[24] Thermodynamics of Ligand–Nucleic Acid Interactions

By TIMOTHY M. LOHMAN and DAVID P. MASCOTTI

I. Introduction

The noncovalent interactions of proteins and other ligands with DNA and RNA are of central importance to DNA metabolism and gene expression in all organisms. As a result, these interactions have been investigated by a variety of approaches in order to understand them at the thermodynamic, kinetic, and structural level. The structural details of the individual macromolecules and their complexes can provide information about some of the molecular contacts that can occur within the complex.[1,2] However, an understanding of the basis for their function and control requires knowledge of their stability, specificity, and mechanisms of interaction, which can only be achieved through thermodynamic and kinetic studies.[3-6]

Protein–nucleic acid interactions can be classified generally as either specific or nonspecific, depending on whether the proteins bind with high selectivity to a particular region of DNA, usually defined by its nucleotide sequence or conformation. However, specific nucleic acid-binding proteins, such as gene regulatory proteins, also bind to nonspecific regions with lower affinity. In fact, the competition of nonspecific DNA binding sites for the *Escherichia coli lac* repressor is believed to play an essential role in the regulation of expression of the *lac* operon *in vivo*,[7-9] and this is likely to be the case for other operons. There is also a general class of nucleic acid-binding proteins that bind to DNA exclusively in a nonspecific manner, with little dependence on the nucleotide sequence.[10] Examples of these are the helix-destabilizing (single-strand DNA binding) proteins,[11-15] histones,[16] as well as a variety of enzymes involved in DNA replication, recombination, and repair.[17]

The stability of specific protein–nucleic acid complexes is dependent on the specific array of hydrogen bonds and other contacts that can form between the protein and the nucleic acid ("direct readout") as well as the conformation of the nucleic acid (single stranded and various duplex forms), which can also be dependent on the nucleotide sequence ("indirect readout").[2] The specificity of the interaction of a protein for a particular nucleic acid site is also dependent on the affinity of the protein for the large number of other nonspecific sites that exist within a linear nucleic acid.[3] However, the stabilities and specificities of protein–nucleic acid interactions, just as for all macromolecular interactions, are also extremely sensitive to the solution environment (salt concentration and type, pH), since these small molecules (e.g., ions) generally interact directly and preferentially with the free protein, free DNA, and their complexes.[9,18,19] Because the specific and nonspecific interactions of any single protein can display differential sensitivities to solution variables, the nucleic acid binding specificity of proteins can also be a function of solution variables.[4,20,21]

The binding of proteins and small positively charged ligands to nucleic acids is particularly sensitive to the ionic environment in solution. This is due to the polyelectrolyte nature of linear nucleic acids, which results in the local accumulation to high concentrations of cations (e.g., K^+, Mg^{2+}) in the vicinity of the nucleic acid,[22-25] some of which are released on

[1] C. O. Pabo and R. T. Sauer, *Annu. Rev. Biochem.* **53,** 293 (1984).
[2] S. C. Harrison and A. K. Aggarwal, *Annu. Rev. Biochem.* **59,** 933 (1990).
[3] M. T. Record, Jr., J.-H. Ha, and M. A. Fisher, this series, Vol. 208, p. 291.
[4] D. R. Lesser, M. W. Kurpiewski, and L. Jen-Jacobson, *Science* **250,** 776 (1990).
[5] T. M. Lohman, *Crit. Rev. Biochem.* *19,* 191 (1986).
[6] P. H. von Hippel, D. G. Bear, W. D. Morgan, and J. A. McSwiggen, *Annu. Rev. Biochem.* **53,** 389 (1984).
[7] P. H. von Hippel, A. Revzin, C. A. Gross, and A. C. Wang, *Proc. Natl. Acad. Sci. U.S.A.* **71,** 4808 (1984).
[8] P. H. von Hippel, *in* "Biological Regulation and Development" (R. F. Goldberger, ed.), Vol. 1, p. 279. Plenum, New York, 1979.
[9] M. T. Record, Jr., and R. S. Spolar, *in* "Nonspecific DNA–Protein Interactions" (A. Revzin, ed.), p. 33. CRC Press, Boca Raton, Florida, 1990.
[10] A. Revzin, *in* "Nonspecific DNA–Protein Interactions" (A. Revzin, ed.), p. 5. CRC Press, Boca Raton, Florida, 1990.
[11] S. C. Kowalczykowski, D. G. Bear, and P. H. von Hippel, *in* "The Enzymes" (P. D. Boyer, ed.), Vol. 14, p. 373. Academic Press, New York, 1981.
[12] T. M. Lohman and W. Bujalowski, *in* "The Biology of Nonspecific DNA–Protein Interactions" (A. Revzin, ed.), p. 131. CRC Press, Boca Raton, Florida, 1990.
[13] R. L. Karpel, *in* "Nonspecific DNA–Protein Interactions" (A. Revzin, ed.), p. 103. CRC Press, Boca Raton, Florida, 1990.
[14] K. R. Williams and J. W. Chase, *in* "Nonspecific DNA–Protein Interactions" (A. Revzin, ed.), p. 197. CRC Press, Boca Raton, Florida, 1990.
[15] J. W. Chase and K. R. Williams, *Annu. Rev. Biochem.* **55,** 103 (1986).
[16] M. J. Behe, *in* "Nonspecific DNA–Protein Interactions" (A. Revzin, ed.), p. 229. CRC Press, Boca Raton, Florida, 1990.
[17] M. M. Cox, *in* "Nonspecific DNA–Protein Interactions" (A. Revzin, ed.), p. 171. CRC Press, Boca Raton, Florida, 1990.
[18] M. T. Record, Jr., T. M. Lohman, and P. H. deHaseth, *J. Mol. Biol.* **107,** 145 (1976).
[19] M. T. Record, Jr., C. F. Anderson, and T. M. Lohman, *Q. Rev. Biophys.* **11,** 103 (1978).
[20] M. T. Record, Jr., P. H. deHaseth, and T. M. Lohman, *Biochemistry* **16,** 4791 (1977).
[21] M. C. Mossing and M. T. Record, Jr., *J. Mol. Biol.* **186,** 295 (1985).
[22] G. S. Manning, *J. Chem. Phys.* **51,** 924 (1969).
[23] G. S. Manning, *Q. Rev. Biophys.* **11,** 179 (1978).

Copyright © 1992 by Academic Press, Inc.
All rights of reproduction in any form reserved.

formation of the protein–nucleic acid complex.[5,18,19] In fact, much of the stability of complexes between gene regulatory proteins and nonspecific DNA is largely due to the large positive entropy changes that result from ion release on formation of the complex.[26] As a result of the dependence of protein–nucleic acid stability and specificity on solution conditions, it is impossible to understand the forces that drive these interactions based solely on structural considerations. Rather, as discussed in the first part of this chapter, the equilibrium binding properties (thermodynamics, energetics) of these interactions must be investigated as a function of solution conditions in order to understand the origins of stability and specificity.

Studies of the thermodynamics of protein– and ligand–nucleic acid interactions can in principle be examined by two approaches: calorimetry and the measurement of equilibrium binding constrants as a function of solution variables. The combined use of both approaches has proved extremely useful in studies of small ligand–DNA interactions.[27,28] However, to date, there have been no calorimetric studies reported for any protein–nucleic acid interaction, and all of the thermodynamic information available on these interactions has been obtained from equilibrium binding studies. Therefore, in this discussion we focus on the use of equilibrium binding studies to obtain such information.

To obtain meaningful thermodynamic information from studies of the dependence of equilibrium binding constants on solution variables, techniques and methods of analysis are required that yield rigorous equilibrium binding isotherms, which can be analyzed to extract equilibrium binding parameters. The rigorous use of spectroscopic methods in this regard has been discussed[29,30] (see also [25] in this volume[31]).

II. Equilibrium Binding Parameters

A. *Equilibrium Binding Constant* (K_{obs})

Consider the interaction of a ligand, L, with a nucleic acid, D, to form a complex, LD, as in Eq. (1):

[24] C. F. Anderson and M. T. Record, Jr., *Annu. Rev. Phys. Chem.* **33**, 191 (1982).

[25] M. T. Record, Jr., M. Olmsted, and C. F. Anderson, *in* "Theoretical Biochemistry and Molecular Biophysics" (D. L. Beveridge and R. Lavery, eds.), p. 285. Adenine Press, Schenectady, New York, 1990.

[26] P. L. deHaseth, T. M. Lohman, and M. T. Record, Jr., *Biochemistry* **16**, 4783 (1977).

[27] L. A. Marky and K. J. Breslauer, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4359 (1987).

[28] W. Y. Chou, L. A. Marky, D. Zaunczkowski, and K. J. Breslauer, *J. Biomol. Struct. Dyn.* **5**, 345 (1987).

[29] W. Bujalowski and T. M. Lohman, *Biochemistry* **26**, 3099 (1987).

[30] T. M. Lohman and W. Bujalowski, this series, Vol. 208, p. 258.

