Electrostatic Contributions to the Binding Free Energy of the λcI Repressor to DNA

Vinod K. Misra, Jonathan L. Hecht, An-Suei Yang, and Barry Honig
Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032 USA

ABSTRACT A model based on the nonlinear Poisson-Boltzmann (NLPB) equation is used to study the electrostatic contribution to the binding free energy of the λcI repressor to its operator DNA. In particular, we use the Poisson-Boltzmann model to calculate the pK_a shift of individual ionizable amino acids upon binding. We find that three residues on each monomer, Glu^{34}, Glu^{83}, and the amino terminus, have significant changes in their pK_a and titrate between pH 4 and 9. This information is then used to calculate the pH dependence of the binding free energy. We find that the calculated pH dependence of binding accurately reproduces the available experimental data over a range of physiological pH values. The NLPB equation is then used to develop an overall picture of the electrostatics of the λcI repressor–operator interaction. We find that long-range Coulombic forces associated with the highly charged nucleic acid provide a strong driving force for the interaction of the protein with the DNA. These favorable electrostatic interactions are opposed, however, by unfavorable changes in the solvation of both the protein and the DNA upon binding. Specifically, the formation of a protein-DNA complex removes both charged and polar groups at the binding interface from solvent while it displaces salt from around the nucleic acid. As a result, the electrostatic contribution to the λcI repressor–operator interaction opposes binding by ~73 kcal/mol at physiological salt concentrations and neutral pH. A variety of entropic terms also oppose binding. The major force driving the binding process appears to be release of interfacial water from the protein and DNA surfaces upon complexation and, possibly, enhanced packing interactions between the protein and DNA in the interface. When the various nonelectrostatic terms are described with simple models that have been applied previously to other binding processes, a general picture of protein/DNA association emerges in which binding is driven by the nonpolar interactions, whereas specificity results from electrostatic interactions that weaken binding but are necessary components of any protein/DNA complex.

INTRODUCTION

Understanding the forces that drive protein-DNA interactions provides insight into the requirements for stable and specific nucleic acid recognition. The main contributions to the free energy of binding can be divided into electrostatic and nonelectrostatic terms, where the nonelectrostatic contribution includes hydrophobic interactions; van der Waals interactions; and translational, rotational, and configurational entropies (Honig et al., 1993). The relative magnitude of each of these contributions to protein-DNA binding is not well defined. However, given the highly charged nature of nucleic acids and the fact that the free energy of protein-DNA binding is generally dependent on both the salt concentration (Misra et al., 1994a,b; Record et al., 1991) and pH (Overman and Lohman, 1994; Record et al., 1991; Senear and Batey, 1991; Bajaj et al., 1990), it is clear that electrostatic interactions play a major role in protein-DNA binding.

The free energy of protein-DNA interactions generally decreases with increasing pH (Senear and Ackers, 1990; Jen-Jacobsen et al., 1983; Record et al., 1978). Because proton-linked effects are a sensitive probe of electrostatic effects on protein-DNA binding (Misra and Honig, 1995), the pH dependence of binding can be used to test the accuracy of theoretical models used to calculate electrostatic free energies. The λcI repressor is well suited for such a study, because its affinity for its operators is highly dependent on pH (Senear and Ackers, 1990) and because pH effects contribute to the discrimination of DNA binding sites by the repressor (Senear and Ackers, 1990).

Our calculations based on the nonlinear Poisson-Boltzmann (NLPB) equation have provided an accurate description of both the magnitude and the salt dependence of electrostatic interactions in nucleic acid systems (Misra and Honig, 1996; Misra and Honig, 1995; Misra et al., 1994a,b; Zacharias et al., 1992). More recently, the NLPB equation has been shown to accurately reproduce the observed pK_a shift of an intercalating drug bound to DNA (Misra and Honig, 1995) and has provided reliable estimates for the binding free energies of anthracycline antibiotics to DNA (Baginiski et al., 1997). In this paper, we will use the NLPB equation to study pH effects on the binding free energy of the λcI repressor–operator system. Our results will be used as the basis for a general discussion of electrostatic contributions to protein-DNA interactions.

Received for publication 6 March 1998 and in final form 5 August 1998.

Address reprint requests to Dr. Barry Honig, Department of Biochemistry and Molecular Biophysics, 630 W. 168 St., New York, NY 10032. Tel.: 212-305-7970; Fax: 212-305-6926; E-mail: honig@trantor.bioc.columbia.edu.

0006-3495/98/11/2262/12 $2.00

METHODS

Theory

Calculation of the total electrostatic binding free energy

The total electrostatic free energy of a molecule in a single ionization state, ΔG_{el}, in a monovalent salt solution can be

calculated using the NLPB equation:

$$\nabla \cdot [\epsilon(\mathbf{r})\nabla \cdot \phi(\mathbf{r})] - (8\pi e^2 I/kT)\sinh[\phi(\mathbf{r})] + 4\pi e\rho^{f}(\mathbf{r})/kT = 0 \quad (1)$$

where ϕ is the dimensionless electrostatic potential in units of kT/e, in which k is Boltzmann's constant, T is the absolute temperature, e is the proton charge, ϵ is the dielectric constant, ρ^{f} is the fixed charge density, and I is the ionic strength of the bulk solution. The quantities ϕ, ϵ, and ρ are all functions of the position vector $\mathbf{r}$ in the reference frame centered on a fixed macromolecule. For any system modeled with the NLPB equation, ΔG_{el} is given by a volume integral over all space (Sharp and Honig, 1990a):

$$\Delta G_{el} = \int\{\rho^{f}\phi^{f}/2 + \rho^{f}\phi^{m} + \rho^{m}\phi^{m}/2 - (\rho^{m}\phi + kTc^{b}[2\cosh(\phi) - 2])\}d\nu \quad (2)$$

where the potential, ϕ, and charge density, ρ, have been split up into contributions from the fixed (f) and mobile (m) charges, and c^{b} is the bulk salt concentration.

The electrostatic free energy of a macromolecule, described by Eq. 2, can be partitioned into salt-independent and salt-dependent terms. The salt-independent contribution to ΔG_{el} is given by (Gilson and Honig, 1988; Gilson et al., 1985)

$$\Delta G_{ns} = \int(\rho^{f}\phi^{f}/2)d\nu \quad (3)$$

and the salt-dependent contribution to ΔG_{el} is given by (Misra et al., 1994b)

$$\Delta G_{s} = \int\{\rho^{f}\phi^{m} + \rho^{m}\phi^{m}/2 - (\rho^{m}\phi + kTc^{b}[2\cosh(\phi) - 2])\}d\nu \quad (4)$$

The electrostatic contribution to the binding free energy can be expressed as the difference in the free energy between the products (the protein-DNA complex, P·DNA) and the reactants (the protein, P, and DNA; Misra et al., 1994b):

$$\Delta\Delta G_{el} = \Delta\Delta G_{ns} + \Delta\Delta G_{s} = \Delta G_{el}^{P\cdot DNA} - \Delta G_{el}^{P} - \Delta G_{el}^{DNA} \quad (5)$$

where $\Delta\Delta G_{ns}$ is the salt-independent electrostatic free energy contribution and $\Delta\Delta G_{s}$ is the salt-dependent contribution. A physically intuitive description of the salt-independent contributions to $\Delta\Delta G_{el}$ has been discussed in prior publications (Misra and Honig, 1995; Gilson and Honig, 1988). The thermodynamic pathway underlying this description is briefly reviewed here. In the initial state, the fully charged molecules are infinitely separated from each other in a solvent of dielectric ϵ_{s}. In the first step, each molecule is partially desolvated by a low dielectric cavity (ϵ_{m}) corresponding to the region that the other molecule will come to occupy in the complex. The electrostatic free energy of this step is equal to the free energy of desolvating each molecule upon binding, $\Delta\Delta G_{d,P} + \Delta\Delta G_{d,DNA}$. In the following step, the charged molecules are transferred into the low dielectric cavity to form the final charged complex. The free energy of this step is given by the intermolecular solvent screened Coulombic free energy, $\Delta\Delta G_{sc}$. The total salt-independent contribution is then given by the sum of the solvent screened Coulombic free energy and the desolvation free energies. The thermodynamic process used to calculate the salt-dependent contributions to $\Delta\Delta G_{el}$ has also been described in detail in a prior publication (Misra et al., 1994b). $\Delta\Delta G_{s}$ can be interpreted simply as the change in the free energy of solvation of each molecule by salt in an aqueous environment (Misra et al., 1994b).

