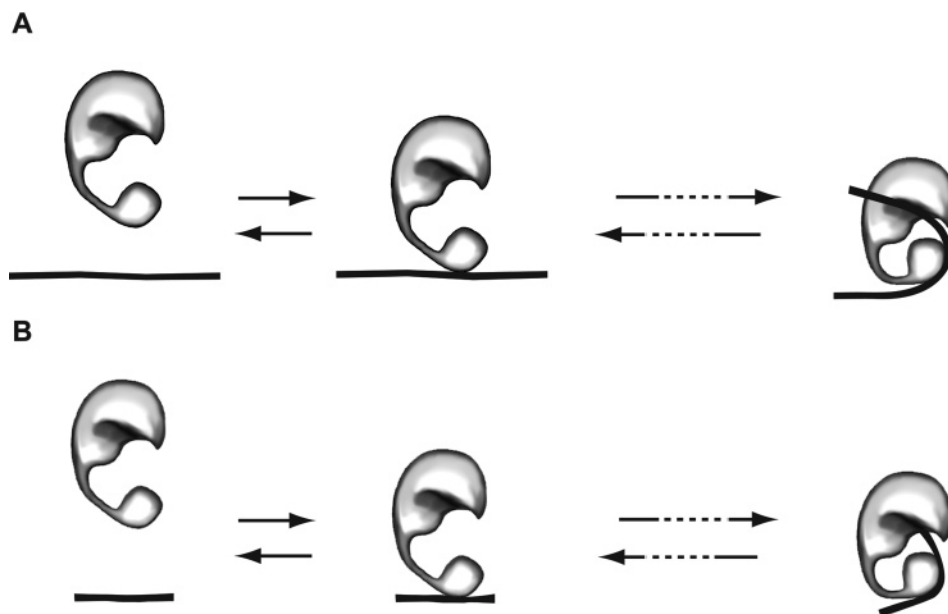


Figure 44. Schematic models of human pol β binding to the ssDNA in the $(\text{pol } \beta)_{16}$ and $(\text{pol } \beta)_5$ binding modes based on the thermodynamic and kinetic data. In both binding modes, the enzyme initially binds the nucleic acid through the small 8-kDa domain. Several conformational transitions of the nucleic acid–enzyme complex are induced at the interface of the 8-kDa domain and the nucleic acid. In the case of the $(\text{pol } \beta)_{16}$ binding mode, that is, when the ssDNA is long enough to engage the total DNA-binding site of the polymerase, the formed intermediates are additionally stabilized through interactions with the DNA-binding subsite located on the large 31-kDa domain of the enzyme, resulting in the final equilibrium complex with both DNA-binding subsites involved in interactions with the nucleic acid (A). In the case of the short ssDNA oligomer, the stability of all conformational states and the final equilibrium complex is based solely on the free energy generated through interactions with the 8-kDa domain, without the involvement of the DNA-binding subsite of the 31-kDa domain (B). Reprinted with permission from ref 95. Copyright 2001 American Chemical Society.

acid and the formation of both binding modes includes kinetically similar steps provides the first indication that the interactions at the interface between the 8-kDa domain and the nucleic acid play a dominant role in transitions between the intermediates in both binding modes (see below). However, although the general kinetic mechanism is the same, there are significant differences in the nature of the formed intermediates between the two binding modes, indicating specific roles played by the DNA-binding subsites located on the 8-kDa and 31-kDa domains.

