Ion Effects on Ligand–Nucleic Acid Interactions

M. Thomas Record, Jr, Timothy M. Lohman and Piter de Haseth

Department of Chemistry
University of Wisconsin, Madison, Wisc. 53706, U.S.A.

(Received 10 February 1976, and in revised form 15 July 1976)

We have developed a general thermodynamic analysis of monovalent ion effects on the observed association constants \( K_{\text{obs}} \) of ligand–nucleic acid interactions. Our approach is based on the binding theory of Wyman (1964) and the polyelectrolyte theory of Manning (1969). In the case of model ligands such as \( \text{Mg}^{2+} \) or short oligolysines, where there is no anion binding by the ligand, the dependence of \( K_{\text{obs}} \) on monovalent ion (\( M^+ \)) concentration results from the release of \( M^+ \) counterions from the nucleic acid in the association reaction. We find that, for these systems, \( \log K_{\text{obs}} \) is a linear function of \( \log [M^+] \). The slope of such a graph yields the number of charge interactions, or ion pairs, formed between ligand and nucleic acid; the intercept of a linear extrapolation to a 1 mM\( M^+ \) standard state yields the non-electrostatic component of the binding free energy. From an analysis of the data of Latt & Sober (1967) on the interactions of oligolysines with polyribonucleotides, we have concluded that the dominant factor driving complex formation between these charged ligands and the nucleic acid is the entropic contribution from the release of counterions. Counterion release also appears to drive the non-specific interactions of proteins with nucleic acids.

1. Introduction

An understanding of the thermodynamic and mechanistic aspects of protein–nucleic acid interactions is central to the elucidation of the molecular basis of the control of gene expression. The problem may be approached from various viewpoints, utilizing kinetic, structural, and thermodynamic studies. In this paper we give a thermodynamic analysis and interpretation of the apparent association constants (\( K_{\text{obs}} \)) of protein–nucleic acid complex formation.

A characteristic property of \( K_{\text{obs}} \) for protein–nucleic acid interactions is its substantial dependence on ion concentrations. For example, Riggs et al. (1970a,b) have found that the binding constant for the interaction of lac repressor protein with lac operator DNA decreases by two orders of magnitude when the KCl concentration in the medium is increased from 0.01 M to 0.18 M. (This work was carried out in the presence of 0.01 M\( \text{MgCl}_2 \); in the absence of the competitive ligand \( \text{Mg}^{2+} \) the effect of \( K^+ \) on the association constant would be still more dramatic.) Similar observations have been made by Latt & Sober (1967) on the interaction of oligolysines with polyribonucleotides, by Poliakow et al. (1972) on the interaction of histones with DNA, by Yamamoto & Alberts (1974) on the binding of the estradiol receptor protein to DNA, and by Jensen & von Hippel (1976) on interactions of RNase and
bacteriophage T4 gene 32 protein with native and denatured DNA. Moreover, in the course of investigating the non-specific interactions of core and holoenzyme forms of RNA polymerase with single and double-stranded DNA, we (de Haseth, Lohman, Record, Wensley & Burgess, manuscript in preparation) have found a large effect of monovalent counterion concentration on these association constants. The non-specific interaction of lac repressor with DNA is similarly sensitive to salt (Revzin & von Hippel, personal communication; de Haseth & Lohman, unpublished experiments). Lin & Riggs (1972,1975) have observed a reduced (though still substantial) effect of monovalent counterion concentration on this interaction in the presence of 0.003 M-MgCl₂. It is the purpose of this paper to develop a general interpretation of these phenomena in terms of modern polyelectrolyte theory (Manning, 1969,1972) and binding theory (Wyman, 1964; Schellman, 1975). In subsequent papers we will apply the theory to our results and literature data on the interactions of RNA polymerase and lac repressor with specific and non-specific sites on DNA.

Our primary focus is on the electrostatic component of the binding reaction, although an evaluation of this component permits us to discuss the non-electrostatic component as well. By the electrostatic component we mean (1) direct ion pair formation between positively charged groups on the ligand and negative phosphate groups on the nucleic acid, and (2) concomitant counterion release from the nucleic acid (and possibly from the ligand) on binding. From our analysis we can obtain an accurate estimate of the number of such ion pairs involved in the interaction, and we can evaluate the entropic contribution of counterion release to the free energy of association. The counterion release effect makes a major contribution to the overall observed binding free energy under the typical ionic conditions of systems in vivo or in vitro.

Latt & Sober (1967) and Daune (1972) have developed aspects of the theory of ion effects on ligand–nucleic acid interactions in previous work. Latt & Sober (1967) recognized the importance of Na⁺ interaction with the polynucleotide. Daune (1972) extended their theoretical treatment to obtain an estimate of the number of ion pairs involved in such interactions. Our present approach utilizes Manning’s (1969) polyelectrolyte theory to treat both the binding and screening interactions of low molecular weight ions with nucleic acids. The resulting theoretical expression for the variation in observed association constant with monovalent counterion concentration has a simple form, and is directly interpretable in terms of the molecular reaction stoichiometry. The analogy to the general binding theory of Wyman (1964) is then clear.