Calculation of the pH dependence of protein-DNA binding

The contribution of individual ionizable groups to $\Delta\Delta G_{el}$ can be calculated from the change in the pK_{a} of all ionizable groups in the free and bound repressors by using a method that has been fully described in previous publications (Yang et al., 1993; Yang and Honig, 1993). We will summarize here the method used to calculate the pH dependence of binding from changes in the pK_{a} of individual ionizable groups on the protein. In reference to the earlier work, which studied the pH dependence of protein stability (Yang et al., 1993; Yang and Honig, 1993), the bound and free states of the protein-DNA complex are analogous to the folded and unfolded states of the protein. Titration of ionizable groups on DNA are unlikely to be involved in the observed pH dependence of binding at physiological values of pH (Senear and Ackers, 1990). Furthermore, calculation of the titration behavior of nucleic acid groups is beyond the scope of the present study.

For a protein with N ionizable residues, a given ionization state, n, where $n = 1\text{–}2^{N}$, can be defined in terms of the vector $\delta_{n}(i)$, $i = 1\text{-}N$. $\delta_{n}(i)$ is zero when the group i is neutral and 1 when it is charged. A reference state of zero free energy is defined as corresponding to the state where all ionizable groups are neutral (i.e., $\delta_{n}(i) = 0$ for all i). ΔG^{n}(bound) is the pH-dependent free energy of the nth state in the protein-DNA complex and is given by (Yang et al., 1993)

$$\Delta G^{n}(\text{bound}) = \sum_{i=1}^{N}\left\{(\delta_{n}(i)\gamma(i)[2.3kT(\text{pH} - \text{pK}_{a_i}^{int})] + \sum_{1\le j<i}\delta_{n}(i)\delta_{n}(j)\Delta G^{ij}\right\} \quad (6)$$

where ΔG^{ij} is the electrostatic interaction energy between groups i and j in their charged state; the term $\gamma(i) = -1$ or $+1$ for an acidic and a basic group, respectively; and pK_{ai}^{int}, the "intrinsic pK_{a}," is defined as the pK_{a} of group i in the hypothetical state when all other ionizable groups are neutral. The term pK_{ai}^{int} can be expressed as

$$\text{pK}_{a_i}^{int} = \text{pK}_{a_i}^{o} - \gamma(i)\Delta G_{i}^{env}/2.3kT \quad (7)$$

in which pK_{ai}^{o} is the pK_{a} of the isolated amino acid in solution and ΔG_{i}^{env} is the change in electrostatic free energy

produced by charging a residue in the all-neutral protein environment (a low dielectric cavity in water) relative to the same charging process of the isolated residue in water. ΔG^{env} includes contributions not only from desolvation resulting from the group's location in the protein, but also from its interactions with polar groups on the protein as well as polar and charged groups on the DNA. The term ΔG^{env} is calculated using the PB equation. An expression similar to Eq. 6 can be written for the unbound (or "free") protein, ΔG^{n}(free).

The total electrostatic free energy of the ionizable groups of the bound protein is obtained from the statistical mechanical expression

$$\Delta G^{ion}(\text{bound}) = -kT \ln(Z^{B}) \tag{8}$$

where Z^{B} is the Boltzmann weighted sum of 2^{N} ionization states of the bound protein:

$$Z^{B} = \sum_{n=1}^{2^{N}} \exp[-\Delta G^{n}(\text{bound})/kT] \tag{9}$$

Similar expressions can be written for the total electrostatic free energy of the ionizable groups of the free protein. The terms ΔG^{ion}(bound) and ΔG^{ion}(free) are calculated with respect to the hypothetical reference state in which all ionizable residues are neutral. Both ΔG^{ion}(bound) and ΔG^{ion}(free) are equal to zero when all ionizable groups on the protein are neutral. The total electrostatic free energy of binding can be written as

$$\Delta\Delta G_{el} = \Delta\Delta G^{neutral} + \Delta\Delta G^{ion} \tag{10}$$

where $\Delta\Delta G^{neutral}$ is the binding free energy in the hypothetical state in which all ionizable protein residues are neutral. It includes the free energy of desolvating both charged (phosphates) and polar (bases) groups on the DNA as well as polar groups on the protein, and the interaction free energy of polar groups on the protein with charges on the DNA, including hydrogen bonds (Yang and Honig, 1993). The term $\Delta\Delta G^{ion}$ is simply equal to the difference ΔG^{ion}(bound) $-$ ΔG^{ion}(free). It should be clear that $\Delta\Delta G^{ion}$ is pH dependent, whereas $\Delta\Delta G^{neutral}$ is pH independent.

Because the summation over 2^{N} states in Eq. 9 is computationally intractable, the number of states used in the calculation is reduced by using an average charge approach (Yang and Honig, 1993). In this approach, the effect of pH on binding is described with a general expression given by Tanford (1970), who treated the problem of pH dependence in terms of multiple equilibria involving acids and bases. The electrostatic contribution to the binding free energy at some pH relative to a reference pH_1, $\Delta\Delta G_{el}(pH, pH_1)$, is written as

$$\Delta\Delta G_{el}(pH, pH_1) = [\Delta G^{ion}(\text{bound}) - \Delta G^{ion}(\text{free})](pH, pH_1)$$

$$= 2.3kT \int_{pH_1}^{pH} \Delta Q(pH)\,dpH \tag{11}$$

where ΔQ(pH) is the difference in average charge between the bound and free proteins at a given pH, as determined from their calculated titration curves (a function of ΔG^{n}(bound) and ΔG^{n}(free)). A reference pH_1 is chosen at a pH where a single ionization state of the protein predominates and there is no change in charged state of the protein, ΔQ, upon binding. At this reference pH, $\Delta\Delta G_{el}$ has little pH dependence and can be calculated considering only a single ionization state. For the λcI repressor, it is convenient to choose a reference pH of 7 in which all ionizable groups are charged (see justification below) and a well-defined structure exists. The electrostatic free energy of this reference state can be formally calculated using Eq. 1. The integral Eq. 11 then offers an efficient means of calculating the electrostatic binding free energy as a function of pH.

Calculating the overall binding free energy

Expanding on an approach that has been used previously to describe ligand binding to proteins and DNA (Baginski et al., 1997; Froloff et al., 1997), the pH-dependent binding of a protein, P, to DNA can be described as a two-step process:

$$\text{DNA} + \text{P} \rightarrow \text{DNA*} + \text{P*} \tag{12}$$

$$\text{DNA*} + \text{P*} \rightarrow \text{DNA*P*} \tag{13}$$

In the first step (Eq. 12), the unbound B-DNA and protein adopt the conformation that they will assume in the complex, DNA* and P*, respectively. The free energy of this step is designated $\Delta\Delta G_{strain}$ (Froloff et al., 1997). It is assumed that all of the energetic consequences of the conformational changes associated with binding, including both electrostatic and nonelectrostatic free energies, are included in this term. Although it will not be directly calculated, we will estimate and discuss the magnitude of $\Delta\Delta G_{strain}$ below.

The binding of these "conformationally strained" species, DNA* and P*, occurs in the second step (Eq. 13). The association of DNA* with P* is assumed to be a rigid body association between two well-defined macromolecules. For solution conditions in which a well-defined ionization state of the protein can be identified, the electrostatic free energy of this step can be explicitly calculated using the full NLPB equation described by Eqs. 2–5. As mentioned above, for the λcI repressor, it is convenient to choose a reference pH of 7 in which all ionizable groups are charged and a well-defined structure exists. The electrostatic binding free energy of this reference state can be obtained from Eq. 5 and is designated as $\Delta\Delta G_{el}(pH_1)$.