The bimolecular rate constants are very similar in both binding modes (Figures 39 and 42). On the other hand, the dissociation rate constant k_{-1} of the $(\text{pol } \beta)_5$ binding mode is higher by a factor of ~ 8 than the value obtained in the $(\text{pol } \beta)_{16}$ binding mode. These data indicate that there is a much higher probability of human pol β being released back into the solution from the $(\text{P}_5\text{-ssDNA})_1$ intermediate in the $(\text{pol } \beta)_5$ binding mode than from the $(\text{P}_{16}\text{-ssDNA})_1$ in the $(\text{pol } \beta)_{16}$ binding mode. The very similar values of k_1 in both binding modes and a much higher value of the dissociation rate constant k_{-1} in the $(\text{pol } \beta)_5$ binding mode result in an ~ 1 order of magnitude lower partial equilibrium constant K_1 characterizing the formation of the $(\text{P}_5\text{-ssDNA})_1$ as compared to the formation of the $(\text{P}_{16}\text{-ssDNA})_1$.⁹⁵ Also, the amplitude analysis indicates that the transition to the first intermediate $(\text{P}_5\text{-ssDNA})_1$ is not associated with a dominant fluorescence change as observed in the case of $(\text{P}_{16}\text{-ssDNA})_1$. Thus, contrary to the $(\text{pol } \beta)_{16}$ binding mode, the formation of $(\text{P}_5\text{-ssDNA})_1$ is not associated with a large nucleic acid conformational change. Recall that equilibrium studies have shown that in the $(\text{pol } \beta)_{16}$ binding mode both the 8-kDa and 31 kDa domains are engaged in interactions with the ssDNA, while only the 8-kDa domain interacts with the nucleic acid in the $(\text{pol } \beta)_5$ binding mode.^{43,44} The

simplest explanation of the kinetic behavior is that in both binding modes, the association reaction (k_1) occurs through the DNA-binding subsite of the 8-kDa domain of the protein. However, in the $(\text{pol } \beta)_{16}$ binding mode, fast engagement of the DNA-binding subsite on the 31-kDa domain in interactions with the DNA leads to a lower value of k_{-1} of the $(\text{P}_{16}\text{-ssDNA})_1$ and to more pronounced conformational changes of the nucleic acid. In the $(\text{pol } \beta)_5$ binding mode only the 8-kDa domain is involved in the formation of the $(\text{P}_5\text{-ssDNA})_1$ without the additional engagement of the 31 kDa domain.

The energetics of the subsequent partial reactions in both binding modes supports this conclusion and provides further mechanistic insight on the nature of the observed intermediates. The first binding step generates a predominant part of the total free energy ΔG° of binding in both binding modes (see above). However, the much higher value of K_1 characterizing the formation of the $(\text{P}_{16}\text{-ssDNA})_1$ as compared to the $(\text{P}_5\text{-ssDNA})_1$ strongly indicates that additional interacting areas are involved in the formation of the $(\text{P}_{16}\text{-ssDNA})_1$. Moreover, with the exception of $(\text{P}_{16}\text{-ssDNA})_4$, the transitions to $(\text{P}_{16}\text{-ssDNA})_2$ and $(\text{P}_{16}\text{-ssDNA})_3$ are accompanied by additional favorable ΔG° changes. On the other hand, all subsequent transitions to $(\text{P}_5\text{-ssDNA})_2$, $(\text{P}_5\text{-ssDNA})_3$, and $(\text{P}_5\text{-ssDNA})_4$ are characterized by the negative contribution to the total free energy of binding. With the ssDNA 10-mer, human pol β exclusively interacts using its 8-kDa domain, that is, it can only form the $(\text{pol } \beta)_5$ binding mode. The obtained results indicate that in the case of this shorter ssDNA oligomer, the polymerase cannot engage in a stable complex beyond the first intermediate. In the $(\text{pol } \beta)_{16}$ binding mode, formed with the ssDNA 20-mer, the nucleic acid is long enough to engage in interactions with the entire total DNA-binding site of the polymerase including the

DNA-binding subsite located on the 31-kDa domain. Thus, the data indicate that the favorable free energy changes in transitions between the intermediates of the $(\text{pol } \beta)_{16}$ binding mode result from additional interactions between the DNA and the DNA-binding subsite located on the large 31-kDa domain.⁹⁵

Notice that these intermediates are also present in the case of the $(\text{pol } \beta)_5$ binding mode formed with the ssDNA 10-mer, indicating that they are, in fact, induced by the interactions between the 8-kDa domain and the nucleic acid, although the 10-mer is not long enough to stabilize them through interactions with the DNA-binding subsite of the 31-kDa domain. In other words, the data suggest that the favorable energy changes in the partial steps of the $(\text{pol } \beta)_{16}$ binding mode formation reflect efficient docking of the nucleic acid in both DNA-binding subsites of the total DNA-binding site of human $\text{pol } \beta$, following the initial association through the 8-kDa domain. In this context, the observed different stability of the $(\text{pol } \beta)_5$ binding mode intermediates would reflect a lack of the extra interaction areas in the DNA-binding subsite of the 31-kDa domain.