2. Theoretical

(a) Interaction of Na⁺ or K⁺ with nucleic acids

Manning (1969,1972) has shown that simple monovalent counterions M⁺ (typically Na⁺ or K⁺) interact with a polyelectrolyte in two ways. There is a direct condensation of counterions onto the polaron to reduce the axial charge density to less than 0.14 net charge per ångström unit (1 charge/7.15 Å). The extent of this condensation varies with the initial axial charge density of the polyelectrolyte. After condensation, the remaining unneutralized polyelectrolyte charges are screened from each other by an atmosphere of mobile counter- and co-ions. Both condensation (c) and screening (s) effects contribute to what we define as the thermodynamic extent of counterion
“binding”, \( \psi \), expressed as the fraction of a counterion bound in the thermodynamic sense per polion charge:

\[
\psi = \psi_c + \psi_s.
\]

Both \( \psi_c \) and \( \psi_s \) are readily calculated for any polyelectrolyte of known axial charge spacing \( b \). Manning defines a parameter \( \xi \) by

\[
\xi = \frac{e^2}{\epsilon k T b}
\]

where \( e \) is the magnitude of the electronic charge, \( k \) is Boltzmann's constant, \( T \) the absolute temperature in deg. K., \( \epsilon \) is the bulk dielectric constant, and \( b \) is the axial phosphate distance along the polynucleotide. If \( \xi > 1 \) for a polion, then counterion condensation will occur to reduce the net value of \( \xi \) to unity. Thus

\[
\psi_c = 1 - \xi^{-1}
\]

gives the fraction of a counterion condensed per original polion charge. Furthermore we will demonstrate that (see eqns (11) and (12))

\[
\psi_s = (2\xi)^{-1}.
\]

Therefore,

\[
\psi = \psi_c + \psi_s = 1 - (2\xi)^{-1}.
\]

A rigorous thermodynamic derivation of equation (5) is given in the Appendix.

Values of \( \xi \) for native and single-stranded DNA, as well as for various synthetic polynucleotides, have been evaluated by Record et al. (1976) from structural data or by a comparison of theoretical and experimental variations of the helix-coil transition temperature \( T_m \) with monovalent counterion concentration \( M^+ \), i.e.

\[
\frac{dT_m}{d\log [M^+]} = \frac{2.3 R T_m^2}{\Delta H^{\text{obs}}_{\text{helix-coil}}} (\psi_{\text{helix}} - \psi_{\text{coil}}).
\]

For the double-helical form of DNA, with two phosphates per 3-4 Å, we calculate that \( \xi = 4.2 \) and \( \psi = 0.88 \), or 0.88 of a counterion is associated in the thermodynamic sense with each phosphate of helical DNA in solution.

(b) General formalism for a ligand–nucleic acid interaction

Consider the interaction of a charged ligand (L) with a nucleic acid (P) in which the association involves the formation of \( m' \) ion pairs. We will assume that the interactions are with \( m' \)† adjacent phosphates, for simplicity. For ligands with small enough binding sites (e.g. \( \text{Mg}^{2+} \), oligolysines), we neglect anion binding to the ligand. Moreover, we assume that the solution is sufficiently dilute in all solute components (including low molecular weight electrolyte) that differential hydration effects are unimportant. Generalization of the analysis to include anion binding and hydration effects can be performed without difficulty if warranted. With these assumptions, the molecular binding reaction is written as:

\[
\text{L} + \text{P (m' sites)} \xrightarrow{k} \text{LP} + m' \psi_c \text{ M}^+,
\]

† Jensen et al. (1976) define a parameter \( m \) as the number of nucleotide units which participate in binding the ligand to the nucleic acid. Our parameter \( m' \) is a measure of the contribution to \( m \) from electrostatic interactions (ion pairs).
where $K$ is a thermodynamic equilibrium constant which is a function of temperature and pressure only and not of ion concentrations. (As written, eqn (6) is applicable to a reaction which does not involve protonation of titratable groups; this effect can be included in a straightforward manner when necessary.)

It follows from equation (6) that

$$K = \frac{a_{LP} \sigma_{M}^{\psi_{o}}}{a_{L} a_{P}} \frac{[LP]}{[L]} \frac{[M^+]^{\sigma_{M}^{\psi_{o}}} \gamma_{L}^{M^+} \gamma_{LP}^{L}}{[P] \gamma_{L} \gamma_{P}}, \quad (7)$$

where the $a$ terms represent activities, the $\gamma$ terms represent activity coefficients, and quantities in brackets denote concentrations. As experimentally determined, the observed equilibrium constant $K_{obs}$ is written as

$$K_{obs} = \frac{[LP]}{[L] [P]} \quad (8)$$

Equation (7) can now be written as

$$K = K_{obs} [M^+]^{\sigma_{M}^{\psi_{o}}} \gamma_{L}^{M^+} \gamma_{LP}^{L}. \quad (9)$$

At constant temperature and pressure, the variation of $K_{obs}$ with $M^+$ concentration is consequently

$$\frac{\partial \log K_{obs}}{\partial \log [M^+]} = m' \psi_{o} + \frac{\partial \log \left(\gamma_{L}^{M^+} \gamma_{LP}^{L}\right)}{\partial \log [M^+]} - \frac{\partial \log \gamma_{P}}{\partial \log [M^+]} \quad (10)$$

All partial derivatives are at constant temperature and pressure, unless specified otherwise.