The pH dependence of binding is assumed to arise from changes in the ionization state of the protein upon binding (Eq. 13). As such, the electrostatic contribution to the binding free energy at some arbitrary pH, $\Delta\Delta G_{el}$, can be calculated using Eq. 11 so that

$$\Delta\Delta G_{el} = \Delta\Delta G_{el}(pH_1) + \Delta\Delta G_{ion}(pH, pH_1) \tag{14}$$

where $\Delta\Delta G_{ion}(pH, pH_1)$ is the difference in $\Delta\Delta G_{ion}$ at an arbitrary pH relative to the reference pH_1 (i.e., ΔG_{ion}(bound) − ΔG_{ion}(free) (pH, pH_1)). In this work, the first term on the right-hand side of Eq. 14 is calculated using the full NLPB equation. The second term is approximated using the statistical mechanical methods outlined here and before (Yang et al., 1993; Yang and Honig, 1993) and only makes use of the linearized PB equation. This latter approximation is required to make the problem computationally tractable, as discussed further below.

Molecular model

Parameter sets

The details of the model used to describe the ligand and the DNA in the finite-difference NLPB method have been given in several publications (Honig et al., 1993; Sharp and Honig, 1990b; Jayaram et al., 1989; Gilson et al., 1988). The bound and free molecules were described by the three-dimensional structure of the λcI repressor–O_L1 operator complexes as described below. Parameter sets characterizing the sizes and partial charges of atoms in the protein-DNA complex were obtained from the AMBER (Weiner et al., 1986) and OPLS (Pranata et al., 1991; Jorgensen and Tirado-Rives, 1988) force fields for the protein and DNA. The difference in the calculated electrostatic binding free energies for these two parameter sets is less than 5%. Results are given for the AMBER charge set.

Each molecule is treated as a low-dielectric medium (ϵ_m) within the volume enclosed by the molecular surface of the macromolecule (probe radius = 1.4 Å). The molecular dielectric constant, ϵ_m, is an important parameter for the calculation of the electrostatic free energy. As has been discussed many times in the past (e.g., see Gilson and Honig, 1986), a dielectric constant of 2 accounts well for the effect of electronic polarizability, and a value of 4 can be used to account for the effect of small dipolar fluctuations that may accompany structural transitions in proteins. The dielectric constant of nucleic acids is also expected to be small (Yang et al., 1995) because base pairing, base stacking, and steric interactions will restrict dipolar fluctuations within the double helix.

Because Eq. 13 describes the interaction of two rigid molecules, a molecular dielectric constant (ϵ_m) of 2 was used to calculate $\Delta\Delta G_{el}(pH_1)$. That is, because all conformational changes, even small ones associated with dipolar relaxation, are included in Eq. 12, only electronic polarizability affects the calculations based on Eq. 13 (Froloff et al., 1997). In contrast, a molecular dielectric constant of 4 was used to calculate the term $\Delta\Delta G_{ion}(pH, pH_1)$ in Eq. 14. For these calculations, a larger value of ϵ_m is needed to describe the dipolar fluctuations associated with multiple ionization states of the protein (Yang et al., 1993). This issue will be discussed further below.

The surrounding solvent was treated as a continuum of dielectric constant (ϵ_s) 80 with a 1:1 electrolyte distributed according to a Boltzmann weighted average of the mean electrostatic potential. A 2.0-Å ion exclusion radius (corresponding roughly to the hydrated radius of a sodium ion) was added to the surface of the macromolecules to account for ion size. For the calculations presented here, a univalent salt concentration of 0.2 M was used.

Structural models used to calculate $\Delta\Delta G_{el}(pH_1)$

For calculations of $\Delta\Delta G_{el}(pH_1)$, atomic coordinates of the NH_2-terminal domain of the λcI repressor (the DNA binding domain) bound to a 20-bp oligonucleotide containing the O_L1 operator were based on the x-ray crystallographic structure of the complex refined to 1.8 Å at a pH of 7.0 (Beamer and Pabo, 1992). The first five NH_2-terminal residues in the nonconsensus half-site of the λcI repressor were missing in the crystallographic structure of the complex. For our purposes, the five amino terminal residues on the nonconsensus half-site were first built with twofold symmetry to the consensus half-site. These five residues were then energy minimized with DISCOVER 3.1 (MSI), with all other residues fixed to optimize the conformation of the nonconsensus NH_2-terminus. In our final structure, the NH_2-terminus on the nonconsensus half-site is slightly farther away from the DNA than is the consensus NH_2-terminus. This is consistent with the crystallographic data, which show that the repressor dimer compensates for differences between the two arms by shifting the residues before helix 1 on the nonconsensus monomer slightly away from the DNA (Beamer and Pabo, 1992).

The following protein groups were treated as titratable: Asp, Glu, Lys, Arg, the NH_2-terminus, and the COOH-terminus (His and Tyr are not present in the λcI repressor). Before assigning partial charges to each atom, protons were added to each molecule according to the crystallographic heavy atom designations and minimized with all heavy atoms fixed to reduce van der Waals overlaps and maximize hydrogen bonds. All ionizable groups on the proteins were fully charged in the free and bound state, consistent with experimental pH titration data at neutral pH (Senear and Ackers, 1990) as well as pH-dependent free energy calculations presented here.

Structural models used to calculate $\Delta\Delta G_{ion}(pH, pH_1)$

The structural model used to calculate the pH dependence of protein equilibria has been described extensively in several prior publications (Yang and Honig, 1993; Yang et al., 1993). A complete treatment of pH-dependent phenomena would require knowledge of the protein conformation of each ionization state at each pH. The simplest approximation is to assume that each of the 2^N possible states at each pH has the same conformation (i.e., a single "rigid" structure whose conformation is independent of pH). However, this approach does not account for structural changes that

may accompany changes in the ionization state of each residue. A less restrictive assumption, which has been found to enhance agreement between theoretically calculated and experimentally determined pK_as, is to model two conformations of the protein, one in which all of the titrable groups are completely ionized and one in which they are all neutralized (Yang and Honig, 1993; Yang et al., 1993). As previously described, these two states can then be used to individually calculate the relative electrostatic free energy of charging each residue in the protein environment—see Eqs. 6 and 7 (Yang and Honig, 1993; Yang et al., 1993).

For these calculations, we start with the complete structure of the λcI repressor–OL1 operator complex based on the crystallographic coordinates as described above. To model the appropriate charge states, molecular dynamics calculations were carried out on the fully charged and fully neutral repressor, both alone and complexed with its operator. With the use of DISCOVER (Biosym Technologies), a 5-Å layer of water (~600 water molecules) was added to each repressor/repressor-operator system. The systems were then energy minimized and equilibrated for 5 ps at 300 K, in 1-fs iterations. This was followed by a 5-ps period in which "snapshots" of the conformation were taken at 0.5-ps intervals. An average conformation for both the all-charged and all-neutral states of the protein was obtained from an arithmetic average of the 10 "snapshots" for each state. All atoms on the DNA were fixed at the crystal structure coordinates during the simulations. Charged and neutral structures of the repressor were constructed in this way.

In molecular dynamics simulations of nucleic acids, the phosphate charges are typically reduced to between 0.2 and 0.32 (Nilsson and Karplus, 1984; Brooks et al., 1983; Tidor et al., 1983) to account for the electrostatic effects of a large accumulation of counterions around DNA that is observed in experiment (Anderson et al., 1978) and predicted by theory (Anderson and Record, 1990; Jayaram et al., 1989, 1990; Manning, 1969). In NLPB calculations for the repressor-complex system, we find that the electrostatic potential of the DNA measured at the ionizable amino acids of the repressor is reduced by 20–80% in 0.2 M salt relative to no salt. Thus no single phosphate scaling factor can fully account for the effect of salt on electrostatic interactions. In the simulations of the complex used in this work, the partial charges assigned to the DNA phosphate groups are scaled to 0.6. The larger value for the phosphate charge used here reflects the exclusion of ions from the high-potential region near the DNA by the repressor. To study the effect of scaling on the simulations, we reduced the phosphate charges to 0.23 and repeated the simulations. The pK_a shifts and the pH dependence of binding calculated with the resulting averaged structures agree to within 10% of the values calculated using a phosphate charge of 0.6. Thus the pK_a calculations are not very sensitive to the charge assigned to the phosphates in the dynamics simulations.