Schematic models of the human $\text{pol } \beta$ binding to the ssDNA in the $(\text{pol } \beta)_{16}$ and $(\text{pol } \beta)_5$ binding modes based on thermodynamic and kinetic data are shown in Figure 44a,b. The initial association of the enzyme with the nucleic acid in both binding modes occurs through the small 8-kDa domain. Interactions between the DNA and the polymerase at the interface of the 8-kDa domain induce conformational transitions of the nucleic acid–enzyme complex. In the case of the short ssDNA oligomer, the stability of these conformational states (intermediates) is solely based on the free energy generated through interactions with the 8-kDa domain (Figure 44b). However, in the case of the $(\text{pol } \beta)_{16}$ binding mode, that is, when the ssDNA is long enough to engage the total DNA-binding site of the polymerase, the formed intermediates are additionally stabilized through interactions with the DNA-binding subsite located on the large 31-kDa domain of the enzyme (Figure 44a).

4. Summary

Understanding ligand–macromolecule interactions, such as those involving proteins and nucleic acids, requires detailed knowledge of the energetics and kinetics of the formed complexes. Spectroscopic methods are widely used in characterizing the energetics (thermodynamics) and kinetics of protein–nucleic acid and, in general, ligand–macromolecule interactions in solution. These methods do not require large quantities of material, are very convenient to use, and, most importantly, *do not perturb the studied processes*. However, spectroscopic methods are indirect, that is, the interactions are measured through monitoring changes of some physicochemical parameter accompanying the formation of studied complexes. Therefore, in such analyses, it is absolutely necessary to determine the relationship between the observed signal and the degree of binding to obtain thermodynamically and kinetically meaningful interaction and spectroscopic parameters. The methods discussed in this review describe general quantitative approaches of the analyses of macromolecular binding and kinetics through spectroscopic measurements, which allow an experimenter to quantitatively determine the degree of binding or the degree of macromolecule saturation and the free ligand concentration and examine spectroscopic properties of the involved intermediates of the reaction. In other words, they

enable the experimenter to construct the *thermodynamically rigorous* binding isotherms to examine structural changes in involved kinetic intermediates. Only when the thermodynamically rigorous isotherm is obtained can it be analyzed by using the thermodynamic models, which incorporate the known molecular aspects of the ligand–macromolecule interactions, cooperativity, allosteric conformational changes, overlap of potential binding sites, etc., completely independent from the relationship with the spectroscopic signal used to monitor the interactions. In turn, such rigorous thermodynamic data are invaluable in the quantitative analyses of the kinetics of the studied interactions and resulting partial equilibria among the kinetic intermediates.

Spectroscopic stopped-flow measurements provide two independent sets of data, the relaxation times and amplitudes characterizing the normal modes of the reaction. Both sets of data contain information about the mechanism and rate constants of the individual kinetic steps. Quantitative studies of the kinetics of ligand–macromolecule interactions require the determination and analyses of both the relaxation times and the amplitudes. Examination of the amplitudes offers additional information about the structure and nature of intermediates, unavailable by any other method. The analysis is greatly facilitated if relaxation times and amplitudes of all normal modes of the reaction are available from the experiment. Nevertheless, it is common that in stopped-flow experiments the relaxation time characterizing the bimolecular step is too fast to be experimentally accessible. However, in favorable cases, where all amplitudes are available over a large range of the ligand or macromolecule concentrations, the rate constants characterizing the bimolecular step can still be estimated with adequate accuracy from the amplitude analysis. The matrix projection operator method allows the experimenter to avoid the cumbersome numerical analysis of the eigenvectors of the coefficient matrix, usually necessary in amplitude analysis, by turning the eigenvector problem into a simpler algebraic calculation of the projection operators.

