Published in final edited form as:

Methods Enzymol. 2011 ; 488: 17–57. doi:10.1016/B978-0-12-381268-1.00002-1.

Macromolecular Competition Titration Method: Accessing Thermodynamics of the Unmodified Macromolecule–Ligand Interactions Through Spectroscopic Titrations of Fluorescent Analogs

Włodzimierz Bujalowski^{*,†} and Maria J. Jezewska^{*}

^{*} Department of Biochemistry and Molecular Biology, The Sealy Center for Structural Biology, Sealy Center for Cancer Cell Biology, The University of Texas Medical Branch at Galveston, Galveston, Texas, USA

[†] Department of Obstetrics and Gynecology, The Sealy Center for Structural Biology, Sealy Center for Cancer Cell Biology, The University of Texas Medical Branch at Galveston, Galveston, Texas, USA

Abstract

Analysis of thermodynamically rigorous binding isotherms provides fundamental information about the energetics of the ligand–macromolecule interactions and often an invaluable insight about the structure of the formed complexes. The Macromolecular Competition Titration (MCT) method enables one to quantitatively obtain interaction parameters of protein–nucleic acid interactions, which may not be available by other methods, particularly for the unmodified long polymer lattices and specific nucleic acid substrates, if the binding is not accompanied by adequate spectroscopic signal changes. The method can be applied using different fluorescent nucleic acids or fluorophores, although the etheno-derivatives of nucleic acid are especially suitable as they are relatively easy to prepare, have significant blue fluorescence, their excitation band lies far from the protein absorption spectrum, and the modification eliminates the possibility of base pairing with other nucleic acids. The MCT method is not limited to the specific size of the reference nucleic acid. Particularly, a simple analysis of the competition titration experiments is described in which the fluorescent, short fragment of nucleic acid, spanning the exact site-size of the protein–nucleic acid complex, and binding with only a 1:1 stoichiometry to the protein, is used as a reference macromolecule. Although the MCT method is predominantly discussed as applied to studying protein–nucleic acid interactions, it can generally be applied to any ligand–macromolecule system by monitoring the association reaction using the spectroscopic signal originating from the reference macromolecule in the presence of the competing macromolecule, whose interaction parameters with the ligand are to be determined.

1. Introduction

Macromolecule–ligand interactions play a fundamental role in metabolic activities of enzymes, structural proteins, and nucleic acids (Boschelli, 1982; Brenowitz *et al.*, 1986; Bujalowski and Jezewska, 1995; Bujalowski and Klonowska, 1993, 1994a,b; Bujalowski and Lohman, 1986, 1987; Bujalowski *et al.*, 1989, 1994; Chabbert *et al.*, 1987; Crothers, 1968; deHaseth *et al.*, 1977; Draper and von Hippel, 1978; Epstein, 1978; Fernando and Royer, 1992; Fried and Crothers, 1981; Garner and Revzin, 1981; Heyduk and Lee, 1990; Kowalczykowski *et al.*, 1981, 1986; Lohman and Bujalowski, 1991; Lohman and Ferrari, 1994; McGhee and von Hippel, 1974; McSwiggen *et al.*, 1988; Menetski and Kowalczykowski, 1985; Porschke and Rauh, 1983; Revzin and von Hippel, 1977; Scatchard, 1949). The interactions can achieve various levels of complexity. For instance, the binding

process may be very specific, limited to a single type of ligand molecule. Specificity may encompass a group of ligand molecules of homologous structures or the interactions may include several ligand molecules, and the binding process may be characterized by cooperativity. In any case, the first step in elucidating the intrinsic energetics of the association, the forces which drive the interactions, the functionally relevant conformational changes of the participating species is the determination of the equilibrium binding isotherm over a range of the total average degree of binding, $\Sigma\theta_i$ (number of moles of the ligand bound per one mole of the macromolecule). Equilibrium binding isotherms, for ligand binding to a macromolecule, represent the relationship between $\Sigma\theta_i$ and the free ligand concentration. A true thermodynamic isotherm reflects only this relationship and does not depend upon any particular binding models or any particular assumptions. The extraction of physically meaningful interaction parameters that characterize the examined complex is only possible when such a relationship is available and is achieved by comparing the experimental isotherms to theoretical predictions that incorporate known molecular aspects of the examined systems (Brenowitz *et al.*, 1986; Bujalowski and Klonowska, 1993, 1994a,b; Bujalowski and Lohman, 1986, 1987; Bujalowski *et al.*, 1989, 1994; Fried and Crothers, 1981; Garner and Revzin, 1981; Heyduk and Lee, 1990; Lohman and Bujalowski, 1991; Lohman and Ferrari, 1994; Revzin and von Hippel, 1977; Scatchard, 1949).

In our discussion, we will concentrate on the thermodynamic analyses of the ligand–macromolecule binding, including protein–DNA interactions, using spectroscopic approaches (Bujalowski and Klonowska, 1993, 1994a,b; Bujalowski *et al.*, 1994; Lohman and Bujalowski, 1991). Spectroscopic methods are widely used in studying the most complex ligand–nucleic acid associations (Boschelli, 1982; Bujalowski and Klonowska, 1993, 1994a,b; Bujalowski and Lohman, 1986, 1987; Bujalowski *et al.*, 1994; Chabbert *et al.*, 1987; Draper and von Hippel, 1978; Fernando and Royer, 1992; Kowalczykowski *et al.*, 1986; Lohman and Ferrari, 1994; McSwiggen *et al.*, 1988; Menetski and Kowalczykowski, 1985; Porschke and Rauh, 1983). This is because the spectroscopic methods are relatively easy to apply, do not require large quantities of materials, allow one to perform solution experiments without perturbing the studied equilibrium, and can be applied at very low, submicromolar concentrations of interacting macromolecules, thus, eliminating any correction for nonideality necessary when high concentrations are required, for example, in transport or chromatographic methods. Moreover, the observed changes of the spectroscopic parameters provide additional information about the structural changes accompanying the examined interactions. We will discuss the quantitative analyses as applied to the use of the fluorescence intensity, which is the most widely applied spectroscopic technique in energetics and dynamics studies of the ligand–macromolecule interactions in solution (Bujalowski and Klonowska, 1993, 1994a,b; Bujalowski and Lohman, 1987; Bujalowski *et al.*, 1994; Lohman and Bujalowski, 1991). Nevertheless, the derived relationships are general and applicable to any physico-chemical signal used to monitor the association (Bujalowski, 2006; Jezewska and Bujalowski, 1996; Lohman and Bujalowski, 1991; Szymanski *et al.*, 2010). The primary limitation of spectroscopic methods for studying many ligand–macromolecule interactions is the lack of a sufficient spectroscopic signal change accompanying the complex formation. Lack of adequate changes in spectroscopic properties of the interacting molecules hinders the rigorous quantitative analyses of such systems. It can be overcome by introducing a fluorescence label into the ligand or the macromolecule. Nevertheless, besides the nonspecific nature of labeling a protein, or a polymer nucleic acid, which limits these approaches to some particular applications, for example, use of short nucleic acids, labeling SH groups of the protein, there is always a problem as to what extent the fluorescent marker perturbs the system, or reflects the interactions between the unmodified components.

The Macromolecular Competition Titration (MCT) method is particularly suited to analyze macromolecule–ligand interactions where the formation of the complexes is not accompanied by any adequate spectroscopic signal change (Bujalowski, 2006; Bujalowski and Jezewska, 2009; Jezewska and Bujalowski, 1996; Szymanski *et al.*, 2010). The method is based on quantitative titrations of the reference fluorescent macromolecule/ligand with a macromolecule/ligand, in the presence of a competing, nonfluorescent macromolecule/ligand, whose interaction parameters are to be determined. The approach allows the determination of the total average degree of binding and the free macromolecule/ligand concentrations, over a large degree of binding range, unavailable by other methods, and construction of a model-independent, thermodynamic binding isotherm. In the case of protein binding to a long, one-dimensional nucleic acid lattice, a more convenient parameter than $\Sigma\theta_i$ is the total average binding density $\Sigma\nu_i$, (moles of ligand bound per mole of bases or base pairs; Bujalowski, 2006; Jezewska and Bujalowski, 1996). Using the MCT method, the determination of $\Sigma\theta_i$, or $\Sigma\nu_i$, is independent of the *a priori* knowledge of the binding characteristics of the macromolecule/ligand to the reference ligand/macromolecule.

2. A Single Titration Curve: Some Simple Considerations of Possible Pitfalls

Fluorescence titration, or any titration method based on observing a physico-chemical signal, is an indirect method of determining the binding isotherm. This is because the association of the ligand with the macromolecule is examined by monitoring changes in the parameter of the system (here, the emission intensity) accompanying the formation of the complex, not the total average degree of binding or the binding density (Bujalowski, 2006; Bujalowski and Jezewska, 2000, 2009; Bujalowski and Klonowska, 1993, 1994a,b; Bujalowski and Lohman, 1987; Bujalowski *et al.*, 1994; Jezewska *et al.*, 1996, 1998a,b, 2001a,b, 2003a,b, 2006; Lohman and Bujalowski, 1991; Rajendran *et al.*, 1998, 2001; Szymanski *et al.*, 2010). These changes must then be correlated with $\Sigma\theta_i$, or $\Sigma\nu_i$, or with the fractional saturation of the macromolecule, and the concentration of the free ligand. Nevertheless, the functional relationship between the observed spectroscopic signal and the total average degree of binding or the total average binding density is never *a priori* known. The only exceptions from this rule are systems where, at saturation, a single ligand molecule binds to the macromolecule. But then, the maximum stoichiometry of the system is already known. In much more general cases, where there is a possibility that multiple ligand molecules can participate in the binding process, the maximum stoichiometry is one of the thermodynamic/structural parameters that must be determined. Moreover, the observed, fractional change of the spectroscopic signal and the total average degree of binding may not, and frequently will not, have a simple and strict proportional relationship (Bujalowski, 2006; Jezewska and Bujalowski, 1996; Szymanski *et al.*, 2010).

For instance, although one can analyze a titration curve by assuming a simple, 1:1 stoichiometry of the binding process, such an analysis is only of a semiquantitative nature, at best, and does not provide any deeper insight into the studied interactions. Figure 2.1A shows a typical titration curve for the ligand binding to six independent sites on the macromolecule, as monitored by the fluorescence emission and performed at a single macromolecule concentration. The selected binding constant, K , and the total macromolecule concentration, $[M]_T$, are $1 \times 10^{-6} M^{-1}$ and $1 \times 10^{-7} M$, respectively. The titration curve is described by Eq. (2.1)

$$\Delta F_{\text{obs}} = \Delta F_{\text{max}} \frac{6K[L]_F}{1 + K[L]_F} = \Delta F_{\text{max}} \sum \theta_i \quad (2.1)$$

where an obvious strict proportionality exists between the observed signal, ΔF_{obs} , and the total average degree of binding, $\Sigma \Theta_i$. The proportionality constant is the maximum observed signal change, $\Delta F_{\text{max}} = 0.5$, for the ligand binding to each of the six sites. The analogous equation for a single ligand molecule binding to a single site on the macromolecule is

$$\Delta F_{\text{obs}} = \Delta F_{\text{max}} \frac{K[L]_F}{1 + K[L]_F} = \Delta F_{\text{max}} \sum \Theta_i \quad (2.2)$$

The corresponding plot of the titration curve, defined by Eq. (2.2), is included in Fig. 2.1A with $K = 8 \times 10^5 \text{ M}^{-1}$ and $\Delta F_{\text{max}} = 3.0$. Within experimental accuracy, the plots are indistinguishable. The error in determining the binding constant (~20%) is significant, but probably not that troublesome in typical biochemical studies. Such a “small” error results from the fact that the concentration of the macromolecule is much lower than the total concentration of the ligand, that is, the concentration of the bound ligand, $[L]_B$, is negligible, as compared to the total ligand concentration, $[L]_T$. However, the character of the binding process, with all conclusions as to the functionality of the macromolecule, is completely missed.