[31] T. M. Lohman and D. P. Mascotti, this volume [25].

$$L + D \overset{K_{obs}}{\rightleftharpoons} LD \qquad (1)$$

This equilibrium can be described by the intrinsic (microscopic) equilibrium constant, K_{obs}, defined in Eq. (2):

$$K_{obs} = \frac{[LD]}{[L][D]} \qquad (2)$$

We emphasize that the equilibrium in Eq. (1), described by the intrinsic equilibrium binding constant K_{obs}, reflects the binding of only one form of the ligand, L, to yield one type of complex, LD. Therefore, if multiple forms of the ligand exist in equilibrium (e.g., a protein that exists in equilibrium between monomer, dimer, and tetramer forms), then a separate intrinsic binding constant is required to describe the interaction of each form of the ligand with the nucleic acid. Similarly, if a ligand can bind to the DNA in a number of different modes, then the binding of each mode is described by a separate intrinsic binding constant. Only if K_{obs} is such an intrinsic equilibrium constant can it be used to calculate the true thermodynamic quantities, $\Delta G°$, $\Delta H°$, and $\Delta S°$, for the equilibrium. The relationships among these quantities are given in Eq. (3), where R is the gas constant and T the absolute temperature:

$$\Delta G° = -RT \ln K_{obs} = \Delta H° - T\Delta S° \qquad (3)$$

If L, D, or LD are involved in equilibria in addition to that defined by Eq. (1), and these other equilibria are not explicitly considered in the analysis of the interaction, then only an "apparent" binding constant will be obtained from an analysis of the binding interaction. This "apparent" binding constant will be a composite parameter reflecting all of the multiple equilibria and hence will not be related in a simple (or necessarily known) manner to the free energy change for the reaction in Eq. (1). To determine whether such multiple equilibria exist, it is generally advisable to examine the ligand–nucleic acid interactions over a range of ligand and nucleic acid concentrations, at constant solution conditions. If L and D are involved only in the equilibrium defined in Eq. (1), then the value of K_{obs} will be independent of the ligand and nucleic acid concentration. If L and D are involved in other equilibria, however, then the experimental binding constant will be a function of the ligand and/or nucleic acid concentrations and therefore will not represent the intrinsic binding constant for the reaction in Eq. (1). In this context, it is essential to understand and characterize the solution behavior of the free ligand (quaternary structure, self-assembly equilibria) and the ligand–nucleic acid complex, quantitatively, before embarking on a quantitative investigation of its equilibrium binding and thermodynamic properties.

B. Ligand Binding Site Size

The ligand binding site size, *n*, represents the number of bases or base pairs that are occupied by the ligand on binding to the nucleic acid. The site size for ligand binding defines the region of the nucleic acid that is occluded by the ligand and hence not accessible for interaction with other ligands. This occluded site size is not necessarily the same as the number of nucleotides contacted by the ligand, although *n* will provide an upper estimate for the number of contacts. Knowledge of the ligand site size is necessary for an accurate description of the nonspecific equilibrium binding of large ligands to long, linear nucleic acids, since the number of free sites on the nonspecific nucleic acid lattice is dependent on *n* as well as the ligand binding density.[32] These aspects are discussed further in Section III,B.

C. Cooperativity of Ligand Binding

Cooperative binding of ligands to linear nucleic acids is possible when multiple ligands (proteins) can bind to the same nucleic acid molecule. Cooperativity is a thermodynamic quantity that reflects the influence of one bound ligand on the binding affinity of a second ligand. If the free energy changes on binding two or more ligands to the same nucleic acid are not independent, then cooperative interactions are indicated. Cooperativity can be either positive or negative; that is, the binding affinity of a second ligand can be either enhanced or reduced with respect to the affinity of the first ligand. True cooperativity reflects changes in the intrinsic binding parameters and should not be confused with statistical effects, such as the apparent negative cooperativity that results from the overlap of potential nonspecific ligand binding sites.[32]

Cooperative effects can be due to nearest-neighbor interactions, namely, interactions resulting from ligands bound to adjacent sites on a linear DNA molecule, or non-nearest-neighbor interactions. Nearest-neighbor cooperativities can result from ligand–ligand interactions and/or short-range conformational changes that are induced in the DNA on ligand binding. Non-nearest-neighbor cooperativity has been observed between DNA-binding proteins that are bound to specific sites that are well separated along the contour length of a linear nucleic acid molecule.[33,34] These cooperativities have generally been explained by models that invoke direct

[32] J. D. McGhee and P. H. von Hippel, *J. Mol. Biol.* **86,** 469 (1974).
[33] J. A. Borowiec, L. Zhang, S. Sasse-Dwight, and J. D. Gralla, *J. Mol. Biol.* **196,** 101 (1987).
[34] S. Oehler, E. R. Eismann, H. Kramer, and B. Müller-Hill, *EMBO J.* **9,** 973 (1990).

interactions between the proteins, resulting in "looping" of the intervening DNA.[35,36]

Positive cooperativity has been observed for a number of nonspecific nucleic acid-binding proteins, including the *E. coli* SSB (single-strand-specific binding) protein,[37–39] the bacteriophage T4 gene 32 protein,[40,41] and the bacteriophage gene 5 protein,[42–44] as well as for some sequence-specific DNA-binding proteins.[45,46] Both positive and negative cooperative effects have been observed for the binding of some small ligands, such as drugs and dyes, to duplex DNA,[47,48] although negative cooperativity in these systems is small and difficult to resolve from the statistical effects that result from overlap of potential nonspecific binding sites. Section III,B discusses different statistical thermodynamic models for the treatment of nearest-neighbor cooperativity between ligands bound nonspecifically to linear nucleic acids, as well as a quantitative definition of cooperativity.

III. Sequence Specific *versus* Nonspecific Binding

Protein (ligand)–nucleic acid interactions can generally be divided into two classes: (1) site-specific interactions, in which the protein binds with high affinity to a specific site(s), usually defined by the nucleotide sequence, and (2) nonspecific interactions in which the protein binds to nucleic acids with the same general affinity, nearly independent of nucleotide sequence. Of course, ligands that bind with sequence specificity can also bind to nonspecific regions of the nucleic acid, although with signifi-

[35] R. Schleif, *Science* **240,** 127 (1988).
[36] G. R. Bellomy and M. T. Record, Jr., *in* "Progress in Nucleic Acid Research and Molecular Biology" (W. Cohn and K. Moldave, eds.), Vol. 39, p. 81. Academic Press, New York, 1990.
[37] N. Sigal, H. Delius, T. Kornberg, M. L. Gefter, and B. Alberts, *Proc. Natl. Acad. Sci. U.S.A.* **69,** 3537 (1972).
[38] T. M. Lohman, L. B. Overman, and S. Datta, *J. Mol. Biol.* **187,** 603 (1986).
[39] W. Bujalowski and T. M. Lohman, *J. Mol. Biol.* **195,** 897 (1987).
[40] B. Alberts and L. Frey, *Nature (London)* **227,** 1313 (1970).
[41] S. C. Kowalczykowski, N. Lonberg, J. W. Newport, and P. H. von Hippel, *J. Mol. Biol.* **145,** 75 (1981).
[42] B. Alberts, L. Frey, and H. Delius, *J. Mol. Biol.* **68,** 139 (1972).
[43] S. J. Cavalieri, K. E. Neet, and D. A. Goldthwait, *J. Mol. Biol.* **102,** 697 (1976).
[44] D. Porschke and H. Rauh, *Biochemistry* **22,** 4737 (1983).
[45] A. D. Johnson, A. R. Poteete, G. Lauer, R. T. Sauer, G. K. Ackers, and M. Ptashne, *Nature (London)* **294,** 217 (1981).
[46] G. K. Ackers, A. D. Johnson, and M. A. Shea, *Proc. Natl. Acad. Sci. U.S.A.* **79,** 1129 (1982).
[47] W. D. Wilson and I. G. Lopp, *Biopolymers* **18,** 3025 (1979).
[48] L. S. Rosenberg, M. J. Carvlin, and T. R. Krugh, *Biochemistry* **25,** 1002 (1986).

cantly lower affinity. The nonspecific binding of otherwise specific DNA-binding proteins appears to play important roles in the regulation of gene expression[7,8] as well as in facilitating the location of the specific DNA site by the protein.[5,49-53]

A. *Site-Specific Equilibrium Binding Isotherms*

The intrinsic observed equilibrium constant, K_{PS}, for a site-specific protein–nucleic acid interaction is given in Eq. (4):

$$K_{PS} = \frac{[PS]}{[P][S]} \tag{4}$$

where P, S, and PS represent the free protein, specific nucleic acid site, and the complex, respectively. In the absence of cooperativity, the fractional occupancy of the specific site, $\Theta = [PS]/S_{TOTAL}$, is described by the simple Langmuir isotherm given in Eq. (5):

$$\Theta = \frac{K_{PS}[P]}{(1 + K_{PS}[P])} \tag{5}$$

However, as the specific site is usually contained within a larger region of nonspecific DNA, one must also account for the competitive binding of protein to the nonspecific sites.[7] For the case of multiple specific binding sites, where cooperativity between proteins bound to the specific sites can occur, then the simple expression in Eq. (5) must be modified as in the case of the bacteriophage λ *cI* repressor, which binds to three specific sites in both the λO_R and λO_L operators.[45,46]