The major effect of the simulations was to produce small movements throughout the protein, which exposed charged side chains at the surface to solvent while maintaining the interactions of side chains at the interface with the DNA. The structure of the complex obtained from the simulations had an RMS deviation of ~1.8 Å compared to the crystal structure of the complex. The structure of the free protein obtained from the simulations had an RMS deviation of ~1.5 Å compared to the protein in the crystal structure of the complex. These differences arise primarily from the increased solvent accessibility of the ionizable side chains on the protein surface during dynamics. The effect of an increased degree of solvation is, in general, to reduce the magnitude of the pK_a shifts relative to the crystal structure.

Numerical calculations

Details of the finite-difference procedure used to calculate electrostatic potentials with the PB equation have been reported previously (Nicholls and Honig, 1991; Jayaram et al., 1989; Gilson et al., 1988). To calculate the electrostatic potentials, the molecular system is first mapped onto a three-dimensional lattice. Parameters are assigned to each lattice point according to the molecular model described above. The finite-difference equations are solved by optimized successive overrelaxation to obtain the potential at all grid points (Nicholls and Honig, 1991).

The term $\Delta\Delta G_{el}(pH_1)$ is calculated using the finite-difference nonlinear PB equation in which the charges and molecular surface were mapped onto a 129^3 lattice. A simple three-step focusing procedure is used to improve the accuracy of the potentials (Gilson et al., 1988). In the initial calculation, the largest dimension of the macromolecule fills 24% of the grid, and the potentials at the lattice points on the boundary of the grid are approximated analytically using the Debye-Hückel equation (Klapper et al., 1986). The final potentials are calculated in three steps in which the grid is made four times finer such that the largest dimension of the macromolecule fills 96% of the grid with the boundary conditions interpolated from the previous step. The final resolution for the ligand-DNA complexes was 1.8 grids/Å. Each electrostatic free energy term is calculated from the electrostatic potentials at each lattice point using the appropriate numerical integrals over the grid as previously described (Misra et al., 1994b; Jayaram et al., 1989; Gilson and Honig, 1988). At these resolutions, the $\Delta\Delta G_{el}$ was found to vary by less than 3% with the position of the molecules on the grid.

As previously described (Yang et al., 1993), the term $\Delta\Delta G_{ion}(pH, pH_1)$ is calculated from ΔG^{ij} and ΔG^{env} in Eqs. 6 and 7, respectively. The latter terms are directly calculated from the electrostatic potentials at each titrating site (Yang et al., 1993). For these calculations, the charges and molecular surface describing the molecules are mapped onto a 65^3 lattice, and the finite difference form of the linearized PB equation is solved. A four-step focusing procedure is used to calculate the electrostatic potentials in which the final resolution of the grid is at least 2.8 grids/Å for each calculation. At this scale, potentials measured at the center of the

grid vary by only 1% with changes in the molecules' positions upon the grid. In principle, the entire calculation of the pH dependence of binding could have been performed using the NLPB equation, but in this case ΔG^{ij} would depend on the overall ionization state of the protein. Taking this interdependence into account requires that the free energy of each ionization state, ΔG^{n}, be calculated separately. Use of a single pairwise interaction energy for every protein charge state would no longer be valid, and the computation becomes intractable. It should be emphasized that the major nonlinear effects result from the high charge density on the DNA (included in $\Delta\Delta G_{el}(pH_1)$), so that ignoring nonlinear effects due to the charge state of the protein is not a severe approximation.

The site potentials calculated above can be parsed to calculate the solvent-screened Coulombic interaction of each ionizable residue, *i*, with DNA, $\Delta\Delta G'_{sc}$, and from changes in the residues' interactions with solvent, $\Delta\Delta G'_{d}$, and salt, $\Delta\Delta G'_{s}$, upon binding. These methods have been extensively described in prior publications (Yang et al., 1993; Gilson and Honig, 1988). This procedure is analogous to that described above for the total electrostatic binding free energy.

RESULTS

The pH dependence of binding

As can be seen in Fig. 1, the calculated electrostatic contribution to the binding free, $\Delta\Delta G_{el}$, becomes more unfavorable with increasing pH. In Fig. 1, the pH dependence of $\Delta\Delta G_{el}$ calculated for the λcI repressor–O_L1 operator interaction is directly compared to that of the experimental binding free energy of the O_R1 operator interaction between pH 5 and 8 at 0.2 M univalent salt concentration (Senear and Ackers, 1990). The interaction of the λcI repressor with the O_L1 operator has been shown to be energetically identical to that of the O_R1 operator interaction (Sarai and Takeda, 1989). We find that the calculated values of $\partial\Delta\Delta G_{el}/\partial$pH over the pH ranges 5–6, 6–7, and 7–8 are in good agreement with the experimental data (Table 1).

The pH dependence of the binding free energy arises from shifts in the pK_a of individual amino acids (Wyman and Gill, 1990). In general, we find that the pK_a values of all ionizable amino acids shift upward upon binding. This is expected because the ionized form of acidic amino acids is destabilized, and the charged form of basic residues is stabilized by protein-DNA interactions. Six residues in λcI repressor are calculated to have pK_a shifts that produce the pH dependence of binding in the experimental pH range (pH 5–8). The pK_a shifts of Glu^{34}, Glu^{83}, Glu^{134}, and Glu^{183} give rise to the pH dependence of binding between pH 5 and 6, whereas the pK_a shifts of the NH_2-terminus of each repressor monomer give rise to the pH dependence of binding between pH 7 and 9. Because the NH_2-terminus on the nonconsensus half-site is slightly farther from the DNA than is the consensus NH_2-terminus, the nonconsensus NH_2-terminus has a smaller pK_a shift (Table 2).

The free energy associated with these pK_a shifts can be divided into contributions from the solvent-screened Coulombic interaction of each ionizable residue, *i*, with DNA, $\Delta\Delta G'_{sc}$, and from changes in the residues' interactions with solvent, $\Delta\Delta'_{d}$, and salt, $\Delta\Delta G'_{s}$, upon binding (Table 2). In each case, the strong Coulombic interaction between the charged amino acid and the DNA, $\Delta\Delta G'_{sc}$, is the dominant contribution to the pK_a shift. However, the smaller solvent-related free energies for each amino acid, $\Delta\Delta G'_{d}$ and $\Delta\Delta G'_{s}$,

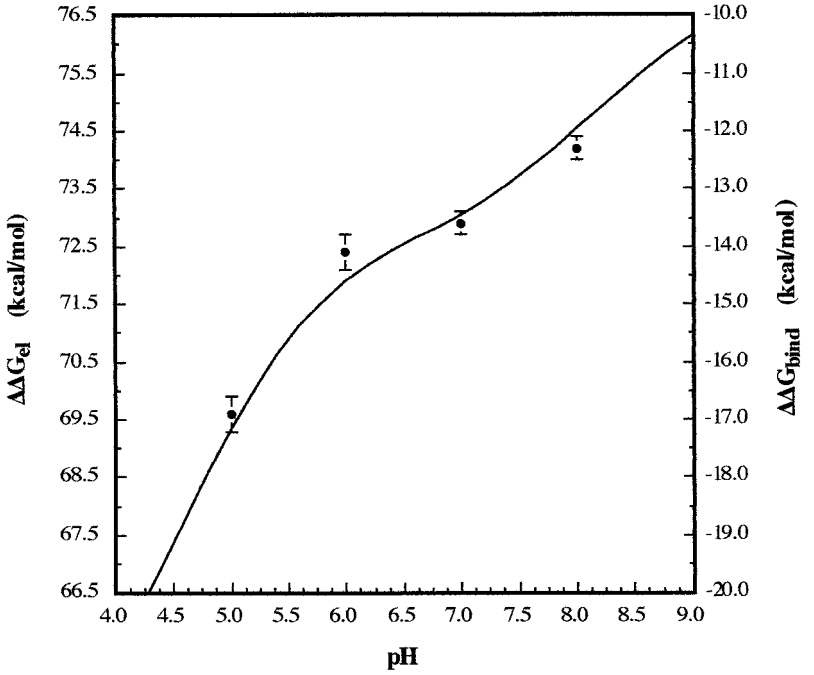

FIGURE 1 The pH dependence of the electrostatic binding free energy ($\Delta\Delta G_{el}$) for the λcI repressor calculated using the PB model (——). The experimental binding free energy ($\Delta\Delta G_{bind}$) for the λcI repressor–O_R1 operator interaction (Senear and Ackers, 1990) is also plotted (●). The error bars represent 65% confidence intervals.