The situation is even worse when the titrations are performed at a single, high macromolecular concentration. In such a case, $[L]_B$ constitutes a significant part of $[L]_T$. In the simplified analysis, which assumes the 1:1 stoichiometry, this would not be *a priori* known to the experimenter and simply ignored. Figure 2.1B shows a simulated titration curve for ligand binding to six independent binding sites on the macromolecule, monitored by the fluorescence emission, with the binding constant, K , and the total macromolecule concentration, $[M]_T$, $1 \times 10^6 \text{ M}^{-1}$ and $1 \times 10^{-6} \text{ M}$, respectively. The plot is described by Eq. (2.1). The single-site binding model cannot fit the curve even for a less quantitatively oriented experimenter (Eq. (2.2)). Moreover, the apparent behavior of the plot would suggest a positive cooperativity in the binding process. The experimenter would be then tempted to use the next simplest model, which would be the ligand binding to two sites with cooperative interactions and accompanied by the same fluorescence emission changes, as defined by

$$\Delta F_{\text{obs}} = \Delta F_{\text{max}} \frac{K[L]_F + 2K^2\sigma[L]_F^2}{1 + K[L]_F + K^2\sigma[L]_F^2} = \Delta F_{\text{max}} \sum \Theta_i \quad (2.3)$$

This model nicely fits the titration curve, although, once again, completely misses the nature of the occurring binding process, which is an independent binding of six ligand molecules to the macromolecule, defined by Eq. (2.1). The obtained binding constant and cooperativity parameter are fitting constants, not the thermodynamic quantities characterizing the system. The example discussed above indicates some pitfalls, but are not, by any means, the most complex possibilities. It simply reflects the fact that a single titration curve, where a spectroscopic or a physico-chemical signal is used to monitor the binding, is not a binding isotherm (see above).

3. Quantitative Equilibrium Spectroscopic Titrations: Thermodynamic Bases

Quantitative analysis of the binding of a ligand to a macromolecule can be performed using two different types of equilibrium spectroscopic titrations, “normal” or “reverse” titration

approach (Boschelli, 1982; Bujalowski, 2006; Bujalowski and Jezewska, 2000, 2009; Lohman and Bujalowski, 1991; Szymanski *et al.*, 2010). In the normal titration, a fluorescing macromolecule is titrated with a nonfluorescing ligand. The total average degree of binding, $\Sigma\theta_i$, increases as the titration progresses. In the reverse titration approach, the fluorescing ligand is titrated with the non-fluorescing macromolecule and $\Sigma\theta_i$ decreases throughout the titration (Bujalowski and Jezewska, 2000; Jezewska and Bujalowski, 1997; Lohman and Bujalowski, 1991). Although both types of titration curves can be quantitatively analyzed, we will limit our discussion to the normal titration method, as more commonly occurring in the experimental practice. Moreover, we will discuss specific ligand–macromolecule interactions examined in our laboratory, which include nucleotide binding to DNA helicases and protein–nucleic acid interactions, where the binding processes have been addressed using protein fluorescence, emission of fluorescent analogs of the nucleotides, and fluorescent derivative of homoadenosine oligomers (Bujalowski, 2006; Bujalowski and Jezewska, 2000, 2009; Bujalowski and Klonowska, 1993, 1994a,b; Bujalowski *et al.*, 1994; Jezewska and Bujalowski, 1996, 1997; Jezewska *et al.*, 1996, 1998a,b, 2001a,b, 2003a,b, 2006; Rajendran *et al.*, 1998, 2001; Szymanski *et al.*, 2010). For multiple binding of nucleotides to the protein, it is rather obvious that the nucleotide cofactor is the ligand and the protein is the macromolecule. Analogously, because usually several protein molecules bind to the examined nucleic acids, the protein is treated as a ligand and the nucleic acid is the macromolecule.

The task is to convert the fluorescence titration curve, for example, a change in the monitored fluorescence as a function of the total protein concentration into a model-independent, thermodynamic binding isotherm, which can be analyzed, using an appropriate binding model to extract binding parameters. We start with the signal and mass conservation equations (Bujalowski, 2006; Bujalowski and Jezewska, 2000, 2009; Jezewska and Bujalowski, 1996; Jezewska *et al.*, 1996, 1998a, b, 2001a, b, 2003a, b, 2006; Lohman and Bujalowski, 1991; Rajendran *et al.*, 1998, 2001; Szymanski *et al.*, 2010). For the total macromolecule concentration, $[M]_T$, the equilibrium distribution of the macromolecule among its different states with a different number of bound ligand molecules, M_i , is determined solely by the free protein concentration, $[L]_F$ (Bujalowski, 2006; Bujalowski and Jezewska, 2000, 2009; Jezewska and Bujalowski, 1996; Jezewska and Bujalowski, 1997; Jezewska *et al.*, 1996, 1998a,b, 2001, 2003, 2006; Lohman and Bujalowski, 1991; Rajendran *et al.*, 1998, 2001; Szymanski *et al.*, 2010). Therefore, at each $[L]_F$, the observed spectroscopic signal, F_{obs} , is the algebraic sum of the macromolecule concentrations in each state, M_i , weighted by the value of the intensive spectroscopic property of that state, F_i . The “signal conservation” equation for the observed fluorescence emission, F_{obs} , is then

$$F_{obs} = F_F[M]_F + \sum F_i[M]_i \quad (2.4)$$

where F_F is the molar fluorescence of the free nucleic acid and F_i is the molar fluorescence of the complex, $[M]_i$, which represents the concentration of the macromolecule with i bound ligand molecules ($i = 1$ to n). Another mass conservation equation relates $[M]_F$ and $[M]_i$ to $[M]_T$ by

$$[M]_T = [M]_F + \sum [M]_i \quad (2.5)$$

Next, we define the partial degree of binding, Θ_i (“ i ” moles of ligand bound per mole of macromolecule), corresponding to all complexes with a given number “ i ” of bound ligand molecules as

$$\Theta_i = \frac{i[M]_i}{[M]_T} \quad (2.6)$$

Therefore, the concentration of the macromolecule with “ i ” protein molecules bound, $[M]_i$, is defined as

$$[M]_i = \left(\frac{\Theta_i}{i} \right) [M]_T \quad (2.7)$$

Introducing Eqs. (2.6) and (2.7) into Eq. (2.4) provides a general relationship for the observed fluorescence, F_{obs} , as

$$F_{\text{obs}} = F_F [M]_T + \left[\sum (F_i - F_F) \left(\frac{\Theta_i}{i} \right) \right] [M]_T \quad (2.8)$$

By rearranging Eq. (2.8), one can define the experimentally accessible quantity, ΔF_{obs} , that is, the fluorescence change normalized with respect to the initial fluorescence intensity of the free macromolecule, as

$$\Delta F_{\text{obs}} = \frac{(F_{\text{obs}} - F_F [M]_T)}{F_F [M]_T} \quad (2.9)$$

and

$$\Delta F_{\text{obs}} = \sum \left(\frac{\Delta F_i}{i} \right) (\Theta_i) \quad (2.10)$$

The quantity, ΔF_{obs} , is the experimentally determined fractional fluorescence change observed at the selected total ligand and macromolecule concentrations, $[L]_T$ and $[M]_T$. The quantity, $\Delta F_i/i$, is the average molar fluorescence change per bound ligand in the complex containing “ i ” ligand molecules. Because $\Delta F_i/i$ is an intensive molecular property of the ligand–macromolecule complex with “ i ” ligand molecules bound, Eq. (2.10) indicates that ΔF_{obs} is only a function of the free ligand concentration through the total average degree of binding, $\Sigma \Theta_i$. Therefore, for a given and specific value of ΔF_{obs} , the total average degree of binding, $\Sigma \Theta_i$, must be the same for any value of $[L]_T$ and $[M]_T$ (Bujalowski, 2006; Bujalowski and Jezewska, 2000, 2009; Jezewska and Bujalowski, 1996, 1997; Jezewska *et al.*, 1996, 1998a,b, 2001, 2003, 2006; Lohman and Bujalowski, 1991; Rajendran *et al.*, 1998, 2001; Szymanski *et al.*, 2010). Thus, if one performs a fluorescence titration of the macromolecule with the ligand, at different $[M]_T$, the same value of ΔF_{obs} at different $[M]_T$ s indicates the same physical state of the macromolecule, that is, the same degree of

macromolecule saturation with the ligand, that is, the same $\Sigma\Theta_i$. Since $\Sigma\Theta_i$ is a unique function of the free ligand concentration, $[L]_F$, the value of $[L]_F$, at the same total average degree of binding, must also be the same. An analogous expression can be derived for the case where the spectroscopic signal originates from the ligand and the binding analysis is performed using the reverse titration method (Bujalowski, 2006; Bujalowski and Jezewska, 2000, 2009; Jezewska and Bujalowski, 1996, 1997; Jezewska *et al.*, 1996, 1998a,b, 2001, 2003, 2006; Lohman and Bujalowski, 1991; Rajendran *et al.*, 1998, 2001; Szymanski *et al.*, 2010). Expression (2.10) indicates a clear method of transforming the fluorescence titration curve into a thermodynamic binding isotherm (Bujalowski, 2006; Bujalowski and Jezewska, 2000, 2009; Jezewska and Bujalowski, 1996, 1997; Jezewska *et al.*, 1996, 1998a,b, 2001, 2003, 2006; Lohman and Bujalowski, 1991; Rajendran *et al.*, 1998, 2001; Szymanski *et al.*, 2010). In optimal and minimal case, one can perform only two titrations at two different total concentrations of the macromolecule, that is, $[M]_{T1}$ and $[M]_{T2}$. At the same value of ΔF_{obs} , the total average degree of binding, $\Sigma\Theta_i$, and the free ligand concentrations, $[L]_F$, must be the same for both titration curves (see above).

Two hypothetical fluorescence titration curves are illustrated in Fig. 2.2A for the binding process where six ligand molecules bind to six identical binding sites on the macromolecule, characterized by intrinsic binding constants, K and the relative fluorescence changes, $\Delta F_{max} = 0.5$, for the ligand binding to each site, respectively, as defined by Eq. (2.1). The values of ΔF_{obs} are plotted as a function of the logarithm of the total ligand concentration, $[L]_T$. At a higher macromolecule concentration, a given relative fluorescence increase, ΔF_{obs} , is reached at higher ligand concentrations, as more ligand is required to saturate the macromolecule at its higher concentration. A set of values of $(\Sigma\Theta_i)_j$ and $([L]_F)_j$ for the selected “ j ” value of $(\Delta F_{obs})_j$ is obtained from these data in the following manner. One draws a horizontal line that intersects both titration curves at the same value of $(\Delta F_{obs})_j$; Fig. 2.2A). The point of intersection of the horizontal line with each titration curve defines two values of the total ligand concentration, $([L]_{T1})_j$ and $([L]_{T2})_j$, for which $([L]_F)_j$ and $(\Sigma\Theta_i)_j$ are the same (see above; Bujalowski, 2006; Bujalowski and Jezewska, 2000, 2009; Jezewska and Bujalowski, 1996, 1997; Jezewska *et al.*, 1996, 1998a,b, 2001, 2003, 2006; Lohman and Bujalowski, 1991; Rajendran *et al.*, 1998, 2001; Szymanski *et al.*, 2010). Then, one has two mass conservation equations for the total concentrations of the protein, $([L]_{T1})_j$ and $([L]_{T2})_j$, as

$$([L]_{T1})_j = ([L]_F)_j + \left(\sum \Theta_i\right)_j [M]_{T1} \quad (2.11a)$$