B. *Nonspecific Equilibrium Binding Isotherms*

The quantitative description of the equilibrium isotherm for a large ligand (protein) binding nonspecifically to linear nucleic acids is more complicated than Eq. (5), since potential nonspecific ligand binding sites overlap. Large ligands are defined as those with site sizes, n, greater than 1 residue (nucleotide or base pair). As a result, the number of available binding sites for the ligand decreases nonlinearly with increasing ligand binding density, ν (moles of ligand bound per residue). This statistical effect results in curvature in the equilibrium binding isotherm, even for a

[49] P. H. Richter and M. Eigen, *Biophys. Chem.* **2,** 255 (1974).
[50] O. G. Berg and C. Blomberg, *Biophys. Chem.* **4,** 367 (1976).
[51] D. R. Dowd and R. S. Lloyd, *J. Biol. Chem.* **265,** 3424 (1990).
[52] O. G. Berg and P. H. von Hippel, *Annu. Rev. Biophys. Biophys. Chem.* **14,** 131 (1985).
[53] O. G. Berg, R. B. Winter, and P. H. von Hippel, *Trends Biochem. Sci.* **7,** 52 (1982).

noncooperative binding ligand, and has been described quantitatively for both infinite nucleic acid lattices[32,54,55] as well as finite nucleic acid lattices.[56,57]

We briefly describe two models for the nonspecific binding of large ligands to homogeneous infinite nucleic acid lattices that consider cooperative ligand binding. Both of these treat cooperative interactions only between nearest-neighbor bound ligands. Although more complicated non-nearest-neighbor models have been described,[58-60] these models require more parameters than can usually be determined from an experimental binding isotherm. The two nearest-neighbor cooperativity models that we consider differ in that the ligands possess either "unlimited" or "limited" cooperativity, as defined below and depicted in Fig. 1.[61] In the unlimited cooperativity model, clusters of bound ligands are limited only by the length of the nucleic acid, whereas in the limited cooperativity model, the ligand cluster size is limited, even for high values of cooperativity.

1. Unlimited Cooperativity. A detailed description of the quantitative aspects of this model has been given[32,55] and Fig. 2 defines the binding parameters for this model. In the "unlimited" cooperativity model, a ligand with site size, n residues (nucleotides or base pairs), can bind to an infinite, linear, homogeneous nucleic acid lattice in any of three modes: an isolated mode, with intrinsic equilibrium constant, K; a singly contiguous mode (one nearest-neighbor ligand), with equilibrium constant $K\omega$; a doubly contiguous mode (two nearest-neighbor ligands), with equilibrium constant, $K\omega^2$. The cooperativity parameter, ω, is a unitless equilibrium constant for the process of moving two bound isolated ligands so that both are bound singly contiguously (see Fig. 2). The cooperativity parameter, ω, is related to the standard free energy change for this process by $\Delta G^\circ_{coop} = -RT \ln \omega$. Positive cooperativity is reflected by $\omega > 1$ ($\Delta G^\circ_{coop} < 0$), negative cooperativity results when $\omega < 1$ ($\Delta G^\circ_{coop} > 0$), and noncooperative interactions are described by $\omega = 1$ ($\Delta G^\circ_{coop} = 0$).

The model predicts that when positive cooperativity exists ($\omega > 1$), the ligands will form clusters of variable length along the nucleic acid,

[54] A. S. Zasedatelev, G. V. Gursky, and M. V. Volkenshtein, *Mol. Biol. (U.S.S.R.)* **5,** 245 (1971).
[55] J. A. Schellman, *Isr. J. Chem.* **12,** 219 (1974).
[56] I. R. Epstein, *Biophys. Chem.* **8,** 327 (1978).
[57] J. A. Schellman, *in* "Molecular Structure and Dynamics" (M. Balaban, ed.), p. 245. International Science Services, Philadelphia, 1980.
[58] Y. Chen, *Biophys. Chem.* **27,** 59 (1987).
[59] G. V. Gursky and A. S. Zasedatelev, *Sov. Sci. Rev. Physicochem. Biol.* **5,** 53 (1984).
[60] Y. D. Nechipurenko and G. V. Gursky, *Biophys. Chem.* **24,** 195 (1986).
[61] T. M. Lohman, W. Bujalowski, and L. B. Overman, *Trends Biochem. Sci.* **13,** 250 (1988).

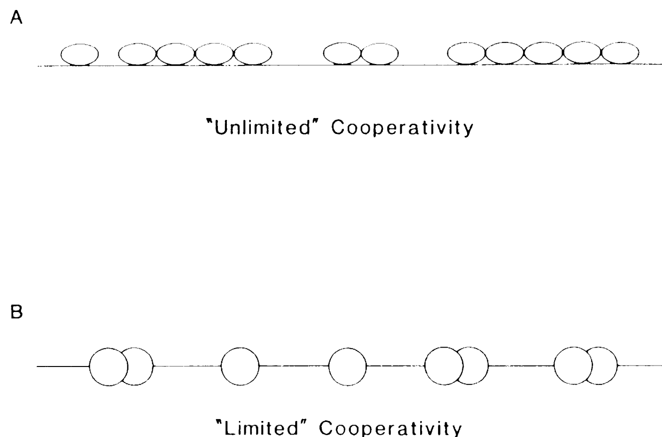


FIG. 1. Diagram illustrating two different types of nearest-neighbor cooperative binding of ligands to linear nucleic acid. (A) "Unlimited" cooperativity, typified by T4 gene 32 protein,[41] in which long clusters of bound protein can form owing to nearest-neighbor interactions occurring on both sides of a bound protein. As the cooperativity parameter, ω (see Fig. 2 for definition), increases to infinity, a single protein cluster will form along the nucleic acid lattice. (B) "Limited" cooperativity, typified by the *E. coli* SSB tetramer binding to single-stranded DNA in its $(SSB)_{65}$ binding mode.[39] In this case, each circle represents an SSB tetramer, and clustering is limited to the formation of dimers of tetramers (octamers), such that no doubly contiguously bound tetramers can form. With this type of cooperativity, it is difficult to saturate the DNA lattice. (From Lohman *et al.*[61])

depending on the binding density, ν (for given values of n, K, and ω), which are limited in size only by the length of the nucleic acid. At a given ligand binding density, the average ligand cluster length is determined by the values of ω and n.[32,55] McGhee and von Hippel[32] obtained separate closed-form expressions for the equilibrium binding isotherms for this model for $\omega = 1$ and $\omega \neq 1$; however, these can be combined into a single expression for the binding isotherm[62] that is valid for all values of $\omega > 0$, and this is given in its Scatchard form in Eq. (6):

$$\nu/L = K(1 - n\nu)\{[2\omega(1 - n\nu)]/[(2\omega - 1)(1 - n\nu) + \nu + R]\}^{(n-1)}\{[1 - (n + 1)\nu + R]/[2(1 - n\nu)]\}^2 \quad (6)$$

where L is the free ligand concentration and $R = \{[1 - (n + 1)\nu]^2 + 4\omega\nu(1 - n\nu)\}^{1/2}$. This nearest-neighbor unlimited cooperativity model provides a good description of the nonspecific, cooperative binding of the bacteriophage T4 gene 32 protein to single-stranded polynucleo-

[62] W. Bujalowski, T. M. Lohman, and C. F. Anderson, *Biopolymers* **28**, 1637 (1989).