The last term in equation (10) can be evaluated from polyelectrolyte theory in the Debye–Hückel approximation (Manning, 1969; Record, 1975). The activity coefficient of the DNA binding site of $m'$ phosphates is

$$\ln \gamma_{P} = m' \frac{\sigma_{M}^{\psi_{o}}}{RT} = -m' \xi^{-1} \ln \kappa, \quad (11)$$

where $\sigma_{M}^{\psi_{o}}$ is the electrostatic free energy per mole of nucleotide phosphates, and $\kappa$ is the Debye–Hückel screening parameter ($\kappa$ is proportional to the square root of the ionic strength). Under conditions of excess $M^+$, where the ionic strength is equivalent to the $M^+$ concentration, equation (11) yields

$$\frac{\partial \log \gamma_{P}}{\partial \log [M^+]} = -m'(2\xi)^{-1} = \psi_{o} \quad (12)$$

With equation (12), equation (10) now reduces to

$$\frac{\partial \log K_{obs}}{\partial \log [M^+]} = m' \psi_{o} + \psi_{o} + \frac{\partial \log \left(\gamma_{L}^{M^+} \gamma_{LP}^{L}\right)}{\partial \log [M^+]} \quad (13)$$

Finally, we expect the activity coefficient ratio $(\gamma_{M}^{\psi_{o}} \gamma_{LP}^{L})/\gamma_{L}$ to be relatively insensitive to $[M^+]$. This can be theoretically verified in the case of model ligands such as $Mg^{2+}$ or short oligonucleotides in the region of validity of the Debye–Hückel approximation for single ion activity coefficients, i.e. $\log \gamma_{L} = AZ_i I^t$, where $Z_i$ is the valence of species $i$, $I$ is the ionic strength and $A$ is a numerical constant. In the
cases of more complex ligands or of ion concentrations above the range of the Debye- Hückel limiting law, which will be of interest here, constancy of the above activity coefficient ratio is a plausible but unproven assumption.

With these simplifications, along with equation (1), equation (13) reduces to

$$\frac{\partial \log K_{\text{obs}}}{\partial \log [M^+]} = m'\psi. \quad (14)$$

Since $\psi$ is known for the particular DNA involved in the interaction (Table 1), determination of the dependence of $K_{\text{obs}}$ on $M^+$ leads directly to a value for $m'$, the number of ion pairs formed. Note that with the definition of $\psi$ as the thermodynamic $M^+$ binding parameter, equation (14) is of the form expected from Wyman’s (1964) binding theory.

If the anion of $M^+$ is bound by the ligand, so that $k$ anions are displaced in the ligand–nucleotide association reaction, then equation (14) is replaced by

$$\frac{\partial \log K_{\text{obs}}}{\partial \log [M^+]} = k + m'\psi. \quad (14a)$$

In the case of a polyelectrolyte ligand, such as polylysine, $k$ would include contributions from condensed and screening counterions, as does $\psi$. If experiments are carried out in concentrated salt solutions, then hydration terms must be included in equation (14) or (14a). Tanford (1969) has extended Wyman’s (1964) binding theory to include this effect. Unless the amount of water of hydration released in the association reaction is very large, hydration effects should be important only at electrolyte concentrations in the molar range.

### 3. Results

(a) $Na^+$ effects on the binding of $Mg^{2+}$ to polynucleotides

Krakauer (1971, 1974) has used a titration procedure in the presence of eriochrome black T as indicator to obtain isotherms for the binding of $Mg^{2+}$ to poly(A), poly(U) and their complexes. His binding constants, calculated in terms of the fraction of a $Mg$ ion bound per phosphate, $n$, show a dependence on both $Na^+$ concentration and $\nu$ for each polynucleotide studied. We have recalculated Krakauer’s data using the large ligand (overlap) model of McGhee & von Hippel (1974), and find that the
dependence of the observed binding constants on \( \nu \) can be eliminated (within experimental uncertainty) by letting each phosphate group be a potential binding site, but allowing a Mg ion to (effectively) cover more than one phosphate. Results of the overlap model calculation are given in Table 2; the physical site size (\( n \)) is in the range 2.0 to 2.7 phosphates, depending on the polymer studied. Representative equilibrium constants calculated from the overlap model, and the derivatives \( -\left( \frac{\partial \log K_{\text{obs}}}{\partial \log [\text{Na}^+] \right) \) obtained from the complete analysis of Krakauer's data (0.1 M \( \geq [\text{Na}^+] \geq 0.01 \) M), are also listed in Table 2. From the values of \( K_{\text{obs}} \), we calculate the ionic interaction parameter \( m' \) using equation (14). The values of \( m' \) cluster around 2, indicating that one \( \text{Mg}^{2+} \) binds to two phosphates, releasing 2\( \phi \) Na ions. The somewhat larger values of \( n \) indicate that generally saturation of the lattices occurs for \( \nu_{\text{max}}^{\text{exp}} = 1/n < 0.5 \). This result is in qualitative accord with Manning's (1969) theory, which predicts that the maximum amount of divalent ion binding to any polyelectrolyte is given by