and

$$([L]_{T2})_j = ([L]_F)_j + \left(\sum \Theta_i\right)_j [M]_{T2} \quad (2.11b)$$

from which one obtains that at a given $(\Delta F_{obs})_j$, the total average degree of binding and the free protein concentration are

$$\left(\sum \Theta_i\right)_j = \frac{([L]_{T2})_j - ([L]_{T1})_j}{[M]_{T1} - [M]_{T2}} \quad (2.12)$$

and

$$([L]_F)_j = ([L]_{Tx})_j - \left(\sum \Theta_i \right)_j ([M]_{Tx}) \quad (2.13)$$

where subscript, x , is 1 or 2 (Bujalowski, 2006; Bujalowski and Jezewska, 2000, 2009; Jezewska and Bujalowski, 1996, 1997; Jezewska *et al.*, 1996, 1998a,b, 2001, 2003, 2006; Lohman and Bujalowski, 1991; Rajendran *et al.*, 1998, 2001; Szymanski *et al.*, 2010). Performing a similar analysis along the titration curves at a selected interval of the observed signal change, one obtains model-independent values of $([L]_F)_j$ and $(\sum \Theta_i)_j$ at any selected “ j ” value of $(\Delta F_{obs})_j$. The most accurate estimates of $([L]_F)_j$ and $(\sum \Theta_i)_j$ are obtained in the region of the titration curves where the concentration of a bound ligand is comparable to its total concentration, $[L]_T$. In our practice, this limits the accurate determination of the total average degree of binding and $[L]_F$ to the region of the titration curves where the concentration of the bound ligand is at least ~10–15% of the $[L]_T$. Selection of suitable concentrations of the macromolecule is of paramount importance for obtaining $([L]_F)_j$ and $(\sum \Theta_i)_j$ over the largest possible region of the titration curves. The selection of the macromolecule concentrations is based on preliminary titrations that provide initial estimates of the expected affinity. Nevertheless, the accuracy of the determination of $(\sum \Theta_i)_j$ is mostly affected in the region of the high concentrations of the ligand, that is, where the binding process approaches the maximum saturation.

Usually, the maximum of the recorded fluorescence changes, that is, the saturation of the observed binding process, is attainable with adequate accuracy. What is mostly unknown is the maximum stoichiometry of the complex at saturation (Bujalowski, 2006; Bujalowski and Jezewska, 1995, 2000, 2009; Bujalowski and Klonowska, 1993; Jezewska and Bujalowski, 1996, 1997; Jezewska *et al.*, 1996, 1998a,b, 2001, 2003, 2006; Lohman and Bujalowski, 1991; Rajendran *et al.*, 1998, 2001; Szymanski *et al.*, 2010). Another unknown is the relationship between the fluorescence change, ΔF_{obs} , and the total average degree of binding (Bujalowski, 2006; Bujalowski and Jezewska, 1995, 2000, 2009; Bujalowski and Klonowska, 1993; Jezewska and Bujalowski, 1996, 1997; Jezewska *et al.*, 1996, 1998a,b, 2001, 2003, 2006; Lohman and Bujalowski, 1991; Rajendran *et al.*, 1998, 2001; Szymanski *et al.*, 2010). Both unknowns can be determined by plotting ΔF_{obs} as a function of $\sum \Theta_i$, as depicted in Fig. 2.2B, for the hypothetical binding model considered here. The plot is clearly linear because we selected the same value of $\Delta F_{max} = 0.5$ for each of the six binding sites. As mentioned above, in practice, the maximum value of $\sum \Theta_i$ cannot be directly determined, due to the inaccuracy at the high-ligand concentration region. However, knowing the maximum increase of the macromolecule fluorescence ΔF_{max} , one can perform a short extrapolation of the plot (ΔF_{obs} vs. $\sum \Theta_i$) to this maximum value of the observed fluorescence change, which establishes the maximum stoichiometry of the formed complex at saturation (Bujalowski, 2006; Bujalowski and Jezewska, 1995, 2000, 2009; Bujalowski and Klonowska, 1993; Jezewska and Bujalowski, 1996, 1997; Jezewska *et al.*, 1996, 1998a,b, 2001, 2003, 2006; Lohman and Bujalowski, 1991; Rajendran *et al.*, 1998, 2001; Szymanski *et al.*, 2010).

4. Nucleotide Binding to the RepA Protein of Plasmid RSF1010

As an example of the real analysis, we consider the binding of the nucleotide cofactors to the RepA protein of plasmid RSF1010 (Jezewska *et al.*, 2005a,b). RepA protein is a hexameric helicase specifically involved in the replication of nonconjugative plasmid RSF1010, which confers bacterial resistance to sulfonamides and streptomycin. 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, binding of nucleotide analogs, 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. The hydrolysis of TNP-ATP is undetectable on the time scale of the binding experiments (Jezewska *et al.*, 2005a,b).

Fluorescence titrations of the RepA hexamer with TNP-ADP, at three different protein concentrations, are shown in Fig. 2.3A. 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 (Jezewska *et al.*, 2005a,b). The spectroscopic titration curves in Fig. 2.3A would indicate an apparent single binding process. However, analysis of the data, using the thermodynamically rigorous approach, as discussed above, reveals a much more complex behavior (Jezewska *et al.*, 2005a,b). Because we have three titration curves, we can select any two of them to determine the total average degree of binding. However, higher accuracy is usually attained with the most separated titration curves, with respect to the total ligand concentrations.

The dependence of the observed fluorescence quenching, ΔF_{obs} , upon the total average degree of binding, $\Sigma\theta_i$, of the TNP-ADP on the RepA hexamer is shown in Fig. 2.3B. The separation of the binding isotherms allows us to obtain the values of $\Sigma\theta_i$ up to ~ 5.1 of the TNP-ADP molecules per RepA hexamer. The plot in Fig. 2.3B is clearly nonlinear. For comparison, the dashed line represents the hypothetical case where a strict proportionality between the total average degree of binding and the quenching of the RepA hexamer fluorescence would exist. Thus, binding of the first three molecules of the nucleotide induces a larger decrease of the protein fluorescence emission than the binding of the subsequent cofactor molecules. Because the maximum quenching is available from the experiment (Fig. 2.3A), 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. In spite of the nonlinear behavior of the plot in Fig. 2.3B, such extrapolation is rather easy because the values of $\Sigma\theta_i$ could be determined up to $\sim 90\%$ of the observed signal. In our practice, we select the last four to five points of the plot and perform a linear regression. The maximum stoichiometry is determined by the interception of the regression line with the horizontal line, which indicates the maximum value of ΔF . For nonlinear experimental curves, which cover $\sim 70\text{--}80\%$ of the observed signal, the error in determination of the maximum stoichiometry is $\sim 10\text{--}15\%$ (Bujalowski, 2006; Jezewska and Bujalowski, 1996; Jezewska *et al.*, 2005a,b). The situation is much simpler with the linear plots where the plot of ΔF versus $\Sigma\theta_i$, covering $\sim 50\%$, provides the maximum stoichiometry of the formed complex with similar, $\sim 10\text{--}15\%$ accuracy.

The determination of the maximum stoichiometry of the nucleoside triphosphate analog, TNP-ATP, binding to the RepA hexamer has been analogously performed. Fluorescence titrations of the RepA protein with TNP-ATP at three different protein concentrations are shown in Fig. 2.3C. At saturation, the maximum quenching of RepA helicase emission is 0.83 ± 0.03 . The dependence of the protein fluorescence quenching, ΔF_{obs} , upon the total average degree of binding of TNP-ATP on the RepA hexamer is shown in Fig. 2.3D. The separation of the isotherms allows us to obtain $\Sigma\theta_i$ up to ~ 4.6 . Similar to the TNP-ADP case (Fig. 2.3B), the plot is nonlinear, indicating that the binding of the first three cofactor molecules is accompanied by a larger quenching of the protein fluorescence than the quenching accompanying the saturation of the remaining nucleotide-binding sites. Extrapolation to the maximum value of $\Delta F_{\text{max}} = 0.83 \pm 0.03$ shows that, at saturation, 6.0 ± 0.5 TNP-ATP molecules bind to the RepA hexamer. Because we have quantitatively determined the average degree of binding, $\Sigma\theta_i$, over $\sim 80\%$ of the spectroscopic titration curves, we can obtain the free nucleotide concentrations, using Eq. (2.13), and construct a

true thermodynamic binding isotherm, that is, the average degree of binding, $\Sigma\theta_i$, as a function of the free nucleotide concentration (Jezewska *et al.*, 2005a,b). The dependence of $\Sigma\theta_i$ upon the free concentration of the [TNP-ADP]_F is shown in Fig. 2.4A. An analogous plot of $\Sigma\theta_i$, as a function of [TNP-ATP]_F, is shown in Fig. 2.4B. In contrast to the fluorescence titrations in Fig. 2.3A and C, the plots in Fig. 2.4A and B are true thermodynamic isotherms.

5. Applying the Statistical Thermodynamic Model for the Nucleotide Binding to the RSF1010 RepA Protein Hexamer

Once the binding isotherms are available, one can address the intrinsic affinities and cooperativities in the system (Bujalowski, 2006; Bujalowski and Jezewska, 2000, 2009; Jezewska and Bujalowski, 1996, 1997; Jezewska *et al.*, 1996, 1998a,b, 2001, 2003, 2005a,b, 2006; Rajendran *et al.*, 1998, 2001; Szymanski *et al.*, 2010). The properties of the binding isotherms (Fig. 2.4A and B) and the known properties of the examined system provide crucial clues as to how to achieve that. A striking feature of the plots in Fig. 2.4A and B is that they extend over more than two orders of magnitude of the free nucleotide concentration, that is, between ~10% and ~90% of the maximum degree of binding. If the six binding sites had the same affinity for the ligand (nucleotide cofactor) then the free ligand concentration could not span more than two orders of magnitude between 10% and ~90% of the maximum degree of binding (Jezewska and Bujalowski, 1996; Jezewska *et al.*, 2005a,b). In other words, a binding system of multiple discrete binding sites, without heterogeneity in their intrinsic affinities, cannot generate such behavior. Therefore, these data unambiguously show that the macroscopic affinities of the nucleotide cofactors for the RepA hexamer are a decreasing function of $\Sigma\theta_i$.

The RepA hexamer is built of six identical protomers forming a ring-like structure with the protomer–protomer contacts limited to two adjacent subunits. The thermodynamic data discussed above show that the RSF1010 RepA hexamer binds six nucleotide molecules and the binding is characterized by the decreasing macroscopic affinity with the increase of the degree of binding. The simplest statistical thermodynamic model, that is, the model containing the smallest number of parameters, that can account for such behavior is the RepA hexamer exhibiting nearest-neighbor negative cooperativity among the otherwise identical six nucleotide-binding sites. This model is referred to as the hexagon model (Bujalowski and Klonowska, 1993, 1994a,b; Bujalowski *et al.*, 1994; Jezewska *et al.*, 2005a,b). The other possible simple model of the nucleotide binding to the six sites of the RepA hexamer, also containing only two binding parameters, is that the hexamer possesses two independent classes of independent binding sites, each class encompassing three independent binding sites. However, further analysis of the system, discussed in the original works, showed that this model does not apply to the RepA hexamer (Jezewska *et al.*, 2005a,b).