TABLE 1 Comparison of the calculated and experimental pH dependence of λcI repressor binding to its operators at 0.2 M [M^+]

	pH 5–6	pH 6–7	pH 7–8
Calculated			
$\partial\Delta\Delta G_{el}/\partial$pH	2.5	1.1	1.4
Experimental*			
O_R1	2.9 ± 0.4	0.5 ± 0.4	1.4 ± 0.3
O_R2	1.9 ± 0.4	1.5 ± 0.3	1.2 ± 0.3
O_R3	1.8 ± 0.5	0.8 ± 0.3	1.1 ± 0.4

*Experimental values of $\partial\Delta\Delta G_{bind}/\partial$pH were calculated from the data of Senear and Ackers at $T = 25°C$ (Senear and Ackers, 1990).

attenuate the effect of $\Delta\Delta G'_{sc}$. These effects are quite similar to those seen in the binding of the simple intercalating ligand 3,8-diamino-6-phenylphenanthridine (DAPP) to DNA (Misra and Honig, 1995). It should be noted that titration of the DNA has not been included in our model. Although the titration of nucleic acid groups may play a role at the extremes of pH, they are unlikely to have a significant role in the proton-linked effects observed at physiological pH values (Senear and Ackers, 1990).

The total electrostatic free energy of binding

At pH 7.0, we find that the total electrostatic free energy, $\Delta\Delta G_{el}$, of the repressor-operator interaction opposes binding by 73 kcal/mol at 0.2 M univalent salt concentration (Fig. 1 and Table 3). The electrostatic binding free energy is described by the balance between the solvent screened Coulombic interaction, $\Delta\Delta G_{sc}$, of the protein with the DNA and the disruption of the solvent ($\Delta\Delta G_d + \Delta\Delta G_s$) around the protein and the DNA upon binding (Misra and Honig, 1995). In Table 3, we have partitioned $\Delta\Delta G_{el}$ into its components for two specific ionization states of the protein: a hypothetical state in which every ionizable group is neutral, and a state in which every ionizable residue is fully charged (see Eq. 10). The latter state corresponds closely to the ionization state of the protein at neutral pH based on the calculated pK_a values.

TABLE 2 The calculated pK_as of residues in the free and bound repressor with shifts between pH 4 and 10

Residues	pK_a (free)	pK_a (bound)	$\Delta\Delta G^i_{sc}$ (kcal/mol)*	$\Delta\Delta G^i_d + \Delta\Delta G^i_s$ (kcal/mol)*	$D(PO_4)^{\#}$ (Å)
NTE 1	7.1	9.3	−10.3	8.3	3.8
Glu^{34}	4.1	4.9	5.7	−4.7	7.2
Glu^{83}	5.1	5.3	6.7	−6.3	11.1
NTE 101	7.3	8.3	−8.3	8.6	5.0
Glu^{134}	4.1	5.6	6.8	−4.5	5.7
Glu^{183}	4.5	5.9	8.5	−7.2	9.3

*The contributions from the solvent-screened coulombic interaction of each ionizable residue, *i,* with DNA, $\Delta\Delta G^i_{sc}$, and from changes in the interaction of each residue with solvent, $\Delta\Delta G^i_d$, and salt, $\Delta\Delta G^i_s$, upon binding. Each term was calculated as the relative binding free energy of the protein with the listed residue charged and neutral.

#The distance to the nearest DNA phosphate oxygen. The distances were calculated between DNA phosphate oxygens and either the backbone nitrogen of the NH_2-terminii or the ε oxygen of the glutamic acids.

TABLE 3 The contributions to the calculated electrostatic binding free energy for the charged and neutral λcI repressor–operator complex*

	Free energy (kcal/mol)	
	Neutral	Charged
$\Delta\Delta G_{sc}$	−98.5	−190.5
$\Delta\Delta G_{d,DNA}$	157.6	157.6
$\Delta\Delta G_{d,prot}$	38.0	90.3
$\Delta\Delta G_s$	−2.5	15.6
$\Delta\Delta G_{el}$	94.7	73.0

*Free energy contributions calculated for a hypothetical "neutral" state in which every ionizable group on the protein is neutral and a state in which every ionizable residue is fully charged. The latter state corresponds to the average ionization state of the protein at neutral pH based on the calculated pK_a values. The calculations were done at 0.2 M univalent salt concentration with $\epsilon_m = 2$.

The solvent screened Coulombic interaction, $\Delta\Delta G_{sc}$, favors the binding of both the neutral and the charged forms of the protein (Table 3). For the neutral protein, this interaction arises from the attraction ($\Delta\Delta G_{sc} = -98.5$ kcal/mol) of the partially charged dipolar protein groups to the DNA and includes the electrostatic component of the hydrogen bonds between the protein and the DNA. For the charged protein, the additional strong attraction between the positively charged amino acids near the binding interface and the DNA phosphates substantially increases $\Delta\Delta G_{sc}$ to −190.5 kcal/mol. As the pH of solution is increased, the basic groups on the protein become increasingly neutralized, so that the absolute magnitude of $\Delta\Delta G_{sc}$ decreases.

Upon binding, the favorable Coulombic interactions described above are strongly opposed by desolvation of the protein and the DNA (Table 3). The desolvation free energy of the DNA ($\Delta\Delta G_{d,DNA}$) is ~158 kcal/mol for both the charged and neutral forms of the protein. This is largely due to the burial of the DNA phosphodiester backbone in the low-dielectric molecular interface. The relatively small desolvation free energy of the neutral protein ($\Delta\Delta G_{d,prot} = 38$ kcal/mol) reflects the burial of dipolar groups at the binding interface of the protein. The larger desolvation energy of the charged protein ($\Delta\Delta G_{d,prot} = 90.3$ kcal/mol) arises from the additional removal of charged amino acids from solvent upon binding.

As discussed previously (Misra et al., 1994a), $\Delta\Delta G_s$ reflects the redistribution of the ion atmosphere around the protein and the DNA that accompanies binding. $\Delta\Delta G_s$ for the neutral protein actually favors binding by −2.5 kcal/mol at 0.2 M univalent salt concentration (Table 3). This effect arises from the energetically favorable accumulation of cations in a solvent-filled pocket with a large negative electrostatic potential at the protein-DNA interface (Misra et al., 1994a). In contrast, the positively charged protein at neutral pH repels the cationic ion atmosphere around the DNA (Misra et al., 1994a) so that $\Delta\Delta G_s$ opposes binding by 15.6 kcal/mol at 0.2 M [M^+] (Table 3). Small differences in this

value (~4%) from previously published values (Misra et al., 1994a) arise from two sources: the first is small differences in the high-resolution crystal structure used here compared to the low-resolution crystal structure analyzed in those reports; the second is the use of a molecular dielectric constant of 2 in this work compared to a value of 4 in the prior work.

DISCUSSION

In Eq. 10, the total electrostatic free energy of binding is written as the sum of $\Delta\Delta G^{neutral}$ and $\Delta\Delta G^{ion}$. The agreement between our calculated values and the experimental values for the pH dependence of the λcI repressor with its operator shows that the NLPB model accurately reproduces $\Delta\Delta G^{ion}$ for multiple charge states of the protein-DNA system. This implies that the Coulombic and solvent-related electrostatic free energy terms for the ionizable amino acids ($\Delta\Delta G_{sc}$, $\Delta\Delta G_d$, and $\Delta\Delta G_s$) are correctly determined. Furthermore, the accuracy of the NLPB model for calculating $\Delta\Delta G^{ion}$ suggests that the parameters used here are, to a first approximation, reasonably accurate.