\[
\nu_{\text{max}}^{\text{theo}} = \frac{1}{2} \left( 1 - (2\phi)^{-1} \right).
\]

Theoretical and experimental values of \( \nu_{\text{max}} \) are also given in Table 2. With the exception of poly(U), the two numbers are in good agreement.

Many binding buffers used for the investigation of protein–nucleic acid interactions have contained \( \text{Mg}^{2+} \) in addition to a monovalent counterion (cf. Riggs et al., 1970a, b; Hinkle & Chamberlin, 1972). \( \text{Mg}^{2+} \) competes with the protein for binding sites on the DNA, and therefore the presence of \( \text{Mg}^{2+} \) reduces the binding constant \( K_{\text{obs}} \) of the protein–nucleic acid interaction. In addition, \( \text{Mg}^{2+} \) binding reduces the level of \( \text{Na}^+ \) condensation (Record, 1975) and so reduces the amount of \( \text{Na}^+ \) released upon binding the protein. Consequently, the \( \text{Na}^+ \) dependence of \( K_{\text{obs}} \) is substantially reduced in the presence of \( \text{Mg}^{2+} \). (\( \text{Mg}^{2+} \) may also have specific effects on the conformation or extent of aggregation of the protein ligand.) We reserve the analysis of these divalent ion effects for a subsequent communication, and consider here only situations in which the binding of ligand releases \( \text{Na}^+ \) or \( \text{K}^+ \) from the DNA.

(b) \( \text{Na}^+ \) effects on the interaction of oligolysines with polynucleotides

The studies by Latt & Sober (1967) on the interaction of oligolysines (~DNP:† Lys(Lys)ₙ, \( N = 3 \) to 8) with poly(A)·poly(U) and poly(I)·poly(C) provide the opportunity to test equation (14) under the relatively defined conditions of this basic polypeptide model system. Latt & Sober showed that at a given \( \text{Na}^+ \) concentration in the range 0-15 to 0-50 M, the logarithm of the observed binding constant increased linearly with \( N \), and that the graphs for the various \( \text{Na}^+ \) concentrations extrapolated to a common intercept at \( N = 0 \), which they identified as the binding constant of the ~DNP-lysine side chain. It therefore appears that only the charged ~NH₃⁺ groups, and not the N-terminal ~NH₂⁺ group, interact with the double helix. Latt & Sober interpreted the slopes \( \frac{\partial \log K_{\text{obs}}}{\partial N} \) in terms of a competitive binding equilibrium between the oligolysine and \( \text{Na}^+ \) ions, and stressed the importance of \( \text{Na}^+ \) binding in interpreting the oligolysine–nucleic acid binding constants. Here we are able to amend their interpretation by introducing the concept of \( \text{Na}^+ \) condensation (Manning, 1969). That is, the extent of \( \text{Na}^+ \) binding by the nucleic acid is determined only by its axial charge density and not by the \( \text{Na}^+ \) concentration in the solution.

In Figure 1, we have replotted \( \log K_{\text{obs}} \) as determined by Latt & Sober (1967) for

† Abbreviation used: DNP, 2,4-dinitrophenyl.
<table>
<thead>
<tr>
<th>Polynucleotide</th>
<th>n</th>
<th>log $K_{obs}$</th>
<th>$\frac{\delta \log K_{obs}}{\delta \log [Na^+]}$</th>
<th>$25^\circ C$</th>
<th>$m'$</th>
<th>$\frac{1}{n}$</th>
<th>$\frac{1}{n}$</th>
<th>$\frac{1}{n}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poly(A) · poly(U)</td>
<td>2.2±0.1</td>
<td>3.6±0.1</td>
<td>1.7±0.1</td>
<td>1.9±0.1</td>
<td>0.45±0.02</td>
<td>0.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poly(U)</td>
<td>2.0±0.1</td>
<td>2.9±0.1</td>
<td>1.2±0.1</td>
<td>1.8±0.2</td>
<td>0.50±0.03</td>
<td>0.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poly(A) · poly(U)$_2$</td>
<td>2.2±0.1</td>
<td>3.7±0.1</td>
<td>1.6±0.2</td>
<td>1.7±0.2</td>
<td>0.45±0.02</td>
<td>0.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poly(A)</td>
<td>2.7±0.3</td>
<td>3.5±0.1</td>
<td>1.6±0.2</td>
<td>2.0±0.2</td>
<td>0.37±0.05</td>
<td>0.39</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the oligolysine–poly(A)·poly(U) interaction, versus log[Na⁺] for the various chain lengths N. The data can be fitted by a family of straight lines, constrained to intersect (in accordance with Latt & Sober's (1967) conclusion) at a common point. The slopes of these lines increase in magnitude in direct proportion to N:

$$\frac{\partial \log K_{obs}}{\partial \log [Na^+] = (0.9 \pm 0.05)N.}$$

(15)

Similar behavior is obtained for interactions of oligolysines with poly(I)·poly(C), except that all logK_{obs} values are slightly larger in magnitude.