In the hexagon model, there are only two parameters that characterize the binding, the intrinsic binding constant, K , and the parameter, σ , describing cooperative interactions (Jezewska *et al.*, 2005a,b). The first ligand can bind to any of six, initially equivalent binding sites. The cooperative interactions are limited to only two neighboring sites. The partition functions, Z_H , is then

$$Z_H = 1 + 6x + 3(3 + 2\sigma)x^2 + 2(1 + 6\sigma + 3\sigma^2)x^3 + 3(3\sigma^2 + 2\sigma^3)x^4 + 6\sigma^4x^5 + \sigma^6x^6 \quad (2.14)$$

where $x = K[N]_F$, and $[N]_F$ is the free nucleotide concentration. The total average degree of binding, $\Sigma\Theta_i$, is defined by the standard statistical thermo-dynamic formula (Bujalowski and Klonowska, 1993, 1994a,b; Bujalowski *et al.*, 1994; Jezewska *et al.*, 2005a,b)

$$\Sigma\Theta_i = \frac{\partial \ln Z_H}{\partial \ln L} \quad (2.15)$$

and

$$\Sigma\Theta_i = \frac{6x + 6(3+2\sigma)x^2 + 6(1+6\sigma+3\sigma^2)x^3 + 12(3\sigma^2+2\sigma^3)x^4 + 30\sigma^4x^5 + 6\sigma^6x^6}{Z_H} \quad (2.16)$$

The solid lines in Fig. 2.4A and B are the nonlinear least square fits of the thermodynamic isotherms for TNP-ADP and TNP-ATP binding the RepA hexamer, according to the hexagon model, with only two binding parameters, using Eqs. (2.14)–(2.16). The obtained values of the intrinsic binding constants and cooperativity parameter σ are $K = (7.7 \pm 1.5) \times 10^6 M^{-1}$, $\sigma = 0.38 \pm 0.05$ and $K = (2.4 \pm 0.2) \times 10^6 M^{-1}$, $\sigma = 0.37 \pm 0.05$, for TNP-ADP and TNP-ATP, respectively.

6. Empirical Function Approach

Direct fitting of the spectroscopic titration curves may refine the obtained binding parameters, as it uses a larger span of the experimental curves. For the direct fitting of the spectroscopic titration curves, using Eqs. (2.14)–(2.16), one would have to know all molar fluorescence intensities of all possible RepA–nucleotide complexes. This is rather a hopeless task. Notice, the plots in Fig. 2.3B and D do not provide any information as to what the values of these parameters should be, with the exception of $\Delta F_{\max}$. On the other hand, the problem of finding all optical parameters can be avoided by using the empirical function approach (Bujalowski and Klonowska, 1993; Jezewska *et al.*, 2005a,b). This approach can be applied to any ligand–macromolecule systems, where the determination of all optical parameters is practically impossible. It is based on introducing the representation of the observed relative fluorescence change of the macro-molecule, ΔF_{obs} , as a function of the total average degree of binding, $\Sigma\Theta_i$, such as in Fig. 2.3B and D, via an empirical function (Bujalowski and Klonowska, 1993; Jezewska *et al.*, 2005a,b). The empirical function is usually a polynomial that accurately relates the experimentally determined ΔF_{obs} , to the experimentally determined, $\Sigma\Theta_i$, as

$$\Delta F_{\text{obs}} = \sum_{j=0}^n a_j (\Sigma\Theta_i)^j \quad (2.17)$$

where a_j are the fitting constants. The generation of the theoretical titration curve for the selected binding model is then accomplished by first calculating the value of $\Sigma\Theta_i$, for a given free ligand concentration and initial estimates of the binding parameters. Next, the obtained $\Sigma\Theta_i$ is introduced into Eq. (2.17) and the value of ΔF_{obs} , corresponding that value of $\Sigma\Theta_i$, is obtained. These calculations are then performed for the entire titration curve. The solid lines in Fig. 2.3A and B are generated using the empirical function approach based on the plots in Fig. 2.3B and D (Bujalowski and Klonowska, 1993; Jezewska *et al.*, 2005a,b).

7. MCT Method: General Considerations

So far, we discussed the methods of analysis of the titration curves that allow the experimenter to generate a model-independent thermodynamic binding isotherm. However, the analysis requires that the ligand binding is accompanied by significant changes in the spectroscopic signal originating from the macromolecule. On the other hand, many examined systems do not have a convenient signal to monitor the binding. One way to overcome this problem is to use fluorescence derivative of one of the interacting species, for example, the nucleic acid and/or protein, although this may not be feasible in all cases. A modification introduced on a protein may affect its activity. Modifications on the short fragments of a nucleic acid may create additional, undesirable interaction areas. Long polymeric nucleic acids provide an additional problem. For instance, while modification of homoadenosine polymers, to obtain fluorescent etheno-derivative, is relatively an easy reaction, obtaining homogeneously labeled fluorescent derivative of, for example, poly(dT) or poly(dC) is not (Baker *et al.*, 1978; Bujalowski, 2006; Bujalowski and Jezewska, 1995; Chabbert *et al.*, 1987; Menetski and Kowalczykowski, 1985; Secrist *et al.*, 1972; Tolman *et al.*, 1974). For this purpose, the MCT method 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 (Bujalowski, 2006; Jezewska and Bujalowski, 1996). Thus, the method is invaluable in addressing base- or sequence-specificity in protein–nucleic acid systems, which are otherwise very difficult to examine. In the first step, the interactions with a modified fluorescent system must be characterized. Next, the examination of interactions with an unmodified component of the reaction can be performed using competition studies (Bujalowski, 2006; Jezewska and Bujalowski, 1996).

The approach is based on the same thermodynamic argument, which leads to Eqs. (2.11a)–(2.13). The titration of a reference fluorescent macromolecule with the ligand is performed in the presence of the second competing, nonfluorescent macromolecule whose interactions with the ligand are to be determined. Although the ligand binds to two different macromolecules, in solution, the observed signal originates only from the fluorescent “reference” macromolecule. The saturation of both macromolecules increases with the increasing concentration of the free ligand. The entire binding system is composed of two macromolecules, reference and unmodified, competing for the ligand (Bujalowski, 2006; Jezewska and Bujalowski, 1996). In the absence of the competing unmodified macromolecule, the total concentration of the ligand, $[L]_{TR}$, at a particular value of the spectroscopic parameter (emission intensity), ΔF_{obs} , is (Fig. 2.5)

$$[L]_{TR} = [L]_F + [L]_{BR} \quad (2.18)$$

and

$$[L]_{TR} = [L]_F + (\sum \Theta_i) [M]_{TR} \quad (2.19)$$

where, $[L]_F$ is the free ligand concentration, $(\sum \Theta_i)_R$ is the total average degrees of binding on the fluorescent reference macromolecule and $[M]_{TR}$ is the total concentration of the reference macromolecule, respectively. In the presence of the unmodified macromolecule, the total concentration of the ligand, $[L]_{TS}$, which provides the same value of ΔF_{obs} , is

$$[L]_{TS} = [L]_F + [L]_{BR} + [L]_{BS} \quad (2.20)$$

and

$$[L]_{TS} = [L]_F + (\sum \Theta_i)_R [M]_{TR} + (\sum \Theta_i)_S [M]_{TS} \quad (2.21)$$

where, $(\sum \Theta_i)_S$ and $[M]_{TS}$ are the total average degrees of binding on the unmodified competing macromolecule, and the total concentration of the unmodified macromolecule, respectively. Subtracting Eq. (2.19) from Eq. (2.21) and rearranging provides

$$(\sum \Theta_i)_S = \frac{[L]_{TS} - [L]_{TR}}{[M]_{TS}} \quad (2.22)$$

The concentration of the free ligand, $[L]_F$, is then

$$[L]_F = [L]_{TR} - [L]_B = [L]_{TS} - (\sum \Theta_i)_R [M]_{TR} \quad (2.23)$$

or

$$[L]_F = [L]_{TS} - [L]_B = [L]_{TS} - (\sum \Theta_i)_R [M]_{TR} - (\sum \Theta_i)_S [M]_{TS} \quad (2.24)$$

Notice, one may determine $(\sum \Theta_i)_S$ without addressing the exact model for the ligand binding to the reference fluorescence macromolecule and treat it only as an indicator of the same free ligand concentration at a given value of the observed relative fluorescence intensity. This can be accomplished by performing titrations of the reference macromolecule at two or more concentrations of the unmodified macromolecule, $[L]_{TS1}$ and $[L]_{TS2}$ (Bujalowski, 2006; Jezewska and Bujalowski, 1996). In such an approach, one has, at a given value of ΔF_{obs} ,

$$[L]_{TS1} = [L]_F + (\sum \Theta_i)_R [M]_{TR} + (\sum \Theta_i)_S [M]_{TS1} \quad (2.25)$$

and

$$[L]_{TS2} = [L]_F + (\sum \Theta_i)_R [M]_{TR} + (\sum \Theta_i)_S [M]_{TS} \quad (2.26)$$

where $(\sum \Theta_i)_R$ and $[M]_{TR}$ have the same meaning as above, $(\sum \Theta_i)_S$, $[M]_{TS1}$, and $[M]_{TS2}$ are the total average degrees of binding on the unmodified competing macromolecule and the

two total concentrations of the unmodified macromolecule, respectively. Subtracting Eq. (2.25) from Eq. (2.26) and rearranging provides

$$(\sum \Theta_i)_S = \frac{[L]_{TS2} - [L]_{TS1}}{[M]_{TS2} - [M]_{TS1}} \quad (2.27)$$

Nevertheless, to obtain the concentration of the free ligand, $[L]_F$, $(\sum \Theta_i)_{RS}$ must be known, providing the free ligand concentration as

$$[L]_F = [L]_{TS1} - [L]_B = [L]_{TS1} - (\sum \Theta_i)_R [M]_{TR} - (\sum \Theta_i)_S [M]_{TS1} \quad (2.28)$$

or

$$[L]_F = [L]_{TS2} - [L]_B = [L]_{TS2} - (\sum \Theta_i)_R [M]_{TR} - (\sum \Theta_i)_S [M]_{TS2} \quad (2.29)$$

8. Application of the MCT Method to the Base Specificity Problem in ASFV Pol X–ssDNA System

We describe the application of the MCT method to examine the base specificity in the interactions of the African swine fever virus polymerase X with the ssDNA (Jezewska *et al.*, 2007). ASFV Pol X is a small DNA repair polymerase involved in maintaining the genome stability of the virus (Jezewska *et al.*, 2007; Maciejewski *et al.*, 2001; Showalter *et al.*, 2003). Binding the enzyme to the etheno-derivative of the homoadenosine nucleic acids is accompanied by a large nucleic acid fluorescence increase. Unfortunately, association of the polymerase with unmodified ssDNA is not accompanied by an adequate fluorescence change of the enzyme. Thus, determination of the ASFV pol X affinity for the unmodified ssDNA oligomers has been performed using the MCT method. In these studies, the fluorescent 16-mer, dεA(pεA)₁₅, has been used as a reference lattice and competition studies are performed with the unmodified ssDNA 20-mers. Both the 16- and 20-mer can only accommodate a single pol X molecule, what greatly simplifies the analysis (Jezewska *et al.*, 2007). Thus, the entire binding system is composed of two short ssDNA lattices, reference and unmodified oligomer, competing for ASFV pol X. The partition function, Z_R , for the reference 16- and unmodified 20-mer are then

$$Z_R = 1 + K_{16} [L]_F \quad (2.30)$$

and

$$Z_S = 1 + K_{16} [L]_F \quad (2.31)$$

where $[L]_F$ is the free ASFV pol X concentration, K_{16} and K_{20} are the macroscopic binding constant for ASFV pol X binding to the 16- and 20-mer, respectively. In the absence and the presence of the competing 20-mer, the total concentrations of the ASFV pol X, which

provide the same value of the fluorescence change of the reference fluorescent 16-mer, are defined by Eqs. (2.18) and (2.20). Using the partition functions (Eqs. (2.30) and (2.31)) one has

$$[L]_{TR} = [L]_F + \left(\frac{K_{16}[L]_F}{1 + K_{16}[L]_F} \right) [M]_{TR} \quad (2.32)$$

and

$$[L]_{TS} = [L]_F + \left(\frac{K_{16}[L]_F}{1 + K_{16}[L]_F} \right) [M]_{TR} + \left(\frac{K_{20}[L]_F}{1 + K_{20}[L]_F} \right) [M]_{TS} \quad (2.33)$$

Therefore, the total average degree of binding of ASFV pol X on the unmodified 20-mer is

$$\left(\sum \Theta_i \right)_s = \frac{K_{20}[L]_F}{1 + K_{20}[L]_F} = \frac{[L]_{TS} - [L]_{TR}}{[M]_{TS}} \quad (2.34)$$

The concentration of the bound ASFV pol X on both nucleic acids is

$$[L]_B = \left(\frac{K_{16}[L]_F}{1 + K_{16}[L]_F} \right) [M]_{TR} + \left(\frac{K_{20}[L]_F}{1 + K_{20}[L]_F} \right) [M]_{TS} \quad (2.35)$$

and the concentration of the free ASFV pol X is in the presence of unmodified 20-mer is

$$[L]_F = [L]_{TS} - [L]_B \quad (2.36)$$

where $[L]_{TS}$ is the total concentration of the ASFV pol X, $[M]_{TR}$ and $[M]_{TS}$ are the total concentrations of the fluorescent reference 16- and the unmodified ssDNA 20-mer, respectively. However, because the system is very simple, we can directly express the observed fluorescence change of the reference 16-mer, ΔF_{obs} , as a function of the ASFV free pol X concentration, as

$$\Delta F_{obs} = \Delta F_{max} \left(\frac{K_{16}[L]_F}{1 + K_{16}[L]_F} \right) \quad (2.37)$$

with $[L]_F$ is defined by Eqs. (2.35) and (2.36).