As we pointed out above, our discussion mainly focused on the fundamental problem of obtaining thermodynamic, kinetic, and spectroscopic parameters free of assumptions about the relationship between the observed signal and the degree of protein or nucleic acid saturation. *It is self-understood that any meaningful conclusions about the molecular interpretation of the behavior of any interacting system are absolutely dependent upon the quantitative determination of the thermodynamic, kinetic, and spectroscopic parameters on which such conclusions are based.* The methods have been discussed as applied to studying protein–nucleic acid interactions using the fluorescence intensity to monitor the binding. The discussed various interacting systems are very different and provided an opportunity to address several specific aspects of the protein–nucleic acid interactions frequently encountered in research practice. However, these approaches can generally be applied to any ligand–macromolecule system, examined using any spectroscopic signal originating from a ligand or a macromolecule.

5. Acknowledgment

We wish to thank Dr. Maria J. Jezewska for discussions of most of the results presented here, her exceptional experimental and analytical skills, and intellectual input that made these works possible. We wish to thank Betty Sordahl

for reading the manuscript. This work was supported by NIH Grants GM-58565 and GM46679.

6. References

- (1) Jen-Jacobson, L. *Biopolymers* **1997**, *44*, 153.
- (2) Record, M. T., Jr.; Anderson, C. F.; Lohman, T. M. *Q. Rev. Biophys.* **1978**, *11*, 103.
- (3) Eftink, M. *Methods Enzymol.* **1997**, *278*, 221.
- (4) Lynch, T. W.; Kosztin, D.; McLean, M. A.; Schulten, K.; Sligar, S. G. *Biophys. J.* **2002**, *82*, 93.
- (5) McAfee, J. G.; Edmondson, S. P.; Zegar, I.; Shriver, J. W. *Biochemistry* **1996**, *35*, 4034.
- (6) Lohman, T. M.; Bujalowski, W. *Methods Enzymol.* **1991**, *208*, 258.
- (7) Hill, T. L. *Cooperativity Theory in Biochemistry*; Springer-Verlag: New York, 1985; Chapter 6.
- (8) Cantor, R. C.; Schimmel, P. R. *Biophysical Chemistry*; W. H. Freeman: New York, 1980; Vol. III, Chapter 15.
- (9) Bujalowski, W.; Lohman, T. M. *Biochemistry* **1987**, *26*, 3099.
- (10) Bujalowski, W.; Klonowska, M. M. *Biochemistry* **1993**, *32*, 5888.
- (11) McGhee, J. D.; von Hippel, P. H. *J. Mol. Biol.* **1974**, *86*, 469.
- (12) Epstein, I. R. *Biophys. Chem.* **1978**, *8*, 327.
- (13) Menetski, J. P.; Kowalczykowski, S. C. *J. Mol. Biol.* **1985**, *181*, 281.
- (14) Kowalczykowski, S. C.; Paul, L. S.; Lonberg, N.; Newport, J. W.; McSwiggen, J. A.; von Hippel, P. H. *Biochemistry* **1986**, *25*, 1226.