To interpret equation (15) in terms of the formalism of equation (14), with ψ for poly(A)·poly(U) equal to 0.89 (Table 1), requires us to assume that there is no anion binding to the oligolysines. This assumption is in agreement with the finding of Latt & Sober (1967) that the addition of excess Cl⁻ (as tetramethyl ammonium chloride, where the tetramethyl ammonium cation is expected to interact with polynucleotides more weakly than Na⁺) did not substantially affect the interaction of oligolysines with poly(I)·poly(C). (We are investigating this point further as a function of oligolysine chain length and anion type.)

Comparison of equations (14) and (15) indicates that $N = m'$; that is, the number of ion pairs formed is equal to the number of $\varepsilon$-NH₄⁺ groups on the oligomer. The quantity

$$\frac{1}{N} \frac{\partial \log K_{obs}}{\partial \log [Na^+] \text{.}}$$
is not precisely constant within each series, but increases slightly with increasing $N$, from 0.85 at $N = 3$ to 0.94 at $N = 7$ for poly(A)·poly(U). More detailed analysis of the equilibrium suggests that this is a result of the variation in oligomer activity coefficient with the square of the oligomer charge. To a first approximation such effects can be neglected, as equation (15) indicates.

One can define a standard state for the oligomer–polynucleotide interaction by extrapolation of the lines of Figure 1 to a Na$^+$ concentration of 1 m. (More precisely, the calculations should involve activities and the extrapolation should be to unit activity of Na$^+$. Since activity coefficients of the oligomeric or polymeric ligands are unknown, we have not attempted this.) Extrapolations of the Latti–Sober data to 1 m-Na$^+$ are shown in Figure 1. To interpret the standard-state behavior of the oligolysine–polynucleotide complexes, we apply equation (14) in integrated form to this situation:

$$K_0 = K_{obs} [\text{Na}^+]^{n/n},$$

where $K_0$ is the value of $K_{obs}$ when the data are linearly extrapolated to 1 m-Na$^+$.

From equation (16)

$$\Delta G_0^0 = \Delta G_{obs}^0 + \Delta G_{Na^+},$$

where $\Delta G_0^0 = -RT \ln K_0$, $\Delta G_{obs}^0 = -RT \ln K_{obs}$ and $\Delta G_{Na^+} = -N_iRT \ln [\text{Na}^+]$. $\Delta G_{Na^+}$ gives the contribution to the thermodynamic quantity $\Delta G_0^0$ of the release of Na$^+$ in the binding reaction. Both $\Delta G_{obs}^0$ and $\Delta G_{Na^+}$ are functions of Na$^+$ concentration; $\Delta G_0^0$ is not, if activity coefficients of ligand and Na$^+$ can be neglected. Moreover, $\Delta G_{obs}^0$ is composed of contributions from the ε-DNP-lysine–polynucleotide interaction ($\Delta G_{DNPN}^0$) and from the N lysine–polynucleotide interactions ($N \Delta G_{DNP}^0$):

$$\Delta G_{obs}^0 = \Delta G_{DNPN}^0 + N \Delta G_{DNP}^0.$$ 

$\Delta G_{DNP}^0$ contains both electrostatic and non-electrostatic components; $\Delta G_{DNPN}^0$ should be non-electrostatic in origin. At [Na$^+$] = 1 m, $\Delta G_{Na^+} = 0$ and $\Delta G_0^0 = \Delta G_{obs}^0$. Also, at the Na$^+$ concentration where the lines of Figure 1 intersect, $\Delta G_{DNPN}^0 = -1.06$ kcal/mol for the interaction with poly(A)·poly(U). Therefore, if $\Delta G_{DNPN}^0$ is independent of Na$^+$ concentration, at 1 m-Na$^+$, we have

$$\Delta G_{obs}^0 (1 \text{ m-Na}^+) = -1.06 \text{ kcal} + N \Delta G_{DNP}^0 (1 \text{ m-Na}^+).$$