Fluorescence titrations of $d\epsilon A(p\epsilon A)_{15}$, with the ASFV pol X, in the presence of two different concentrations of the 20-mer, $dA(pA)_{19}$, are shown in Fig. 2.6A. For comparison, the titration curve in the absence of the competing unmodified 16-mer is also included. Analogous fluorescence titrations of $d\epsilon A(p\epsilon A)_{15}$, with pol X in the presence of two different concentrations of the 20-mer, $dT(pT)_{19}$, are shown in Fig. 2.6B. A large shift of the binding

isotherms, in the presence of $dA(pA)_{19}$ or $dT(pT)_{19}$, indicates an efficient competition between $d\epsilon A(p\epsilon A)_{15}$ and the 20-mers for pol X. The solid lines in Fig. 2.6A and B are nonlinear least squares fits of the experimental titration curves, with a single fitting parameter, K_{20} , using Eqs. (2.30)–(2.37) and $K_{16} = 4.1 \times 10^5 M^{-1}$, independently determined for the reference 16-mer (Jezewska *et al.*, 2007). The obtained binding constants are $K_{20} = (3.0 \pm 0.9) \times 10^7 M^{-1}$ and $(8.0 \pm 1.8) \times 10^6 M^{-1}$ for $dA(pA)_{19}$ and $dT(pT)_{19}$, respectively. Analogous studies with $dC(pC)_{19}$ provided $K_{20} = (3.0 \pm 0.6) \times 10^7 M^{-1}$. Thus, the ASFV pol X does not have a preference for the type of base in the natural ssDNAs (Jezewska *et al.*, 2007).

9. Application of MCT Method to Protein–ssDNA Lattice Binding Systems

The major difference between the analyses discussed above and the ligand binding to long polymer nucleic acids is the replacement of the total average degree of binding, $\Sigma\theta_i$, by the total average binding density, Σv_i (Bujalowski, 2006; Jezewska and Bujalowski, 1996). In other words, one expresses the extent of binding not as the number of moles of the ligand molecules per mole of the macromolecule, but per mole of the monomer of the lattice. Moreover, because more than one ligand molecule bind to the nucleic acid lattice, any statistical thermodynamic binding model must include cooperative interactions. To illustrate the general behavior of such a titration analysis, a series of theoretical spectroscopic 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, unmodified nucleic acid, is shown in Fig. 2.7. Both the reference and the unmodified nucleic acids are treated as long polymers. The titration curves 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 (Bujalowski, 2006; Bujalowski and Lohman, 1986; Crothers, 1968; Epstein, 1978; Jezewska and Bujalowski, 1996; Lohman and Bujalowski, 1991; Lohman and Ferrari, 1994; McGhee and von Hippel, 1974). To simplify the considerations, the protein–nucleic acid complexes for both polymers have the site-size of $n = 20$, cooperativity parameter $\omega = 10$. The intrinsic binding constants are $K = 10^5 M^{-1}$ and $K = 10^6 M^{-1}$ for the reference fluorescent and the nonfluorescent, competing nucleic acid, respectively. Nevertheless, the discussed thermodynamic analysis is independent of any particular binding model for both polymers.

Recall, the measured relative fluorescence increase, ΔF_{obs} , exclusively monitors the saturation of the reference fluorescence nucleic acid with the ligand. Therefore, all curves span the same range of the relative fluorescence change and tend to the same plateau value at selected $\Delta S_{\text{max}} = 3.5$. At very high protein concentrations, all curves are superimposed, because the concentration of the bound protein to both lattices becomes negligibly small, when compared to the total ligand concentration, $[L]_T$. The same value of the relative fluorescence change of the reference nucleic acid, in the presence of the protein, means the same protein binding density, $(\Sigma v_i)_R$ on the reference lattice and the same free protein ligand concentration, $[L]_F$, independent of the presence of the competing unmodified nucleic acid. Thus, in the presence of different concentrations of a competing nucleic acid, and the constant concentration of the reference fluorescence lattice, at a given value of ΔF_{obs} , the concentration of the free protein ligand, $[L]_F$, must be the same and independent of the total concentration of the competing nucleic acid, N_{TS} (dashed line in Fig. 2.7). Hence, one can obtain rigorous measurements of the protein total average binding density, $(\Sigma v_i)_S$, 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, $[M]_{TR}$, with the protein in the presence of two or more concentrations of the

nonfluorescent, competing nucleic acid lattice (Bujalowski, 2006; Jezewska and Bujalowski, 1996).

The experiment provides a set of mass conservation equations for the total protein ligand concentration in solution. In the presence of two different competing nucleic acid concentrations, $[M]_{TS1}$ and $[M]_{TS2}$, the total protein ligand concentrations, $[L]_{T1}$ and $[L]_{T2}$, at which the same relative fluorescence change, ΔF_{obs} , is observed, are defined as

$$[L]_{T1} = (\sum v_i)_R [M]_{TR} + (\sum v_i)_S [M]_{TS1} + [L]_F \quad (2.38a)$$

and

$$[L]_{T2} = (\sum v_i)_R [M]_{TR} + (\sum v_i)_S [M]_{TS2} + [L]_F \quad (2.38b)$$

Subtracting Eq. (2.38a) from Eq. (2.38b) and rearranging provides the expression for the total average binding density of the protein on the competing, unmodified lattice, $(\sum v_i)_S$, and the free protein concentration, $[L]_F$, in terms of known quantities, $(\sum v_i)_R$, total protein, and nucleic acid concentrations as (Bujalowski, 2006; Jezewska and Bujalowski, 1996)

$$(\sum v_i)_S = \frac{[L]_{T2} - [L]_{T1}}{[M]_{TS2} - [M]_{TS1}} \quad (2.39a)$$

And

$$[L]_F = [L]_{Tx} - (\sum v_i)_S [M]_{TSx} - (\sum v_i)_R [M]_{TR} \quad (2.39b)$$

where x is 1 or 2. Notice, the binding density, $(\sum v_i)_R$, of the protein on the reference fluorescent nucleic acid, at any value of ΔF_{obs} , can be determined in an independent titration experiment. Calculations of $(\sum v_i)_S$ and, subsequently, $[L]_F$ can be performed at any value of the observed fluorescence change, ΔF_{obs} , along the titration curves (Fig. 2.7) generating a thermodynamic binding isotherm for protein binding to the competing unmodified nucleic acid. The approach is demonstrated in Section 10, using the experimental data for the binding of the *Escherichia 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 are different for both nucleic acids, at the same value of the measured relative fluorescence change, ΔF_{obs} . Figure 2.8 shows the theoretical dependence of the observed relative fluorescence change, ΔS_{obs} upon the total average binding density of the protein on the competing unmodified nucleic acid, 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 $K = 10^5 M^{-1}$, $\omega = 1$, and $\Delta F_{max} = 3.5$; the ligand binding to the competing lattice is characterized by $K = 10^6 M^{-1}$, $\omega = 1$, and $K = 10^4 M^{-1}$, $\omega = 1$, for high and low affinity cases, respectively. For simplicity, the dependence of ΔF_{obs} upon $(\sum v_i)_R$ for

the reference lattice was selected to be strictly proportional (straight, dashed line). In the case where the macroscopic ligand affinity for the competing lattice is higher than for the reference lattice, the total average binding density, $(\Sigma v_i)_S$, of the ligand on the competing, nonfluorescent lattice is higher, when compared to the reference fluorescent lattice at any value of ΔF_{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 $(\Sigma v_i)_S$, gradually approaching the maximum possible value of ΔF_{obs} . This behavior results from the fact that ΔF_{obs} solely reflects the binding density on the reference lattice, $(\Sigma v_i)_R$, which, due to the higher affinity of the ligand for the reference nucleic acid, saturates with the ligand in advance of the saturation of the competing unmodified nucleic acid. The plots in Fig. 2.8 indicate that, in order to obtain the most accurate estimate of the stoichiometry of protein–unmodified nucleic acid complex, the affinity of the competing nonfluorescent nucleic acid for the protein should be similar or higher than that of the fluorescent reference lattice.

10. Quantitative Analysis of the Binding of the *E. coli* DnaB Helicase to Unmodified Nucleic Acids Using the MCT Method

The *E. coli* DnaB protein is its primary replicative helicase, that is, the factor that unwinds the duplex DNA in front of the replication fork (Jezewska and Bujalowski, 1996). Binding of the enzyme to unmodified nucleic acids, for example, poly(dA), and other polydeoxynucleotides does not cause any significant change in the protein fluorescence (Bujalowski and Jezewska, 1995; Jezewska and Bujalowski, 1996). 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 estimate of the stoichiometry and interaction parameters of the DnaB protein–poly(dεA) complex (Bujalowski and Jezewska, 1995; Jezewska and Bujalowski, 1996). Thus, poly(dεA) can serve as a reference fluorescent nucleic acid in the MCT method.

Figure 2.9 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 significant shift of the titration curves toward higher total $[\text{DnaB}]_T$, in the presence of poly(dA), indicates efficient competition between the two nucleic acids for the enzyme. At the same value of the relative fluorescence increase, the total average binding density, $(\Sigma v_i)_S$, on the competing lattice, poly(dA), and the free DnaB concentration, $[\text{DnaB}]_F$, are the same, independent of the concentration of poly(dA) (Eqs. (2.38a), (2.38b), (2.39a), (2.39b)). Therefore, one can obtain a set of total DnaB concentrations, $[\text{DnaB}]_{T1}$ and $[\text{DnaB}]_{T2}$, which are determined by the intersection of the horizontal line with both titration curves at the same value of the ΔF_{obs} (e.g., horizontal line in Fig. 2.9). Since the total concentrations of poly(dA) are known, one can calculate the true binding density, $(\Sigma v_i)_S$, of the DnaB protein on poly(dA) and the free DnaB concentration, $[\text{DnaB}]_F$, using Eqs. (2.39a) and (2.39b). The same procedure is then repeated over the entire range of ΔF_{obs} , for selected intervals of ΔF_{obs} , providing $(\Sigma v_i)_S$ as a function of $[\text{DnaB}]_F$. In other words, one obtains the thermodynamically rigorous binding isotherm for DnaB helicase association with poly(dA), although the signal used to monitor the binding originates exclusively from the reference fluorescent lattice, poly(dεA).