- (15) Arosio, D.; Costantini, S.; Kong, Y.; Vindigni, A. *J. Biol. Chem.* **2004**, *279*, 42826.
- (16) Di Cera, E.; Kong, Y. *Biophys. Chem.* **1996**, *61*, 107.
- (17) Jensen, D. E.; von Hippel, P. H. *Anal. Biochem.* **1977**, *80*, 267.
- (18) Draper, D. E.; von Hippel, P. H. *J. Mol. Biol.* **1978**, *122*, 321.
- (19) Garner, M. M.; Revzin, A. *Nucleic Acids Res.* **1981**, *9*, 3047.
- (20) Mou, T.-C.; Gray, C. W.; Gray, D. M. *Biophys. J.* **1999**, *76*, 1535.
- (21) Mou, T.-C.; Gray, C. W.; Terwilliger, T. C.; Gray, D. M. *Biochemistry* **2001**, *40*, 2267.
- (22) Holbrook, J. A.; Tsodikov, O. V.; Saecker, R. M.; Record, M. T. *J. Mol. Biol.* **2001**, *310*, 379.
- (23) Garcia-Garcia, C.; Draper, D. E. *J. Mol. Biol.* **2003**, *331*, 75.
- (24) Boschelli, F. J. *Mol. Biol.* **1982**, *162*, 267.
- (25) Liu, D.; Prasad, R.; Wilson, S. H.; DeRose, F.; Mullen, G. P. *Biochemistry* **1996**, *35*, 6188.
- (26) Heyduk, T.; Lee, J. C. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 1744.
- (27) Porschke, D.; Rauh, H. *Biochemistry* **1983**, *22*, 4737.
- (28) Ferrari, M. E.; Bujalowski, W.; Lohman, T. M. *J. Mol. Biol.* **1994**, *236*, 106.
- (29) Witting, P.; Norden, B.; Kim, S. K.; Takahashi, M. *J. Biol. Chem.* **1994**, *269*, 5700.
- (30) Bujalowski, W.; Lohman, T. M. *J. Mol. Biol.* **1989**, *207*, 249.
- (31) Bujalowski, W.; Lohman, T. M. *J. Mol. Biol.* **1989**, *207*, 268.
- (32) Dou, S.-X.; Wang, P.-Y.; Xu, H. Q.; Xi, X. G. *J. Biol. Chem.* **2004**, *279*, 6354.
- (33) Schwarz, G.; Watanabe, F. *J. Mol. Biol.* **1983**, *163*, 467.
- (34) Watanabe, F.; Schwarz, G. *J. Mol. Biol.* **1983**, *163*, 485.
- (35) Bujalowski, W.; Jezewska, M. J. *Biochemistry* **1995**, *34*, 8513.
- (36) Jezewska, M. J.; Kim, U.-S.; Bujalowski, W. *Biochemistry* **1996**, *35*, 2129.
- (37) Jezewska, M. J.; Kim, U.-S.; Bujalowski, W. *Biophys. J.* **1996**, *71*, 2075.
- (38) Jezewska, M. J.; Rajendran, S.; Bujalowski, W. *Biochemistry* **1997**, *36*, 10320.
- (39) Jezewska, M. J.; Rajendran, S.; Bujalowski, W. *Biochemistry* **1998**, *37*, 3116.
- (40) Galletto, R.; Bujalowski, W. *Biochemistry* **2002**, *41*, 8921.
- (41) Jezewska, M. J.; Bujalowski, W. *Biophys. Chem.* **1997**, *64*, 253.
- (42) Bujalowski, W.; Jezewska, M. J. In *Spectrophotometry and Spectrofluorimetry. A Practical Approach*; Gore, M. G., Ed.; Practical Approach Series 325; Oxford University Press: Oxford, U.K., 2000; Chapter 5.
- (43) Jezewska, M. J.; Rajendran, S.; Bujalowski, W. *J. Mol. Biol.* **1998**, *284*, 1113.