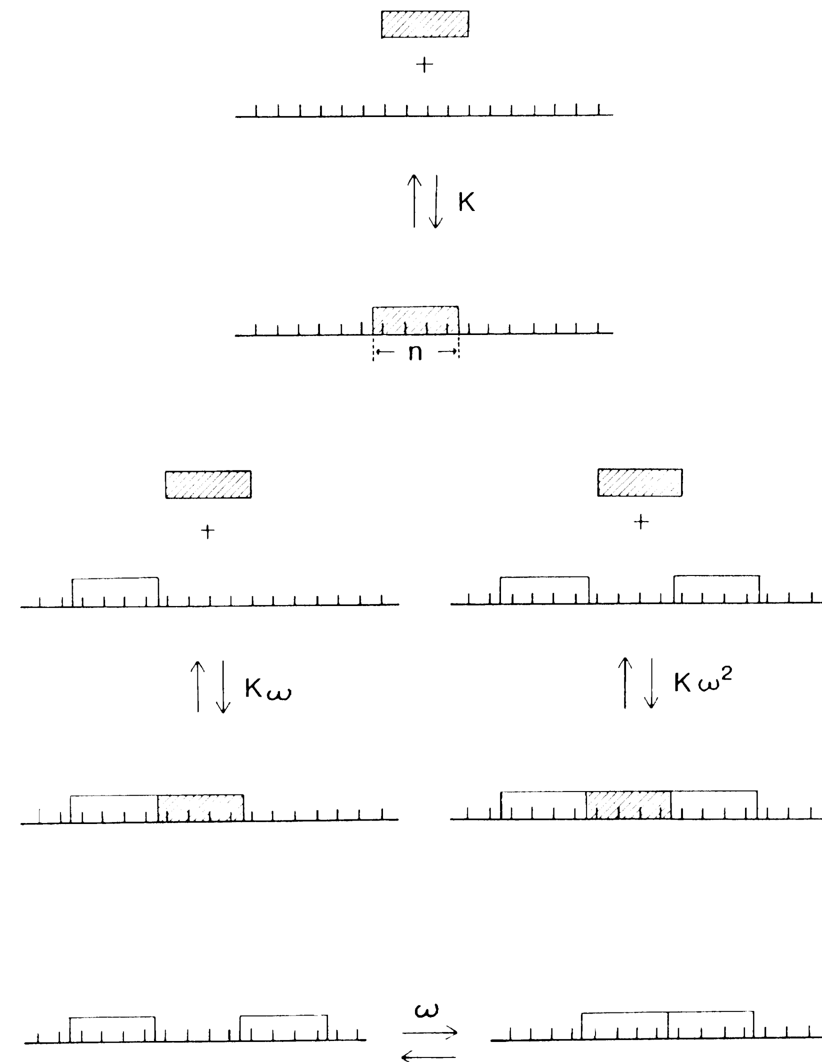


FIG. 2. Definition of the equilibrium binding parameters n (site size), K (intrinsic binding constant), and ω (cooperativity parameter) for the nearest-neighbor, "unlimited" cooperativity model[32] for binding of a ligand to an infinite, one-dimensional, homogeneous lattice.

tides.[11,41,63,64] However, a number of nonspecific cooperatively binding proteins do not seem to conform to this model, even though their equilibrium binding isotherms can be described well by this model. These include the phage fd gene 5 protein, which displays non-nearest-neighbor cooperative interactions,[65] and the *E. coli* SSB protein in its $(SSB)_{65}$ binding mode, which appears to bind with limited cooperativity,[12,39] although it may bind in an unlimited cooperativity mode in its $(SSB)_{35}$ binding mode.[12,38]

An alternative "free sliding ligand" model has been presented[66] to describe the nonspecific binding of large ligands to an infinite linear lattice, including nearest-neighbor cooperativity. In this model, the ligand is not constrained to bind to distinct sites on the nucleic acid lattice, rather it is assumed to translate freely along the nucleic acid lattice. Attempts have also been made to incorporate intrinsic negative cooperative effects that should arise for the binding of small positively charged ligands to a linear nucleic acid at high binding density.[67] However, this approach[67] overestimates the negative cooperativity, at least at low binding densities, for the binding of oligolysines to linear nucleic acids.[68]

2. Limited Cooperativity. In the "limited" cooperativity model, cooperative interactions also occur only between nearest-neighbor bound ligands; however, the interactions are such that clustering is limited to the formation of dimers of ligands; hence, cooperativity is "limited" (see Fig. 1B).[39] The closed-form expression for the binding isotherm (Scatchard form[69]) for this model is

$$\nu/L = K[q^2 - 2\nu q + (1 - \omega)\nu^2]^n/q^{(2n-1)} \tag{7}$$

where $q = [1 - (n - 1)\nu] + \{[1 - (n - 1)\nu]^2 - \nu(1 - \omega)[2 - (2n - 1)\nu]\}^{1/2}$. This limited cooperativity model provides a better description of the nonspecific binding of the *E. coli* SSB tetramer to single-stranded nucleic acids in its $(SSB)_{65}$ binding mode than does the unlimited cooperativity model.[39]

3. Noncooperative Binding. The closed form expression for the equilibrium binding isotherm for noncooperative, nonspecific ligand binding to an infinite, linear nucleic acid can be obtained by substituting $\omega = 1$ into Eq. (6), with the resulting expression[32]

$$\nu/L = K(1 - n\nu)[(1 - n\nu)/(1 - (n - 1)\nu)]^{(n-1)} \tag{8}$$

[63] T. M. Lohman, *Biochemistry* **23,** 4656 (1984).
[64] T. M. Lohman, *Biochemistry* **23,** 4665 (1984).
[65] J. E. Coleman and J. L. Oakley, *Crit. Rev. Biochem.* **98,** 247 (1980).
[66] C. P. Woodbury, *Biopolymers* **20,** 2225 (1981).
[67] R. A. G. Friedman and G. S. Manning, *Biopolymers* **23,** 2671 (1984).
[68] D. P. Mascotti and T. M. Lohman, unpublished, 1992.
[69] G. Scatchard, *Ann. N.Y. Acad. Sci.* **51,** 660 (1949).

We note that Eq. (8) can also be obtained from Eq. (7) on substituting $\omega = 1$; however, the equilibrium constants for the two expressions differ by a statistical factor of 2, since in the unlimited cooperativity model, the ligands are assumed to bind in only one orientation along the nucleic acid, whereas the limited cooperativity model assumes bidirectional ligand binding. However, if the same polarity of ligand binding is assumed for both models then the definitions of K for each model are identical.

IV. Dependence of K_{obs} on Salt Concentration

The intrinsic binding constant defined in Eqs. (1) and (2) is an "observed" binding constant, that is, K_{obs} is defined in terms of only the ligand and nucleic acid species, independent of their interactions with solvent or low molecular weight solutes (e.g., ions, protons).[19] The true thermodynamic equilibrium constant, K_T, is defined in terms of the activities of all species that participate in the reaction and hence is dependent only on temperature and pressure. Therefore, if preferential interactions of cations, anions, water, or protons occur with any of the species, L, D, or LD, then K_{obs}, which is the quantity that is measured experimentally, will be dependent on the bulk salt concentration and pH, as well as temperature and pressure.

A. Cation Effects on Ligand–Nucleic Acid Equilibria

1. Polyelectrolyte Effect: Cation Release from Nucleic Acid. The majority of protein–nucleic acid interactions, as well as positively charged small ligand–nucleic acid interactions, are extremely sensitive to the salt concentration in solution.[5,18,19] This stems from the fact that linear nucleic acids are highly charged polyanions and as a result sequester cations in their vicinity in order to reduce the net charge on the nucleic acid.[22–25,70] The interaction of counterions (M^+) with a linear polyelectrolyte such as DNA is equivalent, thermodynamically, to having a constant fraction of a counterion bound per nucleic acid phosphate.[18] In aqueous media, the fraction of a counterion thermodynamically bound per phosphate, ψ, is dependent only on the structural charge density along the nucleic acid and the counterion valance and is independent of the bulk salt concentration, as long as it is in excess over the phosphate charge.[19,22,23] This has been verified experimentally by ^{23}Na NMR studies of Na–DNA solutions[71] and by examining the salt dependence of a number of oligocations interacting with duplex and single-stranded nucleic acids. For double-stranded B-form

[70] C. F. Anderson and M. T. Record, Jr., *Annu. Rev. Biophys. Chem.* **19,** 423 (1990).
[71] C. F. Anderson, M. T. Record, Jr., and P. A. Hart, *Biophys. Chem.* **7,** 310 (1978).

DNA, ψ has a value of 0.88,[18] whereas for single-stranded nucleic acids, near neutral pH, ψ is approximately 0.70–0.74.[72,73] The lower value of ψ for single-stranded nucleic acids reflects its lower charge density, relative to the duplex form.[74,75]

As a result, the association of a simple oligocation (net charge $+z$) with a linear nucleic acid in the presence of excess univalent salt, MX, can be represented, thermodynamically, by the following equilibrium:

$$L^{z+} + D \rightleftharpoons LD + z\psi M^+ \quad (9)$$

We note that the coefficient $z\psi$ includes contributions due to both the direct release of cations previously bound to the nucleic acid as well as perturbations of the electrostatic screening of the phosphates by the remaining bulk ions (activity coefficient effects). Application of Le Chatelier's principle to Eq. (9) indicates that K_{obs} will increase as the salt (MX) concentration is decreased. The predicted quantitative dependence of K_{obs} on $[M^+]$ is[18,19]

$$\frac{\partial \log K_{obs}}{\partial \log[M^+]} = -z\psi \quad (10)$$

Because the coefficient, $z\psi$, is constant for a given ligand and nucleic acid, this predicts that the logarithm of K_{obs} will decrease as a linear function of the logarithm of the monovalent counterion (M^+) concentration. Furthermore, since $z\psi$ is generally large, the decrease of K_{obs} with increasing salt concentration is also quite dramatic.[18,19]

The salt dependence of K_{obs} for the binding to poly(U) of a series of oligolysines with varying charge, $+z$, is shown in Fig. 3. For each oligopeptide, $\log K_{obs}$ decreases linearly with $\log[K^+]$, and the absolute value of the slope, $|\partial \log K_{obs}/\partial \log[K^+]|$, is proportional to the oligopeptide charge. The value of the proportionality constant determined from the data is $\psi = 0.71 \pm 0.03$, which is consistent with the lower charge density for a single-stranded nucleic acid.[73] Similar behavior has been observed for the interaction of positively charged oligopeptides and polyamines with duplex nucleic acids[18,73,76–79]; however, $\psi = 0.90 \pm 0.05$ due to the higher charge density of the duplex nucleic acids.[18]

[72] M. T. Record, Jr., C. P. Woodbury, and T. M. Lohman, *Biopolymers* **15**, 893 (1976).
[73] D. P. Mascotti and T. M. Lohman, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 3142 (1990).
[74] M. T. Record, Jr., *Biopolymers* **14**, 2137 (1975).
[75] G. S. Manning, *Biopolymers* **11**, 937 (1972).
[76] S. A. Latt and H. A. Sober, *Biochemistry* **6**, 3293 (1967).
[77] S. A. Latt and H. A. Sober, *Biochemistry* **6**, 3307 (1967).
[78] T. M. Lohman, P. H. deHaseth, and M. T. Record, Jr., *Biochemistry* **19**, 3522 (1980).
[79] W. H. Braunlin, T. J. Strick, and M. T. Record, Jr., *Biopolymers* **21**, 1301 (1982).