From the intercepts, we find $\Delta G_{DNP}^0 (1 \text{ m-Na}^+) = 0.22 \pm 0.01$ kcal. The interaction of a lysine group with poly(A)·poly(U) is intrinsically unfavorable at 1 m-Na$^+$, and becomes favorable at lower Na$^+$ concentrations only because of the contribution of the Na$^+$ release term. Per lysine–polynucleotide interaction the change in $\Delta G_{obs}^0$ per tenfold decrease in Na$^+$ concentration is $-1.15 \pm 0.05$ kcal at the experimental temperature of 4°C. Hence for an oligomer of five lysine residues but no ε-DNP-lysine group, we would predict that the extrapolation of a log–log plot of $K_{obs}$ versus [Na$^+$] (for binding to poly(A)·poly(U)) would yield a value of $K_0 = 0.14$ m$^{-1}$ at 1 m-Na$^+$. Furthermore, we would predict a value of $K_{obs} = 4.7 \times 10^3$ m$^{-1}$ at 0.1 m-Na$^+$, and $1.5 \times 10^8$ m$^{-1}$ at 0.01 m-Na$^+$. The conclusions in the case of poly(I)·poly(C) are analogous; here $\Delta G_{DNPN}^0 = -1.29$ kcal, evaluated at 0.74 m-Na$^+$, and $\Delta G_{DNP}^0 (1 \text{ m-Na}^+) = 0.14$ kcal.

We have examined the oligolysine interactions in some detail, not only as a test of the theory but also because the system shows the sort of division between electrostatic and non-electrostatic contributions which may be characteristic of such ligands.
as the lac repressor and RNA polymerase. The dominant electrostatic term is the free energy of release of Na⁺ ions; the effect of this term is minimized by working at high Na⁺ concentrations. The intrinsic free energy change of ion pair formation with the ligand (excluding counterion release contributions) must be small because the polynucleotide is already involved in an ionic interaction with Na⁺. The effect of the principal non-electrostatic interaction (here the ε-DNP interaction) is maximized at high Na⁺ concentration where, of course, the binding constant is small.

To analyze the thermodynamic basis of specificity in repressor–operator or polymerase–promoter interactions, the quantity \( -\langle \frac{\partial \log K_{\text{obs}}}{\partial \log [\text{Na}^+] \rangle} \) must be determined for the specific binding case and for binding to non-specific DNA sites. If these derivatives are the same for the two types of binding, then the number of electrostatic interactions is constant and specificity resides solely in the non-electrostatic component. If the derivatives are different, then the difference in the number of ion pairs can be evaluated directly, and extrapolation of the data to 1 m-Na⁺ will permit an estimation of the non-electrostatic contribution to specificity, if information regarding the anion binding parameter \( k \) is available. In making this estimate, we will use an average \( \Delta G_{\text{ups}}^o \) (1 m-Na⁺) of 0.18 ± 0.04 kcal to correct for the contribution of the ion pair at 1 m-Na⁺.

Brun et al. (1975) have investigated the interaction of the tripeptide Lys-Trp-Lys with poly(A) at pH 7, where the polymer is presumably single-stranded. From their data (Fig. 5 of Brun et al., 1975) we estimate that \( -\log K_{\text{obs}} = 1.8 \pm 2.4 \log [\text{Na}^+] \). Hence \( -\langle \frac{\partial \log K_{\text{obs}}}{\partial \log [\text{Na}^+] \rangle} \) = 2.4 and, since \( \phi_{\text{poly(A)}} = 0.78 \) (Table 1), we calculate that \( n' = 3 \). Evidently single-stranded poly(A) is sufficiently flexible so that all three positive charges on the tripeptide may participate in ion pair formation. We note further that upon extrapolation to 1 m-Na⁺, \( \log K_{\text{obs}} = -1.8 \) and \( \Delta G_{\text{ups}}^o(1 \text{ m-Na}^+) = +2.5 \) kcal. The non-electrostatic component of the free energy of interaction may be estimated by correcting for three lysine-like ion pairs (3\( \Delta G_{\text{ups}}^o(1 \text{ m-Na}^+) \)); it amounts to approximately +2 kcal. Presumably this is comprised of a favorable component due to interaction of the tryptophan side chain with the polynucleotide, and a larger unfavorable effect which may include the entropic contribution of polynucleotide structural distortion required for the formation of three ion pairs. (Brun et al., on the basis of Daune’s (1972) theory, concluded that 2 ion pairs were involved in the tripeptide–nucleic acid interaction.)

(c) Na⁺ effects on non-specific protein–nucleic acid interactions

Jensen & von Hippel (1976) and Jensen et al. (1976) have investigated the non-specific interactions of RNase with native and denatured DNA, and of T4 gene 32 protein with native DNA. In each case, they find that \( \log K_{\text{obs}} \) is a linear function of \( \log [\text{Na}^+] \). The slopes and intercepts of these lines, and the estimated overlap site sizes \( n \), are summarized in Table 3. To interpret these data in terms of ion pairing in the binding reaction, an assumption must be made regarding anion binding. The RNase binding data were obtained at pH 7.7 in the presence of \( 10^{-3} \) m-Na₃HPO₄ and added NaCl. Richards & Wyckoff (1971) state that RNase binds no chloride at pH 9.6 and less than one chloride at pH 6.6. At pH 7.7, the active site binds phosphate ions with a binding constant of only 25 m⁻¹. Consequently, we anticipate no effect on either the slope or intercept of a \( \log K_{\text{obs}}-\log [\text{Na}^+] \) plot from anion binding by RNase. No data are available for T4 gene 32 protein; we assume no anion binding here also.
LIGAND–NUCLEIC ACID INTERACTIONS