It should be 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 a reference fluorescent nucleic acid, due to the same physical nature of both nucleic acids (ssDNA homopolymers). Nevertheless, for completeness, we include the independent determination of this quantity in our discussion. Figure 2.10 shows the dependence of the observed relative fluorescence change, ΔF_{obs} , upon the total average binding density, $(\Sigma v_i)_S$, of the DnaB helicase on poly(dA). For

comparison, the dependence of ΔF_{obs} upon the total average binding density, $(\Sigma v_i)_R$, of the DnaB helicase on the reference lattice, poly(dεA), is also included (dashed straight line). The quantity, $(\Sigma v_i)_S$, could be determined up to the value of ~ 0.044 , which corresponds with $\Delta F_{\text{obs}} \sim 2.6$. Extrapolation to the maximum possible value of $\Delta F_{\text{max}} = 3.6 \pm 0.3$ gives the maximum value of $(\Sigma v_i)_S = 0.05 \pm 0.005$ and the estimation of the site-size poly(dA)–DnaB helicase complex, $\underline{n} = 20 \pm 3$. As expected, this value is the same as the estimated $n = 20 \pm 3$ in identical solution conditions for the poly(dεA)–DnaB complex (Jezewska and Bujalowski, 1996).

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

For binding of a large protein ligand to a long homogeneous nucleic acid lattice, the McGhee–von Hippel model is the simplest statistical thermodynamic description for the binding process (Bujalowski and Lohman, 1987; Bujalowski *et al.*, 1989; Epstein, 1978; Jezewska and Bujalowski, 1996; McGhee and von Hippel, 1974). The lattice site can only exist in two states, free and bound with the ligand in a single type of the complex. The model takes into account the cooperativity of the protein association and the overlap of the potential binding sites. Although the original work provides two expressions for the noncooperative and cooperative binding, a single, generalized equation for the McGhee–von Hippel model, which can be applied to both cooperative and noncooperative binding systems, has been derived and is defined by Bujalowski (2006) and Bujalowski *et al.*, (1989).

$$\sum_{v_i=K(1-n \sum v_i)} \left\{ \frac{[2\omega(1-n \sum v_i)]}{[(2\omega-1)(1-n \sum v_i) + \sum v_i + R]} \right\}^{(n-1)} \left\{ \frac{[1-(n+1)\sum v_i + R]}{2(1-n \sum v_i)} \right\}^2 [L]_F \quad (2.40)$$

where K is the intrinsic binding constant, n is the site-size of the protein–nucleic acid complex (number of bases or base pairs occluded by the bound protein in the complex), ω is the cooperative interactions parameter, and $R = \{[1-(n+1)\sum v_i]^2 + 4\omega \sum v_i(1-n \sum v_i)\}^{0.5}$.

Though it is a simple statistical thermodynamic model, Eq. (2.40) is a formidable expression. In the case of protein binding to two competing and long nucleic acid lattices, the ligand interactions with each nucleic acid are described by an equation in the form of expression (2.40). An attempt to simultaneously use two isotherms, of the type described by Eq. (2.40), is hindered by the fact that they are complex, polynomial implicit functions of the total average binding density, Σv_i , and free ligand concentration. Thus, in order 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.

It should be stressed that the McGhee–von Hippel model is equivalent to the combinatorial model as the length of the homogenous lattice approaches infinity (Bujalowski *et al.*, 1989; Epstein, 1978; McGhee and von Hippel, 1974). To overcome the problem of resorting to complex numerical calculations and to obtain a general method to analyze simultaneous binding of a protein ligand to two or more polymer nucleic acids, one can apply a combined application of the generalized McGhee–von Hippel model, as defined by Eq. (2.40), and the combinatorial theory for large ligand binding to a linear, homogeneous lattice (Bujalowski, 2006; Jezewska and Bujalowski, 1996). In this approach, the ligand binding to the reference fluorescent nucleic acid is described by the generalized McGhee–von Hippel equation. Binding of the protein to the competing nonfluorescent lattice is described by the combinatorial theory for cooperative binding of a large ligand to a linear finite homogeneous

nucleic acid. For a long enough lattice, the isotherm generated using the combinatorial theory, will be, within experimental accuracy, indistinguishable from the isotherm generated, using the generalized Eq. (2.40). For instance, in the case of a protein like the DnaB helicase ($n = 20$), a lattice, which can accommodate ≥ 40 protein molecules (800 nucleotides), represents an “infinite” lattice for any practical purpose.

In the combinatorial theory, the partition function of the protein ligand–nucleic acid system, Z_S , is defined by

$$Z_S = \sum_{k=0}^g \sum_{j=0}^{k-1} S_N(k, j) (K_S [L]_F)^k \omega_S^j \quad (2.41)$$

where g is the maximum number of ligand molecules which may bind to the finite nucleic acid lattice (for the nucleic acid lattice N residues long $g = M/n$), K_S is the intrinsic binding constant characterizing interactions with the unmodified lattice, ω 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

$$S_N(k, j) = \frac{[(M - nk + 1)(k - 1)!]}{[(M - nk + j + 1)(k - j)! j! (k - j - 1)!]} \quad (2.42)$$

The total average binding density, $(\Sigma v_i)_S$ is then obtained by using the standard statistical thermodynamic expression, $(\Sigma v_i)_S = \partial \ln Z_S / \partial \ln L_F$, as Fried and Crothers (1981), Jezewska and Bujalowski (1996), and Bujalowski (2006).

$$(\Sigma v_i)_S = \frac{\sum_{k=1}^g \sum_{j=0}^{k-1} k S_N(k, j) (K_S [L]_F)^k \omega_{S,j}}{\sum_{k=0}^g \sum_{j=0}^{k-1} S_N(k, j) (K_S [L]_F)^k \omega_{S,j}} \quad (2.43)$$

Expressions (2.41)–(2.43) describe the binding of a large ligand to a finite, linear homogeneous nucleic acid. Contrary to the generalized McGhee–von Hippel model (Eq. (2.40)), expression (2.43) 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. (2.40), 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. (2.40) to the reference fluorescent nucleic acid, calculating the free ligand concentration, L_F , for given values of the parameters K , ω , and n , with $(\Sigma v_i)_R$ as the independent variable. Subsequently, the obtained value of $[L]_F$ is introduced into Eq. (2.43) and the binding density, $(\Sigma v_i)_S$, 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 $(\Sigma v_i)_R$, generating the required

$(\Sigma v_i)_S$ for the competing nucleic acid lattice, as a function of $[L]_F$. The experimental binding isotherm, which is the observed relative fluorescence change ΔF_{obs} as a function of the total protein concentration $[L]_T$, is then obtained by calculating for each value of ΔF_{obs} the total protein concentration $[L]_T$, by introducing $(\Sigma v_i)_S$, $(\Sigma v_i)_R$, and $[L]_F$ into expression

$$[L]_T = [L]_F - \left(\sum v_i \right)_R [M]_{\text{TR}} - \left(\sum v_i \right)_S [M]_{\text{TS}} \quad (2.44)$$

The solid lines in Fig. 2.9 are nonlinear least squares fits of the simultaneous binding of the DnaB helicase to two competing nucleic acids, using the procedure outlined above (Jezewska and Bujalowski, 1996). 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. (2.40)). The binding of the protein to the competing poly(dA) has been described, using the combinatorial theory (Eqs. (2.31)–(2.43)). Because the site-size of the DnaB protein–poly(dεA) complex, $n = 20 \pm 3$, has been independently estimated, there are only two parameters, the intrinsic binding constant, K_S , and the cooperativity parameter, ω_S , to be determined (Jezewska and Bujalowski, 1996). It should be pointed out that the analysis of simultaneous binding of a large ligand to competing nucleic acid lattices could also be performed by applying Eqs. (2.31)–(2.43) of the combinatorial theory, to both the reference and the competing nucleic acids. This is a completely equivalent approach to the combined application of the McGhee–von Hippel and combinatorial models. However, the major 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.

12. Using a Single Concentration of a Nonfluorescent Unmodified Nucleic Acid

Using several concentrations of the unmodified nucleic acid increases the accuracy of the binding experiment. However, it may be very time consuming and often impractical, particularly if the availability of the specific nucleic acid substrate is limited. However, the thermodynamic binding density, $(\Sigma v_i)_S$, 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 (Jezewska and Bujalowski, 1996). In Fig. 2.9, one can obtain $(\Sigma v_i)_S$ and $[DnaB]_F$ by simultaneously analyzing titration curves 1 and 2 or 1 and 3, instead of 2 and 3. 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, ΔF_{obs} , of the reference fluorescent nucleic acid is observed, in the absence and presence of competing nonfluorescent lattice, $[L]_{\text{TR}}$, and $[L]_{\text{TSx}}$, respectively, are defined as

$$[L]_{\text{TR}} = \left(\sum v_i \right)_R [M]_{\text{TR}} + [L]_F \quad (2.45a)$$

and

$$[L]_{\text{TS}} = \left(\sum v_i \right)_R [M]_{\text{TR}} + \left(\sum v_i \right)_S [M]_{\text{TS}} + [L]_F \quad (2.45b)$$

Solving the set of Eqs. (2.45a) and (2.45b), for $(\sum v_i)_S$ and $[L]_F$, provides

$$(\sum v_i)_S = [L]_{TS} - \left[\frac{[L]_{TR}}{[M]_{TS}} \right] \quad (2.46a)$$

and

$$[L]_F = [L]_{TS} - (\sum v_i)_R [M]_{TR} - (\sum v_i)_S [M]_{TS} \quad (2.46b)$$

13. Using Short Fluorescent Oligonucleotides in Competition with the Polymer Nucleic Acid

As long as two chemical species compete for the same binding site, they can be used in the competition titration experiments. In the case of the protein–nucleic acid interactions, the MCT analysis can be simplified by using short fragments of a nucleic acid, as a reference fluorescent lattice, in competition titrations with various DNA substrates (forks, template–primers, gapped substrates, long nucleic acid polymers, etc.), which share the same binding site on the ligand (protein). The optimal reference choice is a short nucleic acid fragment, which forms a simple 1:1 complex with the protein and is long enough to exactly span the total site-size of the protein–nucleic acid complex. A selection of the length of the fluorescent reference fragment can be based on the initial estimate of the stoichiometry of the protein–nucleic acid complex obtained using a modified fluorescent polymer nucleic acid or a series of fluorescent oligomers (Bujalowski, 2006; Jezewska and Bujalowski, 1996). In such studies, the degree of saturation of the nucleic acid oligomer with the protein, $\sum \Theta_R$, and the fluorescence change, ΔF_{obs} , accompanying the formation of the complex, are described by

$$(\sum \Theta_i)_R = \frac{K_o [L]_F}{1 + K_o [L]_F} \quad (2.47a)$$

and

$$\Delta F_{obs} = \Delta F_{max} \left(\frac{K_o [L]_F}{1 + K_o [L]_F} \right) \quad (2.47b)$$

where ΔF_{max} 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 protein for the reference fluorescent oligomer. Therefore, in the presence of for example, a competing, unmodified polymer nucleic acid, the observed fluorescence change of the short nucleic acid fragment is thermodynamically linked with the total average binding density, $(\sum v_i)_S$, on the competing polymer through the free protein ligand concentration, $[L]_F$, as defined by

$$[L]_F = [L]_{TS} - \left(\sum_R \Theta_i \right) [O]_{TR} - \left(\sum_S v_i \right) [M]_{TS} \quad (2.48)$$