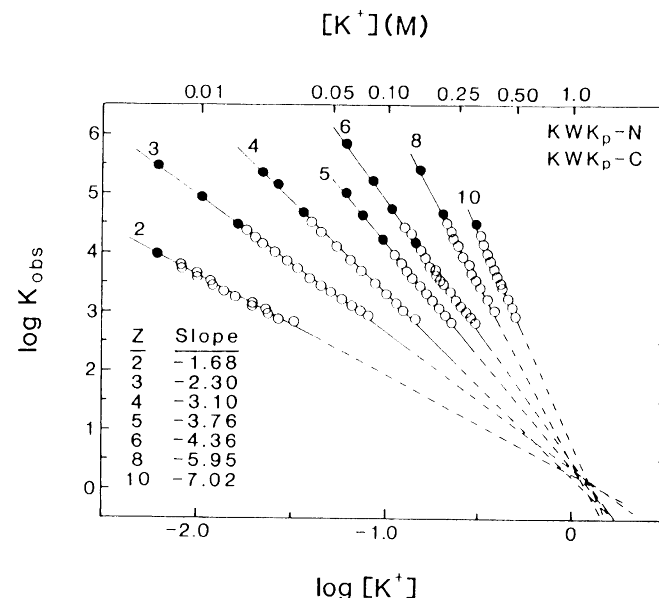


FIG. 3. Dependence of $\log K_{obs}$ on $\log[K^+]$ for a series of oligolysines binding to poly(U) (25.0°, pH 6.0). KCH_3CO_2 is used to vary the K^+ concentration. The oligopeptide sequences are KWK_p-NH_2, with $p = 1, 2, 4, 6$, and 8 ($z = +3, 4, 6, 8$, and 10) and KWK_p-CO_2, with $p = 1$ and 4 ($z = +2$ and 5). The net positive charge of each peptide is indicated for each line, and the slopes of each line are given. (●), Values of K_{obs} determined from an individual reverse titration at a constant salt concentration; (○), values of K_{obs} determined from salt back-titrations. (Data from Mascotti and Lohman.[73])

The release of thermodynamically bound counterions into the bulk salt solution provides a favorable entropic contribution ($\Delta S > 0$) to the free energy of binding (a free energy of dilution).[18] This free energy of dilution has been referred to as the polyelectrolyte effect[80] and is given by Eq. (11) for a ligand–nucleic acid equilibrium occurring in a univalent salt solution[18]:

$$\Delta G^\circ_{PE} = z\psi RT \ln[M^+] \quad (11)$$

Equation (11) indicates that the polyelectrolyte effect is eliminated at 1 M K^+, which is the standard state for the reaction. This behavior has been verified for a number of simple oligocations binding to both duplex[18,78,79] and single-stranded nucleic acids, an example of which is shown in Fig. 3

[80] M. T. Record, Jr., *in* "Unusual DNA Structures" (R. D. Wells and S. C. Harvey, eds.), p. 237. Springer-Verlag, New York, 1988.

for the oligolysine–poly(U) interaction.[73] This polyelectrolyte effect (counterion release from the nucleic acid) provides the major driving force for the interactions of simple oligocations with linear nucleic acids at low salt concentrations.[18,73,78] This also appears to be the case for a number of nonspecific ligand (protein)–nucleic acid interactions, including the *E. coli lac* repressor.[26]

Of course, preferential cation interactions with the nucleic acid will also result from any process that affects the charge spacing of the nucleic acid. The latter effect has been explicitly considered in the interpretation of the salt dependence of K_{obs} for small planar ligands that can intercalate between the base pairs of duplex nucleic acids.[47] In this case, cation release results from both the neutralization of phosphate charge as well as the increase in the axial charge spacing of the duplex nucleic acid due to intercalation, although the former effect dominates the salt dependence.

Various methods have detected small differences in the affinities of different monovalent cations for duplex DNA.[81–84] ^{23}Na NMR measurements[84] indicate the following hierarchy of affinities: $Na^+ < Li^+ < K^+ < Cs^+ < NH_4^+$. However, it has been observed generally that the effects of cations on the equilibrium binding of simple oligocations to linear DNA are not sensitive to the type of monovalent cation, M^+. For example, identical values of K_{obs} and $\partial \log K_{obs}/\partial \log[M^+]$ are observed for the interaction of oligolysines with poly(U), independent of whether the cation is Na^+, K^+, or $NH_3(n\text{-}C_4H_9)^+$.[68] A similar independence of K_{obs} on the cation type (K^+, Na^+, NH_4^+) has been observed for the specific binding of the *E. coli lac* repressor to the O^{sym}[85] and wild-type *lac* operator,[86] although, for the latter interaction, K_{obs} in Cs^+ is larger by a factor of 2 and K_{obs} in Li^+ is smaller by a factor of 10. The latter effects may reflect differences in the affinity of these ions for the DNA; however, preferential interactions of these cations with the protein have also not been ruled out.

In an alternative view, Manning[23] has suggested that the coefficient ψ equals unity, which predicts that one monovalent cation should be released per oligocation positive charge on complex formation, independent of the nucleic acid charge density. However, recent studies of the equilibrium binding of a series of oligolysines to the single-stranded polynucleotide poly(U) indicate that for this interaction ψ is significantly less than unity (see Fig. 3).[73]

2. Electrostatic versus Nonelectrostatic Contributions to ΔG°_{obs}. The standard free energy change for a ligand–nucleic acid interaction, ΔG°_{obs}, can be viewed as having contributions arising from electrostatic and nonelectrostatic (noncoulombic) interactions. In the absence of preferential ion interactions with the ligand, Record *et al.*[18] have described a method for estimating the relative contributions from these interactions. This method is based on the fact that the polyelectrolyte effect [see Eq. (11)] is eliminated at 1 *M* M^+. Therefore, the free energy change on extrapolation to 1 *M* M^+ [ΔG°_{obs}(1 *M* M^+)], should reflect mainly contributions arising from nonelectrostatic interactions [ΔG°(non-el)] and the ionic interaction between the positively charged groups on the ligand and the negatively charged phosphates [ΔG°(ionic)]. This relationship is expressed in Eq. (12):

$$\Delta G^\circ_{obs}(1\ M\ M^+) = \Delta G^\circ(\text{non-el}) + z\Delta G^\circ(\text{ionic}) \quad (12)$$

ΔG°(ionic) is the free energy change per ionic interaction in the absence of counterion release; ΔG°(ionic) is assumed to be independent of salt concentration.

Based on analyses of the monovalent salt dependence of K_{obs} for a series of oligolysines, ΔG°(ionic) has been estimated to be very nearly zero, with a slightly unfavorable value of $+0.2 \pm 0.1$ kcal/mol (per ionic interaction) for interactions with duplex nucleic acids[18,78] and -0.1 ± 0.1 kcal/mol (per ionic interaction) for poly(U).[68] These values of ΔG°(ionic) are presumably close to zero, since the formation of an ionic interaction within the ligand–nucleic acid complex is partially compensated by the loss of a counterion (M^+)–phosphate interaction; in other words, the exchange reaction has a net free energy change near zero (at the 1 *M* M^+ standard state). Therefore, the value of ΔG°_{obs} (1 *M* M^+) provides an estimate of the value of ΔG°(non-el), with a possible small contribution for ΔG°(ionic).[18] However, we emphasize that this approach will not yield meaningful estimates of ΔG°(non-el) if significant preferential interactions with the ligand (protein) exist, since $\log K_{obs}$ will not be a linear function of $\log[M^+]$.

3. Competitive Effects of Monovalent and Divalent Cations for Binding to Nucleic Acids. As yet, we have discussed ligand–nucleic acid equilibria only in the presence of excess monovalent salt, MX. The effects on K_{obs} caused by divalent cations such as Mg^{2+} are qualitatively similar; however, the coefficient ψ in Eq. (9) is replaced by ϕ, which represents the number of divalent counterions thermodynamically associated per phosphate.[26]

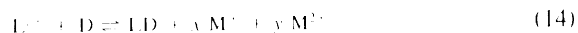
[81] B. Wolf and S. Hanlon, *Biochemistry* **14,** 1661 (1975).
[82] A. Chan, R. Kilkuskie, and S. Hanlon, *Biochemistry* **18,** 84 (1979).
[83] P. Anderson and W. Bauer, *Biochemistry* **17,** 594 (1978).
[84] M. L. Bleam, C. F. Anderson, and M. T. Record, Jr., *Proc. Natl. Acad. Sci. U.S.A.* **77,** 3085 (1980).
[85] J.-H. Ha, Ph.D. Thesis, University of Wisconsin, Madison, Wisconsin (1990).
[86] M. D. Barkley, P. A. Lewis, and G. E. Sullivan, *Biochemistry* **20,** 3842 (1981).