Given these results, we have used the NLPB model to analyze the overall electrostatic contribution to protein-DNA binding. We will first discuss the input parameters used in the NLPB model. We will then discuss the balance of electrostatic forces involved in the binding of the λcI repressor to DNA. Finally, we consider the factors that may be responsible for stable protein-DNA interactions.

The parameters used in the NLPB model

The molecular dielectric constant

As discussed above, a molecular dielectric constant of 2 was used to define the reference state at pH 7.0 ($\Delta\Delta G_{el}(pH_1)$), and a higher dielectric constant (ϵ_m of 4) was used to obtain the pH dependence of binding ($\Delta\Delta G_{ion}(pH, pH_1)$). This apparent inconsistent use of the dielectric constant is, in fact, perfectly reasonable once it is recognized that the binding step (Eq. 13) is assumed to take place between rigid molecules (ϵ_m of 2), whereas the calculation of pH-dependent effects must in some way account for conformational changes. Specifically, it is necessary to account for the fact that there are distinct conformations for each ionization state of the protein (Yang and Honig, 1993; Antosiewicz et al., 1996).

In most current applications, methods that incorporate charge-dependent protein conformational flexibility into the calculation of pH-dependent properties of proteins have used a molecular dielectric constant of 4 (Yang et al., 1993; Alexov and Gunner, 1997). In this paper we have applied the method of Yang et al. (1993), which uses different structures for the fully ionized and fully neutral forms of the protein, and have consequently used a molecular dielectric constant of 4 to evaluate the pH dependence of the binding constant. It should be pointed out that methods that do not explicitly account for conformational changes accompanying changes in ionization state have used much larger dielectric constants in pK_a calculations (Antosiewicz et al., 1994, 1996).

The atomic parameter set

The finite-difference NLPB model also depends on the choice of atomic charges and sizes. The general accuracy of the model requires the development of a consistent set of atomic charges and radii that accurately reproduces the solvation energies of model compounds (Sitkoff et al., 1994). Although pH- and salt-dependent effects do not depend strongly on the atomic charge and size sets used here, more sensitive analyses of the total binding free energies may depend on these parameters. For the analysis done here, the overall electrostatic free energies vary by less than 5% between the AMBER and OPLS parameter sets.

The electrostatic binding free energy

Coulombic attraction

The Coulombic attraction, $\Delta\Delta G_{sc}$, of the positively charged λcI repressor to its operator provides a strong driving force for binding at neutral pH (Table 3). The Coulombic interactions are enhanced by a clustering of basic residues near the DNA binding surface of the repressor protein. To a lesser extent, $\Delta\Delta G_{sc}$ includes local dipolar interactions, such as hydrogen bonds, which also stabilize the protein-DNA complex, as observed in the "neutral" protein-DNA interaction (Table 3). These findings generally agree with other computational studies on the same system (Jayaram et al., 1996). However, this prior work neglected the important role of desolvation discussed below.

Many DNA-binding proteins exploit Coulombic interactions with a predominantly basic DNA binding site that is complementary to the large negatively charged DNA backbone. This phenomenon has been characterized, for example, in the CAP protein (Schultz et al., 1991; Warwicker et al., 1987), the polymerase domains of *E. coli* DNA polymerase I and HIV-1 reverse transcriptase (Yadav et al., 1992, 1994; Warwicker et al., 1985), the processivity factor of *E. coli* DNA polymerase III (Kong et al., 1992), *E. coli* trp repressor (Guenot et al., 1994), and several Fab fragments of immunoglobulins that specifically recognize DNA (Mol et al., 1994a,b). This "electrostatic complementarity" may be particularly important for defining the specificity of interactions (McCoy et al., 1997; Hendsch and Tidor, 1994).

The strong Coulombic interaction between specific ionizable amino acids and the DNA backbone governs the pH dependence of binding by shifting the calculated pK_a's of the ionizable groups to more basic values (Table 2). In fact, the free energy of protein-DNA interactions is generally observed to decrease with increasing pH (Senear and Ackers, 1990; Jen-Jacobsen et al., 1983; Record et al., 1978). Our

calculations for the λcI repressor–operator system agree with this observation both qualitatively and quantitatively.

As expected, the strength of the Coulombic interaction with DNA depends, in part, on the distance of the residue from the DNA (Table 2). For example, the Glu^{134} side chain in the consensus half-site is 1.5 Å closer to the DNA phosphate backbone relative to its nonconsensus counterpart, Glu^{34}. As a result, $\Delta\Delta G^{i}_{sc}$ and the resulting pK_a shift of Glu^{134} are larger than for Glu^{34}. This shows how relatively small structural differences can lead to different interaction free energies via long-range forces. Similar sequence-specific structural differences may play a role in the differential affinity of the repressor for its various operator sites (Senear and Ackers, 1990).

The magnitude of $\Delta\Delta G^{i}_{sc}$ also depends on the three-dimensional shape of the molecular surface around the group *i*. For example, Glu^{83} and Glu^{183} interact almost as strongly with the DNA as Glu^{34} and Glu^{134}, even though they are almost twice as far away. The Glu^{83} and Glu^{183} side chains are found to project into a large solvent-filled pocket between the DNA and the repressor in which the negative electrostatic potentials are very high (Misra et al., 1994a). The surprisingly large Coulombic interaction of Glu^{83} and Glu^{183} is consistent with the experimental finding that mutating Glu^{83} to Lys increases the binding affinity of the repressor for O_L1 by 80-fold at pH 7.3 and 0.2 M ionic strength (Nelson and Sauer, 1985). Long-range electrostatic interactions have also been identified to play a role in the *Escherichia coli* methionine repressor-operator system. In this case, a positively charged corepressor, *S*-adenosylmethionine (SAM), binds to a site on the repressor ~10 Å from the DNA binding interface and increases the affinity of the protein for its DNA by almost 1000-fold (Phillips et al., 1989). These interactions have been found to be governed by the strong electrostatic attraction between SAM and DNA (Phillips and Phillips, 1994).

Desolvation

The formation of the λcI repressor–operator complex desolvates both charged and polar groups at the binding interface and displaces salt from around the DNA. The long-range Coulombic forces governing the interaction of each residue with DNA ($\Delta\Delta G^{i}_{sc}$) are largely offset by changes in the total solvation of each residue upon binding ($\Delta\Delta G^{i}_{d} + \Delta\Delta G^{i}_{s}$), thus attenuating their overall pK_a shift (Table 2). Furthermore, the magnitude of the sum of $\Delta\Delta G_d$ and $\Delta\Delta G_s$ is larger than the magnitude of $\Delta\Delta G_{sc}$ at physiological salt concentrations, so that $\Delta\Delta G_{el}$ actually opposes binding at neutral pH (Table 3).

The desolvation free energy of the DNA, $\Delta\Delta G_{d,DNA}$, is the largest unfavorable contribution to the electrostatic binding free energy (Table 3). The term $\Delta\Delta G_{d,DNA}$ primarily reflects the burial of the DNA sugar-phosphate backbone by the low dielectric protein at the binding interface. The large unfavorable desolvation of DNA phosphates is likely to be a general property of specific protein-DNA binding. The desolvation of charged groups on the protein, $\Delta\Delta G_{d,prot}$, also opposes binding, but to a smaller extent than $\Delta\Delta G_{d,DNA}$ (Table 3). In addition, as is evident from the relative desolvation penalties of the neutral and charged protein in Table 3, $\Delta\Delta G_{d,prot}$ clearly depends on the presence of charged amino acids at the protein-DNA interface.

As discussed above, $\Delta\Delta G_s$ reflects the redistribution of the ion atmosphere around the protein and the DNA upon binding. Although $\Delta\Delta G_s$ for the neutral protein slightly favors binding, salt strongly opposes the binding of the positively charged repressor at neutral pH (Misra et al., 1994a). This effect has previously been discussed in great detail (Misra et al., 1994a).