- (44) Rajendran, S.; Jezewska, M. J.; Bujalowski, W. *J. Biol. Chem.* **1998**, *273*, 31021.
- (45) Rajendran, S.; Jezewska, M. J.; Bujalowski, W. *J. Mol. Biol.* **2001**, *308*, 477.
- (46) Jezewska, M. J.; Rajendran, S.; Bujalowski, W. *J. Biol. Chem.* **2001**, *276*, 1623.
- (47) Jezewska, M. J.; Rajendran, S.; Bujalowski, W. *J. Biol. Chem.* **2000**, *275*, 27865.
- (48) Jezewska, M. J.; Bujalowski, W. *Biochemistry* **2000**, *39*, 10454.
- (49) Jezewska, M. J.; Galletto, R.; Bujalowski, W. *Biochemistry* **2003**, *42*, 5955.
- (50) Woodbury, C. P.; von Hippel, P. H. *Biochemistry* **1983**, *22*, 4730.
- (51) Clore, G. M.; Gronenborn, A. M.; Davies, R. W. *J. Mol. Biol.* **1982**, *155*, 447.
- (52) Wong, I.; Lohman, T. M.; *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5428.
- (53) Wong, I.; Lohman, T. M. *Science* **1992**, *256*, 350.
- (54) Kranz, J. K.; Hall, K. B. *J. Mol. Biol.* **1998**, *275*, 465.
- (55) Luzetti, S. L.; Voloshin, O. N.; Inman, R. B.; Camerini-Otero, D.; Cox, M. M. *J. Biol. Chem.* **2004**, *279*, 30037.
- (56) Schubert, F.; Zettl, H.; Hafner, W.; Krauss, G.; Krausch, G. *Biochemistry* **2003**, *42*, 10288.
- (57) Fried, M.; Crothers, D. M. *Nucleic Acids Res.* **1981**, *9*, 6505.
- (58) Lim, W. A.; Sauer, R. T.; Lander, A. D. *Methods Enzymol.* **1991**, *208*, 196.
- (59) Garner, M. M.; Revzin, A. *Nucleic Acids Res.* **1981**, *9*, 3047.
- (60) Draper, D. E.; von Hippel, P. H. *Biochemistry* **1977**, *18*, 753.
- (61) Wu, J.; Parhurst, K. M.; Powell, R. M.; Brenowitz, M.; Parhurst, L. *J. Biol. Chem.* **2001**, *276*, 14614.
- (62) Villemain, J. L.; Giedroc, D. P. *Biochemistry* **1996**, *35*, 14395.
- (63) Ando, R. A.; Morrical, S. W. *J. Mol. Biol.* **1998**, *283*, 785.
- (64) Dhavan, G. M.; Crothers, D. M.; Chance, M. R.; Brenowitz, M. *J. Mol. Biol.* **2002**, *315*, 1027.
- (65) Clerte, C.; Hall, K. B. *Biochemistry* **2004**, *43*, 13404.
- (66) Showalter, S. A.; Hall, K. B. *J. Mol. Biol.* **2004**, *335*, 465.
- (67) Bailey, M. F.; van der Schans, E. J.; Millar, D. P. *J. Mol. Biol.* **2004**, *336*, 673.
- (68) Purohit, V.; Grindley, N. D. F.; Joyce, C. M. *Biochemistry* **2003**, *42*, 10200.
- (69) VanScyoc, W. S.; Sorensen, B. R.; Rusinova, E.; Laws, W. R.; Ross, J. B. A.; Shea, M. A. *Biophys. J.* **2002**, *83*, 2767.
- (70) Lohman, T. M.; Overman, L. B. *J. Biol. Chem.* **1985**, *260*, 3594.
- (71) Halfman, C. J.; Nishida, T. *Biochemistry* **1972**, *11*, 3493.
- (72) LeBowitz, J. H.; McMacken, R. *J. Biol. Chem.* **1986**, *261*, 4738.
- (73) Baker, T. A.; Funnell, B. E.; Kornberg, A. *J. Biol. Chem.* **1987**, *262*, 6877.