Therefore, in the absence of preferential anion interactions, the relationship in Eq. (13) should hold:

$$\frac{\partial \log K_{obs}}{\partial \log[M^{2+}]} = (\phi/\psi)\frac{\partial \log K_{obs}}{\partial \log[M^{+}]} \quad (13)$$

For duplex B-form DNA, $\phi = 0.47$, $\phi/\psi = -0.53$,[26] whereas for single-stranded poly(U), $\phi = 0.425$, $\phi/\psi = -0.60$. If Eq. (13) does not hold for a particular ligand–nucleic acid interaction, then this suggests that salt effects other than those due to counterion release from the nucleic acid should be considered. These could result from preferential anion or cation interactions with the protein (see below).

In buffers containing a mixture of monovalent, M^+, and divalent, M^{2+}, cations, the dependence of K_{obs} on $[M^+]$ is complicated owing to the competition between monovalent and divalent cations for binding to the nucleic acid.[20,78] In this case, Eqs. (9) and (10) no longer apply, and Eq. (9) is replaced by Eq. (14):



$$L^{z+} + D \rightleftharpoons LD + xM^{+} + yM^{2+} \quad (14)$$

In the absence of preferential ion interactions with the ligand, the result of the M^+/M^{2+} competition for the nucleic acid will be 2-fold: (1) the value of K_{obs}, as well as the absolute value of the monovalent salt dependence of K_{obs}, $|\partial \log K_{obs}/\partial \log[M^+]|$, will be lower than the values in the absence of divalent cations; and (2) $\partial \log K_{obs}/\partial \log[M^+]$ will no longer be independent of M^+ concentration, that is, a plot of $\log K_{obs}$ versus $\log[M^+]$ will display curvature. Both of these effects result from the fact that $\partial \log K_{obs}/\partial \log[M^+]$ measures only the release of M^+ [the stoichiometric coefficient x in Eq. (14)]. Therefore, when both M^+ and M^{2+} are bound to the nucleic acid, less M^+ will be released on formation of the ligand–nucleic acid complex. Furthermore, the binding of M^{2+} itself will be a function of $[M^+]$, since it also is a positively charged ligand; therefore, the coefficients x and y in Eq. (14) will change with $[M^+]$, resulting in a nonlinear plot of $\log K_{obs}$ versus $\log[M^+]$.[20,78]

B. Cation and Anion Effects on Protein–Nucleic Acid Equilibria

1. General Ion Effects. In general, the equilibrium binding constants for protein–nucleic acid interactions will display a more complicated range of effects owing to changes in salt concentration than is observed for simple positively charged oligocations. This is due to the fact that cations, anions, protons, and water interact with most proteins; hence, preferential interactions of these ions with the protein must be considered in addition

to the preferential interactions of cations and water with the nucleic acid. These additional preferential interactions are indicated in Eq. (15),

$$P + D \rightleftharpoons PD + \Delta c\, M^{+} + \Delta a\, X^{-} + \Delta h\, H^{+} + \Delta w\, H_2O \quad (15)$$

where the coefficients Δc, Δa, Δh, and Δw represent the net preferential interactions of cations, anions, protons, and water. The preferential interaction parameter for cation binding is defined as $\Delta c = (c_{PD} - c_P - c_D)$, where c_i represents the moles of cations bound thermodynamically per mole of species i; Δa, Δh, and Δw are defined similarly. As defined, the preferential interaction parameters include contributions from both direct binding and nonideality (activity coefficient) effects. The dependence of K_{obs} on monovalent salt concentration at constant temperature, pressure, and pH for this case is then given by[19]

$$\left(\frac{\partial \log K_{obs}}{\partial \log[MX]}\right) = -(\Delta c + \Delta a) + \frac{2[MX]}{[H_2O]}\Delta w \quad (16)$$

The justification for the use of monovalent salt concentration, rather than mean ionic activities, in Eq. (16) has been discussed.[18,19] We note that the individual terms Δc, Δa, and Δw can be either positive or negative, indicating a net release or uptake of that species, respectively; however, for positively charged ligands, cation release from the nucleic acid will always be a major component of Δc. The terms Δc and Δw can have contributions from both the protein and the nucleic acid, whereas any anion effects will be due to preferential interactions with the protein. We also note that each of the terms Δc, Δa, and Δw represent net preferential interactions. Therefore, one can imagine the case in which a net release of cations on formation of PD can have separate contributions owing to release of cations from the nucleic acid, which is partially offset owing to an uptake of cations by the protein. This appears to be the case for *E. coli* SSB binding to single-stranded nucleic acids in its $(SSB)_{65}$ binding mode.[87] We also note that Δa and Δc are not constrained by any electroneutrality relationship and can assume any values, depending on the particular interaction.

In general, each of the coefficients in Eq. (15) can vary with the bulk salt concentration, hence the derivative ($\partial \log K_{obs}/\partial \log[MX]$) can also be a function of salt concentration. The one exception to this is the contribution owing to cation release from the nucleic acid, which seems to be constant over a wide range of salt concentration[18,22,71,73] as discussed above. However, for a particular protein–nucleic acid interaction, Eq. (16) can usually

[87] L. B. Overman, W. Bujalowski, and T. M. Lohman, *Biochemistry* **27**, 456 (1988).

be simplified, since not all of the species are likely to interact preferentially. In particular, except at high salt concentrations (≥0.5 *M*), the term arising from preferential hydration is expected to be negligible when compared to the ion release terms.

For simple oligocations,[73] such as polyamines and oligolysines, the net positive charge on the oligocation determines the salt dependence of K_{obs}; however, this is definitely not the case for proteins. This is obvious from the fact that many nucleic acid-binding proteins have a net negative charge at neutral pH (p*I* < 7) and yet still bind tightly to DNA with a net release of cations. An illustration of this point is shown in Fig. 4B for the interaction of the *E. coli* SSB tetramer with poly(U) to form the $(SSB)_{65}$ binding mode. The salt dependence of K_{obs} in NaF is essentially independent of pH over the range from 5.5 to 9.0, even though the net charge of the SSB tetramer should vary drastically over this range, perhaps even changing sign since the protein has a p*I* of approximately 6.0.[88]

The dependences of K_{obs} on salt concentration for the various ligand–nucleic acid interactions discussed in the previous sections are not due to simple ionic strength effects. The sole use of ionic strength as a parameter to describe the effects of salt concentration on ligand–nucleic acid interactions, or any macromolecular interaction for that matter, implies that the effects of salt are due only to long-range screening effects, without contributions from the direct binding of ions. Although the ionic strength of a solution is the important variable for the description of interactions among ions in simple salt solutions as indicated in Debye–Hückel theory,[89–91] its use to describe salt effects on ligand–nucleic acid interactions is incorrect. This is most clear from the fact that dramatically different values of K_{obs} are obtained in buffers containing monovalent cations versus divalent cations, even though the solutions are at the same "ionic strength." Therefore, theoretical approaches that attempt to explain salt effects on ligand–nucleic acid interactions, or other macromolecular interactions, solely in terms of ionic strength effects cannot be correct, quantitatively.

Decomposition of Salt Dependence into Contributions from Anions versus Cations. To distinguish the relative contributions to the salt dependence of K_{obs} for a ligand–nucleic acid interaction, arising from preferential cation versus anion interactions, a number of experiments can be performed. First, the effects on K_{obs} and $(\partial \log K_{obs}/\partial \log[MX])$ of a series of

[88] J. Weiner, L. Bertsch, and A. Kornberg, *J. Biol. Chem.* **250**, 1972 (1975).
[89] P. Debye and E. Huckel, *Physik. Z.* **24**, 185 (1923).
[90] P. Debye and E. Huckel, *Physik. Z.* **24**, 384 (1923).
[91] P. Debye and E. Huckel, *Physik. Z.* **25**, 97 (1924).