In practice, two experimental situations usually arise. The site-size of the protein on the polymer nucleic acid, or the maximum stoichiometry on a specific DNA substrate is known and the experimenter wants to examine the intrinsic affinities and cooperativities in some specific solution conditions. The competition titration curve, which is ΔF_{obs} as a function of the total protein concentration, $[L]_{TS}$, in the presence of the competing nucleic acid, is then fitted to Eq. 2.47b), with $[L]_F$ treated as independent variable. The values of $[L]_{TS}$ are then calculated using Eq. (2.48) and the selected statistical thermodynamic binding model for $(\Sigma v_i)_S$. However, if the binding to the competing unmodified nucleic acid is unknown, then a complete analysis is in place. As in the case of the reference polymer nucleic acid, the same value of ΔF_{obs} at different competing 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 $(\Sigma v_i)_S$. Therefore, performing two fluorescence titrations of a short reference 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, $[N]_{TS1}$ and $[N]_{TS2}$, allows the experimenter to obtain the total average binding density, $(\Sigma v_i)_S$, and free protein concentration, $[L]_F$, analogously, as described above for a polymer reference lattice. Expressions (2.38a) and (2.38b) are

$$[L]_{T1} = \left(\sum_R \Theta_i \right) [O]_{TR} + \left(\sum_S v_i \right) [M]_{TS1} + [L]_F \quad (2.49a)$$

and

$$[L]_{T2} = \left(\sum_R \Theta_i \right) [O]_{TR} + \left(\sum_S v_i \right) [M]_{TS2} + [L]_F \quad (2.49b)$$

The total average binding density, $(\Sigma v_i)_S$, and the free protein concentration are described by Eqs. (2.39a) and (2.39b). Analogously, if the reference fluorescent oligomer is titrated with the protein in the presence of a single concentration of the competing, long polymer nucleic acid at the concentration $[N]_{TS}$, the total concentrations of the protein ligand in the absence and presence of the competing nucleic acid and at the same value of the fluorescence intensity, ΔF_{obs} , are (Jezewska and Bujalowski, 1996)

$$[L]_{TR} = \left(\sum_R \Theta_i \right) [O]_{TR} + [L]_F \quad (2.50a)$$

and

$$[L]_{TS} = \left(\sum_R \Theta_i \right) [O]_{TR} + \left(\sum_S v_i \right) [M]_{TS} + [L]_F \quad (2.50b)$$

The total average binding density, $(\Sigma v_i)_S$, obtained by subtracting Eq. (2.50a) from Eq. (2.50b), and the free protein concentration are described by (Eq. (2.46a) and (2.46b))

$$\left(\sum v_i\right)_S = \frac{[L]_{TS} - [L]_{TR}}{[M]_{TS}} \quad (2.51a)$$

and

$$[L]_F = [L]_{TS} - \left(\sum \Theta_i\right)_R [O]_{TR} - \left(\sum v_i\right)_S [M]_{TS} \quad (2.51b)$$

14. Conclusions

The MCT method enables one to obtain, rigorously and quantitatively, interaction parameters of protein–nucleic acid interactions, which may not be available by other methods, particularly for the unmodified long polymer lattices and specific DNA substrates, if the binding is not accompanied by adequate spectroscopic signal changes. Nevertheless, the method is not limited by the size of the reference nucleic acid. A simple analysis of the competition titration experiments can be performed in which the fluorescent, short fragment of nucleic acid, spanning the exact site-size of the protein–nucleic acid complex, and binding with only a 1:1 stoichiometry to the protein, is used as a reference macromolecule. We have discussed the MCT method, predominantly, as applied to studying protein–nucleic acid interactions. However, the method can generally be applied to any ligand–macromolecule system, by monitoring the binding using the spectroscopic signal originating from a reference macromolecule, in the presence of the competing macromolecule whose interaction parameters with the ligand are to be determined, and whose spectroscopic properties do not change upon complex formation with the ligand.

Acknowledgments

We thank Gloria Drennan Bellard for reading the chapter. This work was supported by NIH Grants GM46679 and GM58565 (to W. B.).

References

- Baker BM, Vanderkooi J, Kallenbach NR. Base stacking in A fluorescent dinucleoside monophosphate: ϵ ApeA. *Biopolymers*. 1978; 17:1361.
- Boschelli FJ. Lambda phage cro repressor Non-specific DNA binding. *Mol Biol*. 1982; 162:267.
- Brenowitz M, Senear DF, Shea MA, Ackers GK. *Methods Enzymol*. 1986; 130:132. [PubMed: 3773731]
- Bujalowski W. *Chem Rev*. 2006; 106:556. [PubMed: 16464018]
- Bujalowski W, Jezewska MJ. *Biochemistry*. 1995; 34:8513. [PubMed: 7612593]
- Bujalowski, W.; Jezewska, MJ. *Spectrophotometry and Spectrofluorimetry: A Practical Approach*. Gore, MG., editor. Vol. Chapter 5. IRL Press (Oxford University Press); Oxford, UK: 2000. p. 141-165.
- Bujalowski W, Jezewska MJ. *Methods Enzymol*. 2009; 466:294.
- Bujalowski W, Klonowska MM. *Biochemistry*. 1993; 32:5888. [PubMed: 8504109]
- Bujalowski W, Klonowska MM. *Biochemistry*. 1994a; 33:4682. [PubMed: 8161526]
- Bujalowski W, Klonowska MM. *J Biol Chem*. 1994b; 269:31359. [PubMed: 7989300]
- Bujalowski W, Lohman TM. *Biochemistry*. 1986; 25:7779.
- Bujalowski W, Lohman TM. *Biochemistry*. 1987; 26:3099. [PubMed: 3300771]
- Bujalowski W, Lohman TM, Anderson CF. *Biopolymers*. 1989; 28:1637. [PubMed: 2775853]

- Bujalowski W, Klonowska MM, Jezewska MJ. J Biol Chem. 1994; 269:31350. [PubMed: 7989299]
- Chabbert M, Cazenave C, Hélène C. Biochemistry. 1987; 26:2218. [PubMed: 3304422]
- Crothers DM. Biopolymers. 1968; 6:575. [PubMed: 5644787]
- deHaseth PL, Gross CA, Burgess RR, Record MT Jr. Biochemistry. 1977; 16:4777. [PubMed: 911788]
- Draper DE, von Hippel PH. J Mol Biol. 1978; 122:321. [PubMed: 357731]
- Epstein IR. Biophys Chem. 1978; 8:327. [PubMed: 728537]
- Fernando T, Royer C. Biochemistry. 1992; 31:3429. [PubMed: 1554725]
- Fried M, Crothers DM. Nucleic Acids Res. 1981; 9:6505. [PubMed: 6275366]
- Garner MM, Revzin A. Nucleic Acids Res. 1981; 9:3047. [PubMed: 6269071]
- Heyduk T, Lee JC. Proc Natl Acad Sci USA. 1990; 87:1744. [PubMed: 2155424]
- Jezewska MJ, Bujalowski W. Biochemistry. 1996; 35:2117. [PubMed: 8652554]
- Jezewska MJ, Bujalowski W. Biophys Chem. 1997; 64:253. [PubMed: 9127949]
- Jezewska MJ, Kim US, Bujalowski W. Biochemistry. 1996; 35:2129. [PubMed: 8652555]
- Jezewska MJ, Rajendran S, Bujalowski W. Biochemistry. 1998a; 37:3116. [PubMed: 9485465]
- Jezewska MJ, Rajendran S, Bujalowski W. J Mol Biol. 1998b; 284:1113. [PubMed: 9837730]
- Jezewska MJ, Rajendran S, Bujalowski W. Biochemistry. 2001; 40:3295. [PubMed: 11258949]
- Jezewska MJ, Galletto R, Bujalowski W. Biochemistry. 2003; 42:5955. [PubMed: 12741854]
- Jezewska MJ, Lucius AL, Bujalowski W. Biochemistry. 2005a; 44:3865. [PubMed: 15751962]
- Jezewska MJ, Lucius AL, Bujalowski W. Biochemistry. 2005b; 44:3877. [PubMed: 15751963]
- Jezewska MJ, Marcinowicz A, Lucius AL, Bujalowski W. J Mol Biol. 2006; 356:121. [PubMed: 16337650]
- Jezewska MJ, Bujalowski PJ, Bujalowski W. Biochemistry. 2007; 46:12909. [PubMed: 17941646]
- Kowalczykowski SC, Lonberg N, Newport JW, von Hippel PH. J Mol Biol. 1981; 145:75. [PubMed: 7265204]
- Kowalczykowski SC, Paul LS, Lonberg N, Newport JW, McSwiggen JA, von Hippel PH. Biochemistry. 1986; 25:1226. [PubMed: 3486003]
- Lohman TM, Bujalowski W. Methods Enzymol. 1991; 208:258. [PubMed: 1779838]
- Lohman TM, Ferrari ME. Annu Rev Biochem. 1994; 63:527. [PubMed: 7979247]
- Maciejewski M, Shin R, Pan B, Marintchev A, Denninger A, Mullen MA, Chen K, Gryk MR, Mullen GP. Nat Struct Biol. 2001; 8:936. [PubMed: 11685238]
- McGhee JD, von Hippel PH. J Mol Biol. 1974; 86:469. [PubMed: 4416620]
- McSwiggen JA, Bear DG, von Hippel PH. J Mol Biol. 1988; 199:609. [PubMed: 2451028]
- Menetski JP, Kowalczykowski SC. J Mol Biol. 1985; 181:281. [PubMed: 3981638]
- Porschke D, Rauh H. Biochemistry. 1983; 22:4737. [PubMed: 6354266]
- Rajendran S, Jezewska MJ, Bujalowski W. J Biol Chem. 1998; 273:31021. [PubMed: 9813000]
- Rajendran S, Jezewska MJ, Bujalowski W. J Mol Biol. 2001; 308:477. [PubMed: 11327782]
- Revzin A, von Hippel PH. Biochemistry. 1977; 16:4769. [PubMed: 20938]
- Scatchard G. Ann NY Acad Sci. 1949; 51:660.
- Secrist JA, Barrio JR, Leonard NJ, Weber G. Biochemistry. 1972; 11:3499. [PubMed: 4340904]
- Showalter AK, Byeon IJ, Su MI, Tsai MD. Nat Struct Biol. 2003; 8:942.
- Szymanski MR, Jezewska MJ, Bujalowski W. J Mol Biol. 2010; 398:8. [PubMed: 20156448]
- Tolman GL, Barrio JR, Leonard NJ. Biochemistry. 1974; 13:4869. [PubMed: 4373039]

Figure 2.1.

(A) The dashed line is the theoretical fluorescence titration of a macromolecule that has six discrete identical binding sites with a ligand at the macromolecule concentration $[M]_T = 1 \times 10^{-7} M$. The intrinsic binding constant, characterizing the binding, $K = 1 \times 10^6 M^{-1}$. Association of the ligand induces at each binding site the relative change of the macromolecule fluorescence, $\Delta F = 0.5$, providing the maximum change of the macromolecule fluorescence at saturation, $\Delta F_{\max} = 3.0$. The solid line is the same titration curve but approximated using a single-site binding titration curve, with the binding constant $K = 8 \times 10^5 M^{-1}$ and $\Delta F_{\max} = 3.0$. (B) The dashed line is the theoretical fluorescence titration of a macromolecule, which has six discrete, identical binding sites, characterized by the intrinsic binding constants, $K = 1 \times 10^6 M^{-1}$, with a ligand. The macromolecule concentration, $[M]_T = 1 \times 10^{-6} M$. The binding of the ligand induces the relative change of the macromolecule fluorescence, at each binding site, $\Delta F = 0.5$, providing the maximum change of the macromolecule fluorescence at saturation, $\Delta F_{\max} = 3.0$. The solid line is the same titration curve but approximated using a single-site binding titration curve with the binding constant $K = 2.8 \times 10^5 M^{-1}$ and $\Delta F_{\max} = 3.0$. The line (---) is the approximation of the titration curve using a simple binding model of two cooperative binding sites with the intrinsic binding constant, $K = 1.8 \times 10^5 M^{-1}$, the cooperativity parameter, $\sigma = 4$, and the relative fluorescence change accompanying the binding to each site $\Delta F = 1.5$.