- (74) Johnson, S. K.; Bhattacharyya, S.; Grieb, M. A. *Biochemistry* **2000**, *39*, 736.
- (75) Nakayama, N.; Arai, N.; Kaziro, Y.; Arai, K. *J. Biol. Chem.* **1984**, *259*, 88.
- (76) San Martin, M. C.; Stamford, N. P. J.; Dammerova, N.; Dixon, N. E.; Carazo, J. M. *J. Struct. Biol.* **1995**, *114*, 167.
- (77) Bujalowski, W.; Klonowska, M. M.; Jezewska, M. J. *J. Biol. Chem.* **1994**, *269*, 31350.
- (78) Jezewska, M. J.; Bujalowski, W. *J. Biol. Chem.* **1996**, *271*, 4261.
- (79) Yu, X.; Jezewska, M. J.; Bujalowski, W.; Egelman, E. H. *J. Mol. Biol.* **1996**, *259*, 7.
- (80) Yang, S.; Yu, X.; VanLoock, M. S.; Jezewska, M. J.; Bujalowski, W.; Egelman, E. H. *J. Mol. Biol.* **2002**, *321*, 839.
- (81) Jezewska, M. J.; Rajendran, S.; Bujalowski, W. *J. Biol. Chem.* **1998**, *273*, 9058.
- (82) Jezewska, M. J.; Rajendran, S.; Bujalowska, D.; Bujalowski, W. *J. Biol. Chem.* **1998**, *273*, 10515.
- (83) Jezewska, M. J.; Galletto, R.; Bujalowski, W. *Biochemistry* **2003**, *42*, 5955.
- (84) Pelletier, H.; Sawaya, M. R.; Kumar, A.; Wilson, S. H.; Kraut, J. *Science* **1994**, *264*, 1891.
- (85) Wiebauer, K.; Jiricny, J. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 5842.
- (86) Matsumoto, Y.; Bogenhagen, D. F. *Mol. Cell. Biol.* **1989**, *9*, 3750.
- (87) Matsumoto, Y.; Bogenhagen, D. F. *Mol. Cell. Biol.* **1991**, *11*, 4441.
- (88) Bujalowski, W.; Lohman, T. M.; Anderson, C. F. *Biopolymers* **1989**, *28*, 1637.
- (89) Lohman, T. M.; Overman, L. B.; Ferrari, M. E.; Kozlov, A. G. *Biochemistry* **1996**, *35*, 5272.
- (90) Lifson, S. *J. Chem. Phys.* **1964**, *40*, 3705.
- (91) Bradley, D. F.; Lifson, S. *Molecular Associations in Biology*; Pullman, B., Ed.; Academic Press: New York, 1968; p 261.
- (92) Schellman, J. A. *Molecular Structure and Dynamics*; Balaban, M., Ed.; International Science Services: Philadelphia, PA, and Jerusalem, 1980; p 245.
- (93) Jezewska, M. J.; Galletto, R.; Bujalowski, W. *Biochemistry* **2003**, *42*, 11864.
- (94) Jezewska, M. J.; Rajendran, S.; Galletto, R.; Bujalowski, W. *J. Mol. Biol.* **2001**, *313*, 977.
- (95) Jezewska, M. J.; Galletto, R.; Bujalowski, W. *Biochemistry* **2001**, *40*, 11794.
- (96) Jezewska, M. J.; Galletto, R.; Bujalowski, W. *Cell Biochem. Biophys.* **2003**, *38*, 125.
- (97) Jezewska, M. J.; Galletto, R.; Bujalowski, W. *J. Biol. Chem.* **2002**, *277*, 20316.
- (98) Garcia-Diaz, M.; Dominguez, O.; Lopez-Fernandez, L. A.; de Lera, L. T.; Saniger, M. L.; Ruiz, J. F.; Parraga, M.; Garcia-Ortiz, M. J.; Kirchhoff, T.; del Mazo, J.; Bernad, A.; Blanco, L. *J. Mol. Biol.* **2000**, *301*, 851.