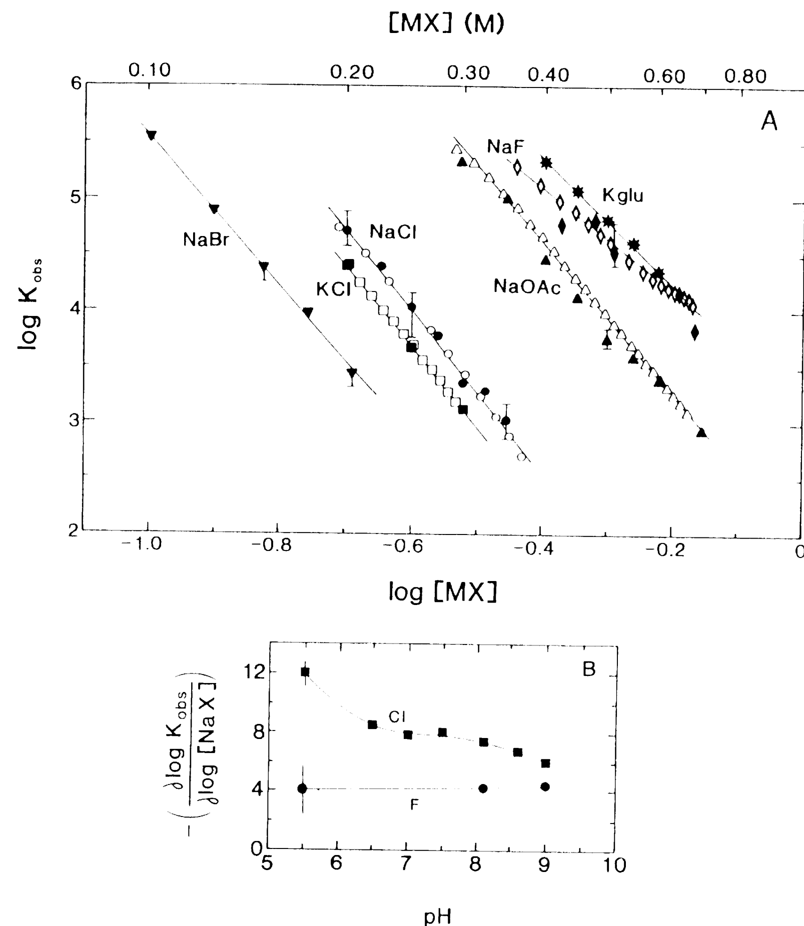


FIG. 4. (A) Dependence of log K_{obs} on total monovalent cation concentration, M^+, for a series of salts differing in anion type, for the equilibrium binding of the *E. coli* SSB tetramer to poly(U) to form the $(SSB)_{65}$ binding mode (25.0°, pH 8.1). Linear least-squares lines are drawn through each set of data. (Data from Overman *et al.*[87]) (B) Dependence of $(\partial \log K_{obs}/\partial \log[NaX])$ on pH for the equilibrium binding of the *E. coli* SSB tetramer to poly(U) to form the $(SSB)_{65}$ binding mode (25.0°) in NaF and NaCl. [Data from L. B. Overman, Ph.D. Thesis, Texas A & M University, College Station, Texas (1989).]

monovalent salts which differ only in the anion type (e.g., KBr, KCl, KCH_3CO_2, KF) can be examined. The interactions of anions with proteins generally follows the Hofmeister series ($F^- < CH_3CO_2^- < Cl^- < Br^-$),[19,92-94] with fluoride showing the weakest preferential interaction.[87] Figure 4A illustrates the dramatic effect of different anions on the interaction of the *E. coli* SSB tetramer with poly(U). In contrast, no effects of anions have been observed for the interaction of simple, short oligolysines ($z \leq 10$) with either duplex or single-stranded nucleic acids.[73,78] This approach should be coupled with changing the pH, since anion binding sites on the protein are generally titratable. Fig. 4B shows data from such an approach for the interaction of the *E. coli* SSB tetramer with poly(U) to form the $(SSB)_{65}$ binding mode.[95] At 25°, the salt dependence in NaCl ($\partial \log K_{obs}/\partial \log[NaCl]$) is observed to change dramatically with pH, from a value of -6.7 ± 0.2 at pH 8.6 to -12.0 ± 0.8 at pH 5.5. However, ($\partial \log K_{obs}/\partial \log[NaF]$) equals -4.4 ± 0.7, independent of pH. This indicates that the preferential interaction of Cl^- increases with decreasing pH. Furthermore, F^- has a low preferential interaction, and the salt dependence in NaF appears to reflect only the net Na^+ release on complex formation.

Second, the dependence of K_{obs} on monovalent cation concentration, $\partial \log K_{obs}/\partial \log[M^+]$, can be compared to its dependence on divalent cation concentration, $\partial \log K_{obs}/\partial \log[M^{2+}]$. If Eq. (13) does not hold, then preferential interactions of either cations or anions with the protein are indicated.

Third, mixtures of salts with a common anion but different monovalent and divalent cations (e.g., KCl and $MgCl_2$) can be used to vary the anion concentration over a wider range than is accessible with a single salt type. The use of this approach generally requires that preferential cation interactions with the protein are negligible, so that the effect of the divalent cation, M^{2+}, is only due to its competition with monovalent cation, M^+, for binding to the nucleic acid.[20,78]

V. Obtaining Thermodynamic Parameters by van't Hoff Analysis

Thermodynamic parameters for a ligand–nucleic acid interaction can be obtained from studies of the temperature dependence of the equilibrium

[92] F. Hofmeister, *Arch. Exp. Pathol. Pharmakol.* **24,** 247 (1888).
[93] P. H. von Hippel and T. Schleich, *in* "Biological Macromolecules" (G. Fasman and S. Timasheff, eds.), p. 417. Dekker, New York, 1969.
[94] K. D. Collins and M. W. Washabaugh, *Q. Rev. Biophys.* **18,** 323 (1985).
[95] L. B. Overman, Ph.D. Thesis, Texas A&M University, College Station, Texas (1989).

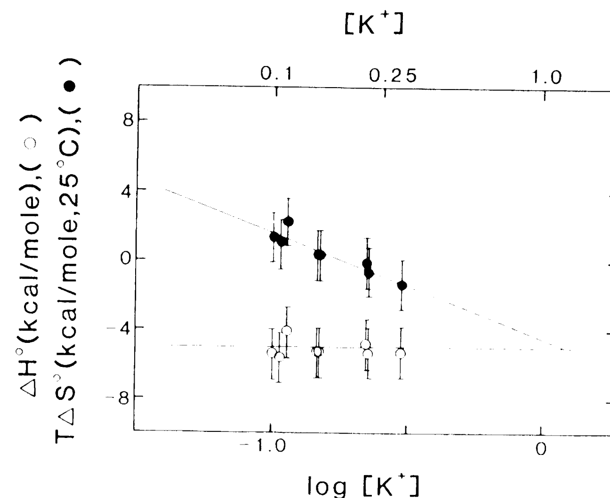


FIG. 5. Dependence of ΔH°_{obs} and $T\Delta S^\circ_{obs}$ on $\log[K^+]$ for the interaction of KWK_4-NH_2 with poly(U) (25.0°, pH 6.0, KCH_3CO_2). The values of ΔH°_{obs} at each salt concentration were determined from van't Hoff analysis.[68] Values of $T\Delta S^\circ_{obs}$ were calculated from $T\Delta S^\circ_{obs} = \Delta H^\circ_{obs} + RT \ln K_{obs}$.

constant, K_{obs}, as indicated by

$$[\partial \ln K_{obs}/\partial(1/T)]_P = -\Delta H^\circ_{obs}/R \quad (17)$$

Therefore, the slope of a plot of $\ln K_{obs}$ versus $1/T$ (van't Hoff plot) yields the standard enthalpy change, ΔH°_{obs}, for the reaction at that temperature. If ΔH°_{obs} is independent of temperature, then the van't Hoff plot will be linear.

For simple oligocations, such as polyamines and oligolysines, which bind exclusively electrostatically, ΔH°_{obs} is nearly zero and independent of salt concentration. This reflects a near balance between the enthalpy change associated with neutralization of a phosphate by a bound counterion (M^+) versus a positive charge on the oligocation.[18,78,79] Figure 5 shows that independence of ΔH° on KCH_3CO_2 concentration for the equilibrium binding of a positively charged oligopeptide, L-Lys-L-Trp-(L-Lys)$_4$-NH_2 (KWK_4-NH_2), to poly(U).[68] In this case, the salt dependence of ΔG° is due solely to the salt dependence of ΔS°. The slightly negative value of ΔH°, -5 ± 1 kcal/mol, is due to the interaction of the Trp residue.