**Table 3**

Non-specific protein–nucleic acid interactions

<table>
<thead>
<tr>
<th>System</th>
<th>$\frac{\partial \log K_{\text{obs}}}{\partial \log [\text{Na}^+]_{1\text{ m}}}$</th>
<th>$\log K_0$ (1 m-Na+, 24°C)</th>
<th>n</th>
<th>m'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene 32 protein/native DNA</td>
<td>1.73†</td>
<td>1.03</td>
<td>10†</td>
<td>2.0</td>
</tr>
<tr>
<td>RNase/native DNA</td>
<td>3.90‡</td>
<td>-1.15</td>
<td>7-10‡</td>
<td>3.9</td>
</tr>
<tr>
<td>RNase/denatured DNA</td>
<td>4.95‡</td>
<td>-1.60</td>
<td>10-13‡</td>
<td>7.0</td>
</tr>
</tbody>
</table>

Symbols are defined in the text. The value of $\log K_0$ at 1 m-Na+ is estimated by extrapolation.

† Data of Jensen et al. (1976).
‡ Data of Jensen & von Hippel (1976).

With the assumption that there is no anion binding, we have calculated $m'$ (the number of ion pairs) for these non-specific binding reactions from equation (14). For gene 32 protein, it appears that two ion pairs are involved in complex formation with helical DNA, and there is a substantial favorable non-electrostatic component of the interaction. Upon extrapolation to 1 m-Na+, $\Delta G^0 = -2.2$ kcal. Upon correction for two lysine-like ion pairs, we estimate the non-electrostatic free energy of binding to be $-2.6$ kcal. This indicates that a substantial component of the binding free energy under physiological conditions results from non-specific interactions between the protein and base or sugar residues on the polynucleotide. This interaction may be relevant to an understanding of the mechanism of denaturation of poly[d(A-T)] by this protein.

In the case of RNase, approximately 4 and 7 ion pairs appear to be involved in the interaction of the enzyme with native and denatured DNA, respectively, at pH 7.7. The extrapolated values of $\log K_{\text{obs}}$ at 1 m-Na+ are unexceptional, and may be accounted for almost entirely by the intrinsic free energy of the lysine-like ion pairs. His12, His119, and Lys41 form the active site of RNase (Richards & Wyckoff, 1971). At pH 7.7, only lysine bears a positive charge, and it is therefore clear from the observed values of $m'$ that the polynucleotide chains are interacting with neighboring regions of the enzyme as well. From the data of Raju & Davidson (1969) on the pH dependence of binding of denatured DNA to RNase, $(\partial \log K_{\text{obs}}/\partial \text{pH}) = -0.9$ in the pH range 7 to 8. Therefore, one proton is taken up in the binding reaction, presumably by one of these histidine residues. There are at least three other lysines (Lys77, Lys37, Lys69) and two arginines (Arg10, Arg39) in the general vicinity of the catalytic site. It therefore appears possible to account for the observed values of $m'$. The rigidity of helical DNA and its larger molecular diameter presumably restrict its possibilities for ion pair formation relative to single-stranded DNA, which should mimic RNA in its binding pattern.

4. Discussion

The analysis presented above enables one to separate and quantify the electrostatic and non-electrostatic components of the free energy of ligand–nucleic acid complex formation. In effect, low molecular weight ions serve as thermodynamic probes of the interaction. The formation of ion pairs between ligand and nucleic acid releases counterions. The entropic effect of counterion release makes a major
contribution to the overall free energy of formation of complexes between polynucleotides and such model ligands as Mg$^{2+}$, oligolysines, and RNase. We have discussed these systems in detail in previous sections. Here we will consider the applicability of this analysis to protein–nucleic acid interactions in general.

Small proteins, with small binding sites and a physical site size $n$ which is comparable to the number of ionic interactions $m'$, should be adequately treated by the analysis in its present form. Larger proteins, where $m'$ may be large and $n$ may be substantially larger than $m'$, may require the consideration of anion binding by the ligand and of the condition of phosphate groups which are covered by the ligand but not involved in ion pairing. The extent of anion association with a polycation ligand can be predicted from polyelectrolyte theory, but anion binding by the DNA binding site of a protein such as lac repressor or Escherichia coli RNA polymerase will have to be determined experimentally as a function of the nature and concentration of the anion added.