The stability of the λcI repressor–operator complex

The most surprising result of this work is that the overall electrostatic free energy opposes λcI repressor–operator binding under physiological conditions. This results primarily from the large unfavorable desolvation energy of burying charged groups at the binding interface. We find that the desolvation energy is not fully compensated by favorable electrostatic interactions between the protein and the DNA. This result is likely to be a general feature for the binding of low-dielectric macromolecules in a high-dielectric solvent. In fact, very similar findings have been presented for the binding of the simple intercalating ligand 3,8-diamino-6-phenanthridine (DAPP) to DNA (Misra and Honig, 1995), for the binding of anthracycline antibiotics to DNA (Baginski et al., 1997), and for the binding of peptides to MHC class I proteins (Froloff et al., 1997).

Protein-DNA interactions are somewhat unique in that a large number of charged groups, the DNA phosphates, are necessarily buried in the low-dielectric protein-DNA interface. As shown here, the resulting electrostatic desolvation penalty can be compensated in part by an excess of positively charged groups at the DNA-binding interface of the protein. Nucleic acids may also employ other strategies to minimize the free energy cost of burying phosphate groups. For example, the Trp repressor has been shown to retain a layer of water molecules in the binding interface with its operator (Ladbury et al., 1994; Shakked et al., 1994). Similarly, the λCI repressor–operator seems to have an extended network of solvent-mediated contacts between the protein and the DNA in the crystal (Beamer and Pabo, 1992).

Specifically bound cations and/or water molecules between the protein and DNA may influence the electrostatic free energy in several ways. First, as mentioned above, they can help "solvate" buried DNA phosphates at the interface. Second, they can similarly influence the pK_a of nearby buried ionizable groups (Gibas and Suramaniam, 1996; Yang et al., 1993). Finally, the binding process may be accompanied by the uptake or release of specifically bound solvent molecules, which can affect the overall free energy

of both the solvent and the macromolecule. As a macroscopic model used to calculate the overall electrostatic free energy of macromolecular systems, PB theory in its simplest form cannot be used to explicitly describe these microscopic processes (Honig et al., 1993). However, it should be realized that the PB model includes many enthalpic and entropic contributions to the electrostatic solvation free energy, such as hydrogen bonding, dielectric entropy, and solvent reorganization entropies (Sharp, 1995). Therefore, the inclusion of additional terms to describe these processes must be done with great care.

Although our current analysis has included only simple univalent ion effects, divalent ions play an important role in protein-DNA interactions. In the NLPB model, divalent ions in solution have been shown to destabilize the λcI repressor–operator complex by decreasing the overall electrostatic binding free energy (Chen and Honig, 1997). However, experimentally it has been observed that low concentrations of Ca^{2+} actually stabilize the λcI repressor–O_R operator complex by up to 4 kcal/mol (Sauer et al., 1990). Thus divalent ions may play an important role in compensating for the free energy cost of burying DNA phosphate groups. Modeling these divalent ion effects is an important area of future research.

Given our results, the major driving force for the formation of the λcI repressor–operator complex is almost certainly due to the release of water molecules from the surface of both the protein and DNA. Such effects have been shown to be an important driving force for the binding of several minor-groove binding ligands (Ding and Ellestad, 1991; Boger et al., 1990) and proteins (Spolar and Record, 1994; Lundback et al., 1993; Ha et al., 1989) to DNA. In the standard thermodynamic cycle used here to calculate electrostatic free energies (Friedman and Honig, 1995; Froloff et al., 1997), there is a contribution to binding in a hypthetical state where both interacting molecules are "discharged electrostatically." This contribution is hydrophobic in nature, although both polar and nonpolar surface areas are treated on the same footing, because at this state in the cycle the entire complex is nonpolar. It is worth pointing out that that the entropic effect of water release from nonpolar surfaces is distinct from entropic contributions to the dielectric response of water, which are accounted for directly in the PB calculations (Sharp, 1995) and are treated at different steps in the thermodynamic cycle involved in charging and discharging the interacting molecules.

We estimate the nonpolar contributions to the binding free energy ($\Delta\Delta G_{np}$) from the curvature-corrected buried surface area at the molecular interface, ΔCA, and the microscopic surface tension parameter, γ_{aw}, using the "solvent-accessible surface area" (SASA) method ($\Delta\Delta G_{np} = \gamma_{aw}(\Delta CA)$; Nicholls et al., 1991; Sharp et al., 1991). In previous applications, a value for γ_{aw} of 50 cal/mol/Å^2 has been used to account for both the enhanced van der Waals interactions in the complex (relative to solvent) and the hydrophobic effect (Honig et al., 1993; Friedman and Honig, 1995). This approach has been used to account for nonpolar forces in the binding of anthracycline antibiotics to DNA (Baginski et al., 1997), as well as for the binding of peptides to MHC class I proteins (Froloff et al., 1997), and is described in detail in those publications. In the specific case of the λcI repressor–operator complex, 3884 Å^2 of solvent-accessible surface area is buried upon binding. The corresponding value of ΔCA is 2996 Å^2. Using these values, we find that $\Delta\Delta G_{np}$ drives binding by ~150 kcal/mol.

To fully account for the overall binding free energy, we must also account for a number of entropic contributions. The loss of translational and rotational degrees of freedom upon binding opposes binding by ~10 kcal/mol, based on several previous estimates (Brady and Sharp, 1997; Gilson et al., 1997; Finkelstein and Janin, 1989; Baginski et al., 1997). The loss of side-chain degrees of freedom is accounted for with the empirical scale of Pickett and Sternberg (1993), which is based on the number of rotameric states available to a side chain in the unbound state and the degree of burial of each group. For the λcI repressor–operator complex, 39 amino acids are buried more than 40% at the protein-DNA interface; the resulting loss of side-chain conformational entropy on binding is ~50 kcal/mol.

Finally, the free energy of conformational changes in the protein and DNA upon binding, $\Delta\Delta G_{strain}$, described by Eq. 12, are also expected to oppose binding (Froloff et al., 1997). The term $\Delta\Delta G_{strain}$ includes both electrostatic and nonelectrostatic contributions to the free energy of structural changes that accompany binding. Although it is not explicitly calculated, an upper limit of ~10 kcal/mol for $\Delta\Delta G_{strain}$ can be estimated from the folding free energy of most proteins (Froloff et al., 1997).

Adding the various free energy terms (+73 kcal/mol for electrostatics, −150 kcal/mol for the hydrophobic effect, +10 kcal/mol for translational and rotational entropy, +50 kcal/mol for side-chain degrees of freedom, and +10 kcal/mol for strain energy) yields a total binding free energy of −7 kcal/mol. Empirically accounting for the experimentally observed divalent ion effects gives a total binding free energy of −3 kcal/mol. Given the many uncertainties and approximations inherent in this model, it is remarkable that we arrive at a value for the total binding free energy that is the same order of magnitude as the observed value. It is reassuring, however, that similar values have been obtained with the same physical model applied to protein-peptide interactions (Froloff et al., 1997) and in recent work on antibiotic-DNA interactions (Baginski et al., 1997).

Although the generality of the results obtained here can only be ascertained with further study, it is likely that many of the features of the λcI repressor–operator complex will be found for other systems. Specifically, the high electrostatic cost of desolvating phosphate groups suggests that the electrostastic interactions will generally be found to oppose protein-DNA association, even given the strong Coulombic attractions between the oppositely charged moieties at the macromolecular interface. This conclusion suggests that a strong nonpolar driving force resulting from the burial of

interfacial surface area will prove to be a crucial feature for the stability of all protein-DNA complexes.

This work was supported by National Institutes of Health Grant GM-40371.

REFERENCES

Alexov, E. and M. Gunner. 1997. Incorporating protein conformational flexibility in the calculation of pH dependent protein properties. *Biophys. J.* 75:2075–2093.