- (99) Dominguez, O.; Ruiz, J. F.; Lain de Lera, T.; Garcia-Diaz, M.; Gonzalez, M. A.; Kirchoff, T.; Martinez, A. C.; Bernad, A.; Blanco, L. *EMBO J.* **2000**, *19*, 1731.
- (100) Showalter, A. K.; Byeon, I.-J.; Su, M.-I.; Tsai, M.-D. *Nat. Struct. Biol.* **2001**, *8*, 942.
- (101) Jezewska, M. J.; Bujalowski, W. *Biochemistry* **1996**, *35*, 2117.
- (102) Secrist, J. A.; Bario, J. R.; Leonard, N. J.; Weber, G. *Biochemistry* **1972**, *11*, 3499.
- (103) Leonard, N. J. *Crit. Rev. Biochem.* **1984**, *15*, 125.
- (104) Ando, T.; Asai, H. *J. Biochem.* **1980**, *88*, 255.
- (105) Eftink, M. R.; Ghiron, C. A. *Anal. Biochem.* **1981**, *114*, 199.
- (106) Bujalowski, W.; Klonowska, M. M. *Biochemistry* **1993**, *32*, 5888.
- (107) Jezewska, M. J.; Lucius, A. L.; Bujalowski, W. *Biochemistry* **2005**, *44*, 3865.
- (108) Jezewska, M. J.; Lucius, A. L.; Bujalowski, W. *Biochemistry* **2005**, *44*, 3877.
- (109) De Gaaf, J.; Crossa, J. H.; Heffron, F.; Falkow, S. *J. Bacteriol.* **1978**, *134*, 1117.
- (110) Guerry, P.; van Embden, J.; Falkow, S. *J. Bacteriol.* **1974**, *117*, 987.
- (111) Scherzinger, E.; Ziegelin, G.; Barcena, M.; Carazo, J. M.; Lurz, R.; Lanka, E. *J. Biol. Chem.* **1997**, *272*, 30228.
- (112) Roleke, D.; Hoier, H.; Bartsch, C.; Umbach, P.; Scherzinger, E.; Lurz, R.; Saenger, W. *Acta Crystallogr.* **1997**, *D53*, 213.
- (113) Niedenzu, T.; Roleke, D.; Bains, Scherzinger, E.; Saenger, W. *J. Mol. Biol.* **2001**, *306*, 479.
- (114) Bernasconi, C. F. *Relaxation Kinetics*; Academic Press: New York, 1976; Chapters 3–7.
- (115) Strehlow, H. *Rapid Reactions in Solution*; VCH: New York, 1992; Chapter 4.
- (116) Bujalowski, W.; Jezewska, M. J. *J. Mol. Biol.* **2000**, *295*, 831.
- (117) Galletto, R.; Bujalowski, W. *Biochemistry* **2002**, *41*, 8907.
- (118) Bjorson, K. P.; Hsieh, J.; Amaratunga, M.; Lohman, T. M. *Biochemistry* **1998**, *37*, 891.
- (119) Kozlov, A. G.; Lohman, T. M. *Biochemistry* **2002**, *41*, 6032.
- (120) Kleinschmidt, C.; Tovar, K.; Hillen, W.; Porschke, D. *Biochemistry* **1988**, *27*, 1094.
- (121) Pilar, F. L. *Elementary Quantum Chemistry*; McGraw-Hill: New York, 1968; Chapter 9.
- (122) Amundson, N. R. *Mathematical Methods in Chemical Engineering. Matrices and their Application*; Prentice-Hall: Englewood Cliffs, NJ, 1966; Chapter 5.
- (123) Galletto, R.; Jezewska, M. J.; Bujalowski, W. *Biochemistry* **2003**, *43*, 11002.
- (124) Berry, R. S.; Rice, S. A.; Ross, J. *Physical Chemistry*; J. Wiley & Sons: New York, 1980; Chapter 15.
- (125) Moore, J. W.; Pearson, R. G. *Kinetics and Mechanism*; J. Wiley & Sons: New York, 1981; Chapter 6.
- (126) Smoluchowski, M. *Z. Phys. Chem.* **1917**, *92*, 129.
- (127) Tanford, C. *Physical Chemistry of Macromolecules*; John Wiley & Sons: New York, 1961; Chapter 6.
- (128) Bujalowski, W.; Klonowska, M. M. *J. Biol. Chem.* **1994**, *269*, 31359.
- (129) Tolman, G. L.; Barrio, J. R.; Leonard, N. J. *Biochemistry* **1974**, *13*, 4869.
- (130) Baker, B. M.; Vanderkooi, J.; Kallenbach, N. R. *Biopolymers* **1978**, *17*, 1361.
- (131) Bujalowski, W.; Klonowska, M. M. *Biochemistry* **1994**, *33*, 4682.
- (132) Porschke, D. *Eur. J. Biochem.* **1973**, *39*, 117.
- (133) Porschke, D. *Biopolymers* **1978**, *17*, 315.
- (134) Jones, J. M.; Nakai, H. *J. Mol. Biol.* **2001**, *312*, 935.
- (135) Jones, J. M.; Nakai, H. *J. Mol. Biol.* **1999**, *289*, 503.
- (136) Mulsch, A.; Colpan, M.; Wollny, E.; Gassen, H. G.; Riesner, D. *Nucleic Acids Res.* **1981**, *9*, 2376.
- (137) Riesner, D.; Pingoud, A.; Boehme, D.; Peters, F.; Maass, G. *Eur. J. Biochem.* **1976**, *68*, 71.
- (138) Krauss, G.; Riesner, D.; Maass, G. *Eur. J. Biochem.* **1976**, *68*, 81.

CR040462L