Even with these additional complexities, analysis of the effects of NaCl concentration on the observed association constants of specific and non-specific binding of repressor or polymerase to DNA should indicate the origin of specificity in these systems. We (de Haseth, Lohman, Record, Wensley & Burgess, manuscript in preparation) have found very large dependences of log$K_{obs}$ on log[Na$^+$] for the non-specific interactions of lac repressor and RNA polymerase with DNA, and have concluded that electrostatic interactions contribute substantially to the binding free energy in these systems. If we assume that the same site on the protein is involved in specific and non-specific binding, then anion effects should be similar in both modes of interaction. If ($\partial \log K_{obs}/\partial \log [\text{Na}^+]$) has the same constant value for both modes of binding, then it can be concluded that the number of ionic interactions is the same in both cases, and that specificity originates in the non-electrostatic component. Different values of ($\partial \log K_{obs}/\partial \log [\text{Na}^+]$) for specific and non-specific binding would indicate differences in the number of ionic interactions, and suggest that ion pairing is a component of specificity in that system. The differential binding data presented by Lin & Riggs (1972, 1975) for the interactions of lac repressor with lac operator and non-operator DNA, although obtained in the presence of Mg$^{2+}$ which complicates the numerical analysis substantially (de Haseth, Lohman & Record, unpublished results), indicate that non-specific binding of repressor is more sensitive to salt concentration than is specific binding. From their data one can tentatively conclude that non-specific binding of repressor involves more ion pairs than does the specific interaction of repressor and operator DNA. The quantitative treatment of these data and their interpretation will be published elsewhere.

5. Appendix

The development of equation (5) proceeds directly from the consideration of the free energy of a polyelectrolyte solution containing $n_p$ moles of polynucleotide phosphates and $n_{\text{NaCl}}$ moles of added NaCl. The free energy $G$ of this solution is (Manning, 1972)

$$G = n_p \mu_p + (n_{\text{NaCl}} + n_p \xi_p^{-1}) \mu_{\text{Na}^+} + n_{\text{NaCl}} \mu_{\text{Cl}^-} + n_{\text{H}_2\text{O}} \mu_{\text{H}_2\text{O}}$$

(19)

The quantity $n_p \xi_p^{-1}$, where $\xi_p$ is the reduced charge density parameter defined in the
text, gives the number of phosphates on the helix that are not neutralized by Na⁺ condensation. Hence \( n_p \xi_p^{-1} \) is also the number of moles of Na⁺ released into the solution per mole of NaDNA added, and \( 1 - \xi_p^{-1} \equiv \psi_c \) is the fraction of a Na ion remaining condensed per phosphate.

To obtain \( \psi_a \) and \( \psi = \psi_c + \psi_a \), we consider the form of \( \mu_p \). Among the contributions to \( \mu_p \) enumerated by Manning (1972), only the Debye–Hückel screening term depends on [Na⁺]:

\[
\mu_p = \mu_p^\kappa - \xi_p^{-1} RT \ln \kappa,
\]

where \( \mu_p^\kappa \) includes all other contributions to \( \mu_p \), and \( \kappa \) is the Debye–Hückel screening parameter. Since \( \kappa^2 \) is proportional to \( n_{Na^+} \),

\[
\mu_p = \mu_p^{\kappa} - (2\xi_p)^{-1} RT \ln n_{Na^+},
\]

where \( \mu_p^{\kappa} \) collects the terms in \( \kappa \) independent of [Na⁺]. Moreover

\[
\mu_{Na^+} = \mu_{Na^+}^a + RT \ln n_{Na^+},
\]

where \( \mu_{Na^+}^a \) is independent of Na⁺ concentration. Therefore, from equations (19) to (22)

\[
G = n_p \mu_p^{\kappa} + (n_{NaCl} + n_p \xi_p^{-1}) \mu_{Na^+}^a \\
+ RT (n_{NaCl} + n_p (\xi_p^{-1} - (2\xi_p)^{-1})) \ln n_{Na^+} \\
\quad + n_{Na} \mu_{NaCl}^- + n_{H_2O} \mu_{H_2O}.
\]

Focusing on the coefficient of the \( \ln n_{Na^+} \) term in equation (23), we see that, whereas the fraction \( \xi_p^{-1} \) of a Na ion is directly released per phosphate of NaDNA added to the solution, the fraction \( (2\xi_p)^{-1} \) of a Na ion is reclaimed (i.e. does not contribute to \( \mu_{Na^+} \)) per DNA phosphate by the screening effect. Consequently, only \( \xi_p^{-1} - (2\xi_p)^{-1} = (2\xi_p)^{-1} \) Na ions are thermodynamically released to the solution per DNA phosphate or \( 1 - (2\xi_p)^{-1} \) Na ions are thermodynamically bound per phosphate. That is

\[
\psi_a = (2\xi_p)^{-1}
\]

and

\[
\psi = \psi_c + \psi_a = 1 - (2\xi_p)^{-1}.
\]

For double-stranded DNA, \( \xi^{-1} = 0.24 \) and we obtain a value for \( \psi = 0.88 \) for the case where only Na counterions are present.

We thank Drs G. S. Manning and P. H. von Hippel for their comments on this manuscript. Support from the National Science Foundation (GB43249) and from the National Institutes of Health Biomedical Sciences Support program administered by the University of Wisconsin is gratefully acknowledged.

**REFERENCES**