In general, for a protein–nucleic acid interaction, it is difficult to dissect ΔH°_{obs} into its separate contributions from hydrogen bonds, ionic, hydrophobic interactions, etc. This is primarily because ΔH°_{obs}, as with all

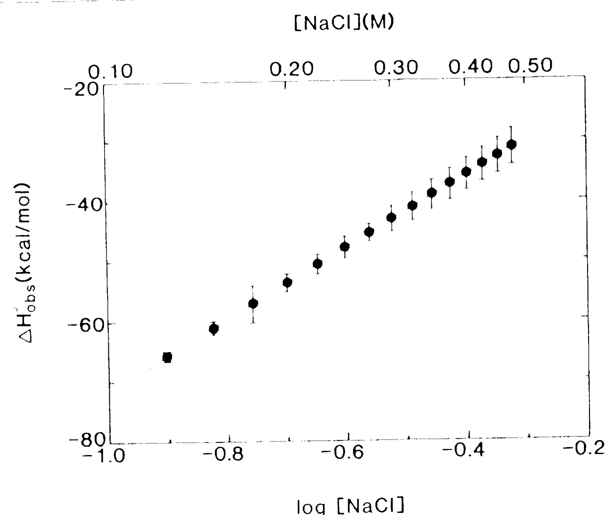


FIG. 6. Dependence of the van't Hoff enthalpy, ΔH°_{obs}, on log[NaCl] for the equilibrium binding of the *E. coli* SSB tetramer to poly(U) to form the $(SSB)_{65}$ binding mode (pH 8.1). The linear least-squares line through the data is $\Delta H^{\circ}_{obs} = 60 \log[NaCl] - 12$ (kcal/mol). (Data from Overman.[95])

thermodynamic quantities, reflects all changes that occur within the system, not simply those occurring locally at the protein–nucleic acid interface. Furthermore, ΔH°_{obs} (as well as ΔS°_{obs}) is not generally constant, but rather is a function of solution conditions (e.g., salt and pH). This can result if protonation or deprotonation events, which can also be linked to ion binding, are coupled to the formation of the ligand (protein)–nucleic acid complex. Therefore, just as for K_{obs} (ΔG°_{obs}), important information about the thermodynamics of the interaction will be obtained from studies of the dependence of ΔH°_{obs} on solution variables, whereas a determination of ΔH°_{obs} under only one set of solution conditions provides limited information. One must examine how ΔH°_{obs} varies with solution conditions (pH, salt concentration and type, temperature) in order to begin to determine the various contributions to complex stability. For example, the ΔH°_{obs} for the interaction of the *E. coli* SSB tetramer with poly(U) to form the $(SSB)_{65}$ binding mode is quite dependent on the NaCl concentration, ranging from -61 kcal/mol (SSB tetramer) at 0.15 *M* NaCl to -31 kcal/mol at 0.48 *M* NaCl as shown in Fig. 6.[95] This seems to reflect the fact that formation of the $(SSB)_{65}$ complex with poly(U) is coupled to preferential ion binding to the protein, and the SSB ion binding sites are titratable.[95] Clearly, an

understanding of the origins of the ΔH°_{obs} for this interaction would be impossible, based on measurements at only a single NaCl concentration.

The temperature dependence of ΔH°_{obs} is related to the change in heat capacity for the binding reaction, ΔC°_{P}, as in Eq. (18):

$$(\partial \Delta H^{\circ}_{obs}/\partial T)_{P} = \Delta C^{\circ}_{P,obs} \quad (18)$$

Therefore, information about the heat capacity change for a reaction can be obtained, in principle, from a study of the temperature dependence of K_{obs} (ΔG°_{obs}). In practice, it is often difficult to study a binding equilibrium over a wide enough range of temperature to obtain accurate values of $\Delta C^{\circ}_{P,obs}$. In these cases, $\Delta C^{\circ}_{P,obs}$ is better determined by calorimetric means, although there have yet been no studies of protein–nucleic acid interactions by calorimetry. In any event, a binding equilibrium that yields a linear van't Hoff plot indicates that $\Delta C^{\circ}_{P,obs}$ equals 0 for the reaction. In general, ligands that bind to nucleic acids in an entirely electrostatic mode will have values of ΔH°_{obs} and ΔC°_{P} close to zero, as has been observed for the interaction of simple oligocations with nucleic acids.[68,78] Significant values of $\Delta H^{\circ}_{obs} < 0$ have been measured for the nonspecific binding of *E. coli lac* repressor to duplex DNA,[26] as well as the nonspecific binding of the *E. coli* SSB protein to poly(U)[95]; however, in these cases, it appears that the nonzero ΔH°_{obs} results from a linkage of binding to protonation of residues on the protein. On the other hand, significant curvature has been observed in van't Hoff plots for a number of sequence-specific protein–DNA interactions (*lac* repressor, *Eco*RI endonuclease, RNA polymerase), indicating $\Delta C^{\circ}_{P} \ll 0$ on formation of these complexes. It has been suggested that this large $\Delta C^{\circ}_{P} \ll 0$ reflects the hydrophobic effect,[96,97] that is, the removal of large regions of the protein from contact with water, resulting in the release of bound water, which contributes a large favorable entropy change ($\Delta S^{\circ}_{obs} \gg 0$). The extent to which these large heat capacity effects are general is not yet clear.

VI. Summary

Ligand– and protein–DNA equilibria are extremely sensitive to solution conditions (e.g., salt, temperature, and pH), and, in general, the effects of different solution variables are interdependent (i.e., linked). As a result, an assessment of the basis for the stability and specificity of ligand– or protein–DNA interactions requires quantitative studies of these

[96] J.-H. Ha, R. S. Spolar, and M. T. Record, Jr., *J. Mol. Biol.* **209**, 801 (1989).

[97] R. S. Spolar, J.-H. Ha, and M. T. Record, Jr., *Proc. Natl. Acad. Sci. U.S.A.* **86**, 8382 (1989).

interactions as a function of a range of solution variables. Many of the most dramatic effects on the stability of these interactions result from changes in the entropy of the system, caused by the preferential interaction of small molecules, principally ions which are released into solution on complex formation. A determination of the contributions of these entropy changes to the stability and specificity of protein– and ligand–DNA interactions requires thermodynamic approaches and cannot be assessed from structural studies alone.

Acknowledgments

We thank Tom Record for sharing unpublished data from his laboratory and for many discussions of aspects of this material and Lisa Lohman for help in preparing the figures. Research from the author's laboratory was supported by grants from the National Institutes of Health (GM39062, GM30498) and the American Cancer Society (NP-756). T.M.L. is a recipient of American Cancer Society Faculty Research Award FRA-303.

[25] Nonspecific Ligand–DNA Equilibrium Binding Parameters Determined by Fluorescence Methods

By TIMOTHY M. LOHMAN and DAVID P. MASCOTTI

I. Introduction

Spectroscopy provides sensitive and convenient methods to monitor protein (ligand)–nucleic acid interactions. In this chapter we discuss the use of spectroscopic probes to monitor nonspecific ligand–nucleic acid interactions as well as rigorous methods of analysis to obtain model-independent equilibrium binding isotherms from spectroscopic titrations.[1-3] Such isotherms can then be analyzed, using appropriate statistical thermodynamic models, to yield equilibrium binding parameters for the interaction. Studies of the dependence of the binding parameters on temperature and other solution conditions can then provide thermodynamic information, which is necessary to understand the basis for stability and specificity of the interactions.[4] See also [24] in this volume for further discussion.[5]

[1] C. J. Halfman and T. Nashida, *Biochemistry* **11,** 3493 (1972).
[2] W. Bujalowski and T. M. Lohman, *Biochemistry* **26,** 3099 (1987).
[3] T. M. Lohman and W. Bujalowski, this series, Vol. 208, p. 258.
[4] M. T. Record, Jr., J.-H. Ha, and M. A. Fisher, this series, Vol. 208, p. 291.
[5] T. M. Lohman and D. P. Mascotti, this volume [24].

Copyright © 1992 by Academic Press, Inc.
All rights of reproduction in any form reserved.

This chapter focuses on the use of steady-state fluorescence techniques to monitor changes in the ligand (protein) that accompany binding, with examples drawn from studies of ligands and proteins that bind nonspecifically to nucleic acids. Spectroscopic methods, such as fluorescence, UV absorbance, and circular dichroism, although indirect, offer many advantages for the study of ligand–nucleic acid as well as other macromolecular interactions. In the case of steady-state fluorescence, one has access to a rapid, nonradioactive method, which can be used at fairly low concentrations of ligand and nucleic acid. For example, micromolar concentrations of tryptophan can be detected easily, hence binding studies with proteins or peptides containing multiple tryptophans can often be performed at concentrations in the 10 n*M* range. A further general advantage of spectroscopic methods is that titrations of a single solution can be monitored continuously, thus increasing the precision and ease of data collection. These are essential considerations if one is to attempt to investigate systematically the thermodynamics of a protein– or ligand–nucleic acid interaction over a wide range of solution variables. Finally, although assumptions regarding the relationship between the spectroscopic signal and the degree of binding are often made when using spectroscopic techniques, these assumptions are not necessary, since thermodynamic, model-independent methods of analysis are available which enable the determination of absolute binding isotherms.[1-3] Therefore, rigorous investigations of the equilibrium binding and thermodynamic properties of ligand (protein)–nucleic acid interactions that rely on convenient spectroscopic probes are possible.

II. Model-Independent Determination of Ligand–Nucleic Acid Binding Isotherms Using Spectroscopic Approaches to Monitor Binding

A variety of experimental approaches have been described for the study of protein–nucleic acid interactions[6]; however, we shall focus on the use of spectroscopic approaches, specifically changes in steady-state fluorescence, that accompany formation of the complex. To use a change in a spectroscopic signal that is induced on formation of a ligand–nucleic acid complex to obtain a true equilibrium binding isotherm, either the relationship between the signal change and the degree of binding must be known, or a method of analysis must be used that does not require knowledge of this relationship. If some relationship between the signal change and the degree of binding is assumed (e.g., linear), then the resulting

[6] A. Revzin, *in* "Nonspecific DNA–Protein Interactions" (A. Revzin, ed.), p. 5. CRC Press, Boca Raton, Florida, 1990.