Figure 2.2.

(A) Theoretical fluorescence titrations of a macromolecule that has six independent and discrete binding sites with a ligand obtained at two different macro-molecule concentrations, $[M]_{T1} = 3 \times 10^{-7} M$, and $[M]_{T2} = 3 \times 10^{-6} M$, respectively. The binding sites are characterized by the intrinsic binding constants, $K = 1 \times 10^5 M^{-1}$. Association of the ligand induces at each binding site the relative change of the macromolecule fluorescence, $\Delta F = 0.5$, providing the maximum change of the macromolecule fluorescence at saturation, $\Delta F_{\max} = 3.0$. The arrows indicate the total ligand concentrations, $[L]_{T1}$ and $[L]_{T2}$, at the same selected value of the observed fluorescence change, marked by the dashed horizontal line, ΔF_{obs} , at which the total average degree of binding of the ligand, $\Sigma\Theta_i$, is the same for both titration curves. (B) Dependence of the relative fluorescence of the macromolecule, ΔF_{obs} , upon the total average degree of binding, $\Sigma\Theta_i$, of the ligand for the binding model in (A). Because the value of $\Delta F = 0.5$ is the same for each binding site, the plot is linear over the entire range of $\Sigma\Theta_i$.

Figure 2.3.

(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: (■) $5 \times 10^{-7} M$; (□) $1 \times 10^{-6} M$; (●) $3 \times 10^{-6} M$ (hexamer). The solid lines are nonlinear least squares fits of the titration curves, according to the hexagon model (Eqs. (2.14)–(2.16)) using a single set of binding parameters with the intrinsic binding constant $K = 8 \times 10^6 M^{-1}$ and cooperativity parameter $\sigma = 0.36$ (Jezewska *et al.*, 2005a,b). (B) Dependence of the relative fluorescence quenching, ΔF_{obs} , upon the average degree of binding of TNP-ADP on the RepA hexamer, $\Sigma\theta_i$, (■). The point (□) and the straight horizontal line indicate the location of the maximum value of the observed ΔF . The line (—) is the linear regression using the last five points of the plot. The solid line is the nonlinear least squares fit, using the second-degree polynomial function representing the empirical function (Jezewska *et al.*, 2005a,b). The dashed line is the hypothetical dependence of ΔF_{obs} upon $\Sigma\theta_i$ that assumes a strict linear relationship between the observed fluorescence quenching and the average degree of binding. The maximum value of $\Delta F_{\max} = 0.87 \pm 0.03$. (C) Fluorescence titration of the RepA helicase with TNP-ATP in 50 mM Tris/HCl (pH 7.6, 10 °C), containing 10 mM NaCl and 1 mM MgCl₂, at different RepA protein concentrations: (■) $5 \times 10^{-7} M$; (□) $1 \times 10^{-6} M$; (●) $6 \times 10^{-6} M$ (hexamer). The solid lines are nonlinear least squares fits of the titration curves, according to the hexagon model (Eqs. (2.14)–(2.16)) using a single set of binding parameters with the intrinsic binding constant $K = 2.38 \times 10^6 M^{-1}$ and cooperativity parameter $\sigma = 0.37$ (Jezewska *et al.*, 2005a,b). (D) Dependence of the relative fluorescence quenching, ΔF_{obs} , upon the average degree of binding of TNP-ADP on the RepA hexamer, $\Sigma\theta_i$, (■). The point (□) and the straight horizontal line indicate the location of the maximum value of the observed ΔF . The line (—) is the linear regression using the last four points of the plot. The solid line is the nonlinear least squares fit, using the second-degree polynomial function representing the empirical function (Jezewska *et al.*, 2005a,b). The dashed line is the theoretical dependence of ΔF_{obs} upon $\Sigma\theta_i$ that assumes a strict linear relationship between the observed fluorescence quenching and the average degree of binding. The maximum value of $\Delta F_{\max} = 0.83 \pm 0.03$.

Figure 2.4.

(A) The average number of TNP-ADP molecules bound per RepA hexamer as a function of the free concentration of the cofactor. The solid line is a nonlinear least squares fit according to the hexagon model (Eqs. (2.14)–(2.16)) with the intrinsic binding constant $K = 7.7 \times 10^6 \text{ M}^{-1}$ and cooperativity parameter $\sigma = 0.38$. (B) The average number of TNP-ATP molecules bound per RepA hexamer as a function of the free concentration of the cofactor. The solid line is a nonlinear least squares fit according to the hexagon model (Eqs. (2.14)–(2.16)) with the intrinsic binding constant $K = 2.38 \times 10^6 \text{ M}^{-1}$ and cooperativity parameter $\sigma = 0.4$ (Jezewska *et al.*, 2005a,b).

Figure 2.5.

Theoretical fluorescence titrations of a fluorescent reference macromolecule with a ligand, at concentration, $[M]_{TR}$, in the absence (solid line) and presence of a competing unmodified macromolecule, at concentration, $[M]_{TS}$ (dashed line). The arrows indicate the total ligand concentrations, $[L]_{TR}$ and $[L]_{TS}$, at the selected value of the observed fluorescence change, marked by the dashed horizontal line, ΔF_{obs} , at which the total average degree of binding of the ligand, $\Sigma \theta_i$, on the reference fluorescence macromolecule and the free ligand concentration, $[L]_F$, are the same for both titration curves.

Figure 2.6.

(A) Fluorescence titrations of the ssDNA 16-mer, $d\epsilon A(p\epsilon A)_{15}$, ($\lambda_{\text{ex}} = 325 \text{ nm}$, $\lambda_{\text{em}} = 410 \text{ nm}$) with the ASFV pol X in 10 mM sodium cacodylate (pH 7.0, 10 °C), containing 100 mM NaCl and 1 mM MgCl_2 , in the absence (■) and the presence of the unmodified ssDNA 20-mer, $dA(pA)_{19}$. The concentration of $d\epsilon A(p\epsilon A)_{15}$ is $1 \times 10^{-6} \text{ M}$. The concentration of unmodified ssDNA oligomer, $dA(pA)_{19}$ are (□) $1 \times 10^{-6} \text{ M}$ (●) $5 \times 10^{-6} \text{ M}$ (oligomer). The solid lines are nonlinear least squares fits of the experimental fluorescence titration curves, using the MCT method (Jezewska *et al.*, 2007). (B) Fluorescence titrations of the ssDNA 16-mer, $d\epsilon A(p\epsilon A)_{15}$, ($\lambda_{\text{ex}} = 325 \text{ nm}$, $\lambda_{\text{em}} = 410 \text{ nm}$) with the ASFV pol X in 10 mM sodium cacodylate (pH 7.0, 10 °C), containing 100 mM NaCl and 1 mM MgCl_2 , in absence (■) and the presence of the unmodified ssDNA 20-mer, $dT(pT)_{19}$. The concentration of $d\epsilon A(p\epsilon A)_{15}$ is $1 \times 10^{-6} \text{ M}$. The concentrations of unmodified ssDNA oligomer, $dT(pT)_{19}$ are: (□) $1 \times 10^{-6} \text{ M}$ (●) $5 \times 10^{-6} \text{ M}$ (oligomer). The solid lines are nonlinear least squares fits of the experimental fluorescence titration curves, using the MCT method (Jezewska *et al.*, 2007).

Figure 2.7.

Theoretical fluorescence titration curves of the reference fluorescent nucleic acid with the protein ligand in the presence of different concentrations of the competing, nonfluorescent nucleic acid lattice. Binding of the ligand to the reference lattice is described by the McGhee–von Hippel model for large ligand binding to a linear, homogeneous nucleic acid, using intrinsic binding constant $K = 10^5 M^{-1}$, cooperativity parameter $\omega = 1$, and site-size $n = 20$. The maximum increase of the nucleic acid fluorescence intensity upon saturation with the ligand is $\Delta F_{\text{max}} = 3.5$ (Jezewska and Bujalowski, 1996). Binding of the ligand to a competing, nonfluorescent nucleic acid is described by the combinatorial theory using intrinsic binding constant $K = 10^5 M^{-1}$, cooperativity parameter $\omega = 1$, and site-size $n = 20$. The selected length of the nucleic acid is 1600 nucleotides. The concentration of the competing nucleic acid (nucleotide) is: (O) 0; (A) $2 \times 10^{-5} M$; (B) $4 \times 10^{-5} M$; (C) $1.2 \times 10^{-4} M$; (D) $2.4 \times 10^{-4} M$. The concentration of the reference fluorescent nucleic acid is $2 \times 10^{-5} M$ (Nucleotide). The horizontal dashed line connects points on all titration curves characterized by the same value of the relative fluorescence increase, ΔF_i .

Figure 2.8.

Computer simulation of the dependence of the relative fluorescence increase of the reference fluorescent nucleic acid, ΔF , upon the binding density, $(\Sigma \nu_i)_S$, of the ligand on the competing, nonfluorescent lattice where the competing, nonfluorescent lattice has higher (— — —) and lower (——) affinity toward the ligand. The binding of the ligand to the reference lattice is described by the McGhee–von Hippel model of large ligand binding to a linear, homogeneous nucleic acid, using intrinsic binding constant $K = 10^5 M^{-1}$, cooperativity parameter $\omega = 1$, and site-size $n = 20$. The maximum increase of the nucleic acid fluorescence intensity upon saturation with the ligand is $\Delta F_{\max} = 3.5$. Binding of the ligand to competing, nonfluorescent nucleic acid is described by the combinatorial theory, using $\omega = 1$ and $n = 20$; $K = 10^4 M^{-1}$ and $10^6 M^{-1}$ for lower and higher affinity cases, respectively. The selected length of the nucleic acid is 1600 nucleotides (Jezewska and Bujalowski, 1996).

Figure 2.9.

Fluorescence titrations ($\lambda_{\text{ex}} = 325 \text{ nm}$, $\lambda_{\text{em}} = 410 \text{ nm}$) of poly(d ϵ A) with the DnaB helicase in 50 mM Tris/HCl (pH 8.1, 10 °C) containing 50 mM NaCl and 1 mM AMP-PNP, in the presence of different concentrations of poly(dA) (Nucleotide): (■) 0; (●) $2 \times 10^{-5} \text{ M}$; (◆) $1.2 \times 10^{-4} \text{ M}$. The concentration of the poly(d ϵ A) is $2 \times 10^{-5} \text{ M}$ (Nucleotide). The solid lines separate the sets of data points and do not have theoretical bases. The horizontal dashed line connects points on all titration curves characterized by the same value of the relative fluorescence increase, ΔF_i . The intersection points of the dashed horizontal line with the titration curves, in the presence of poly(dA), define the total DnaB concentrations, P_{T2} and P_{T2} , at which the binding density on poly(d ϵ A), $(\Sigma v_i)_R$, the binding density on poly(dA), $(\Sigma v_i)_S$, and the free helicase concentration, P_F , are constant. The total concentrations of nucleic acids at each titration are included in the figure (Jezewska and Bujalowski, 1996).

Figure 2.10.

The dependence of the relative fluorescence increase of poly(dεA) upon the binding density, $(\Sigma v_i)_S$, of the DnaB helicase on competing poly(dA) in 50 mM Tris/HCl (pH 8.1, 10 °C) containing 50 mM NaCl and 1 mM AMP-PNP. The dependence of the relative fluorescence increase of poly(dεA) upon the binding density, $(\Sigma v_i)_R$, of the DnaB helicase on poly(dεA), in the same buffer conditions, is also included (dashed, straight line). The solid line is the computer 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 (see below), for the simultaneous binding of the DnaB helicase to poly(dεA) and poly(dA) (**30**). The binding of the helicase to the reference 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 $n = 20$. 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 $K = 1.4 \times 10^6 \text{ M}^{-1}$.