Anderson, C. F., and M. T. Record. 1990. Ion distributions around DNA and other cylindrical polyions: theoretical descriptions and physical implications. *Annu. Rev. Biophys. Biophys. Chem.* 19:423–465.

Anderson, C. F., M. T. Record, and P. A. Hart. 1978. Sodium-23 NMR studies of cation-DNA interactions. *Biophys. Chem.* 7:301–316.

Antosiewicz, J., J. A. McCammon, and M. K. Gilson. 1994. Prediction of the pH-dependent properties of proteins. *J. Mol. Biol.* 238:415–436.

Antosiewicz, J., J. A. McCammon, and M. K. Gilson. 1996. The determinants of pK_as in proteins. *Biochemistry.* 35:7819–7833.

Baginski, M., F. Fogolari, and J. M. Briggs. 1997. Electrostatic and nonelectrostatic contributions to the binding free energies of anthracycline antibiotics to DNA. *J. Mol. Biol.* 274:253–267.

Bajaj, N. P., M. J. McLean, M. J. Waring, and E. Smekal. 1990. Sequence-selective, pH-dependent binding to DNA of benzophenanthridine alkaloids. *J. Mol. Recognit.* 3:48–54.

Beamer, L. J., and C. O. Pabo. 1992. Refined 1.8 Å crystal structure of the λ repressor-operator complex. *J. Mol. Biol.* 227:177–196.

Boger, D. L., B. J. Invergo, R. S. Coleman, H. Zarrinmayer, P. A. Kitos, S. A. Thompson, T. Leong, and L. W. McLaughlin. 1990. A demonstration of the intrinsic importance of stabilizing hydrophobic binding and non-covalent van der Waals contacts dominant in the non-covalent CC-1065/B-DNA binding. *Chem. Biol. Interact.* 73:29–52.

Brady, G. P., and K. A. Sharp. 1997. Entropy in protein folding and in protein-protein interactions. *Curr. Opin. Struct. Biol.* 7:215–221.

Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy minimization and dynamic calculations. *J. Comp. Chem.* 4:187–217.

Chen, S. W., and B. Honig. 1997. Monovalent and divalent salt effects on electrostatic free energies defined by the nonlinear Poisson-Boltzmann equation: application to DNA binding reactions. *J. Phys. Chem. B.* 101:9113–9118.

Ding, W.-D., and G. A. Ellestad. 1991. Evidence for hydrophobic interaction between calicheamicin and DNA. *J. Am. Chem. Soc.* 113: 6617–6620.

Finkelstein, A. V., and J. Janin. 1989. The price of lost freedom: entropy of bimolecular complex formation. *Protein Eng.* 3:1–3.

Friedman, R., and B. Honig. 1995. A free energy analysis of nucleic acid base stacking in aqueous solution. *Biophys. J.* 69:1528–1535.

Froloff, N., A. Windemuth, and B. Honig. 1997. On the calculation of binding free energies using continuum methods: application to MHC class I protein-peptide interactions. *Protein Sci.* 6:1293–1301.

Gibas, C. J., and S. Subramaniam. 1996. Explicit solvent models in protein pK_a calculations. *Biophys. J.* 71:138–147.

Gilson, M. K., J. A. Given, B. L. Bush, and J. A. McCammon. 1997. The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophys. J.* 72:1047–1069.

Gilson, M. K., and B. H. Honig. 1986. The dielectric constant of a folded protein. *Biopolymers.* 25:2097–2119.

Gilson, M. K., and B. Honig. 1988. Calculation of the total electrostatic energy of a macromolecular system: solvation energies, binding energies, and conformational analysis. *Proteins Struct. Funct. Genet.* 4:7–18.

Gilson, M. K., A. Rashin, R. Fine, and B. Honig. 1985. On the calculation of electrostatic interactions in proteins. *J. Mol. Biol.* 183:503–516.

Gilson, M. K., K. A. Sharp, and B. H. Honig. 1988. Calculating the electrostatic potential of molecules in solution: method and error assessment. *J. Comput. Chem.* 9:327–335.

Guenot, J., R. J. Fletterick, and P. A. Kollman. 1994. A negative electrostatic determinant mediates the association between the *Escherichia coli* trp repressor and its operator DNA. *Protein Sci.* 3:1276–1285.

Ha, J.-H., R. S. Spolar, and M. T. Record. 1989. Role of the hydrophobic effect in stability of site-specific protein-DNA complexes. *J. Mol. Biol.* 209:801–816.

Hendsch, Z. S., and B. Tidor. 1994. Do salt bridges stabilize proteins? A continuum electrostatic analysis. *Protein Sci.* 3:211–226.

Honig, B., K. Sharp, and A. Yang. 1993. Macroscopic models of aqueous solutions: biological and chemical application. *J. Phys. Chem.* 97: 1101–1109.

Jayaram, B., F. M. DiCapua, and D. L. Beveridge. 1991. A theoretical study of polyelectrolyte effects in protein-DNA interactions: Monte Carlo free energy simulations on the ion atmosphere contribution to the thermodynamics of λ repressor-operator complex formation. *J. Am. Chem. Soc.* 113:5211–5215.

Jayaram, B., K. A. Sharp, and B. Honig. 1989. The electrostatic potential of DNA. *Biopolymers.* 28:975–993.

Jayaram, B., S. Swaminathan, D. L. Beveridge, K. A. Sharp, and B. Honig. 1990. Monte Carlo simulation studies on the structure of the counterion atmosphere of B-DNA. Variations on the primitive dielectric model. *Macromolecules.* 23:3156–3165.

Jen-Jacobsen, L., M. Kurpiewski, D. Lesser, J. Grable, H. W. Boyer, J. M. Rosenberg, and P. J. Greene. 1983. Coordinate ion pair formation between *Eco*RI endonuclease and DNA. *J. Biol. Chem.* 258: 14638–14646.

Jorgensen, W. L., and J. Tirado-Rives. 1988. The OPLS potential functions for proteins. Energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* 110:1657–1666.

Klapper, I., R. Hagstrom, R. Fine, K. Sharp, and B. Honig. 1986. Focusing of electric fields in the active site of CuZn superoxide dismutase: effects of ionic strength and amino acid modification. *Proteins Struct. Funct. Genet.* 1:47–59.

Kong, X. P., R. Onrust, M. O'Donnell, and J. Kuriyan. 1992. Three-dimensional structure of the beta subunit of *E. coli* DNA polymerase III holoenzyme: a sliding DNA clamp. *Cell.* 69:425–437.

Ladbury, J. E., J. G. Wright, J. M. Sturtevant, and P. B. Sigler. 1994. A thermodynamic study of the trp repressor-operator interaction. *J. Mol. Biol.* 238:669–681.

Lundback, T., C. Cairns, J.-A. Gustafsson, J. Carlstedt-Duke, and T. Hard. 1993. Thermodynamics of the glucocorticoid receptor-DNA interaction: binding of wild-type GR DBD to different response elements. *Biochemistry.* 32:5074–5082.

Manning, G. S. 1969. Limiting laws and counterion condensation in polyelectrolyte solutions. II. Self-diffusion of the small ions. *J. Chem. Phys.* 51:934–938.

McCoy, A. J., V. C. Epa, and P. M. Colman. 1997. Electrostatic complementarity at protein/protein interfaces. *J. Mol. Biol.* 268:570–584.

Misra, V. K., J. L. Hecht, K. A. Sharp, R. A. Friedman, and B. Honig. 1994a. Salt effects on protein-DNA interactions: the λcI repressor and *Eco*RI endonuclease. *J. Mol. Biol.* 238:264–280.

Misra, V. K., and B. Honig. 1995. On the magnitude of the electrostatic contribution to ligand-DNA interactions. *Proc. Natl. Acad. Sci. USA.* 92:4691–4695.

Misra, V. K., and B. Honig. 1996. The electrostatic contribution to the B to Z transition of DNA. *Biochemistry.* 35:1115–1124.

Misra, V. K., K. A. Sharp, R. A. Friedman, and B. Honig. 1994b. Salt effects on ligand-DNA binding: minor groove binding antibiotics. *J. Mol. Biol.* 238:245–263.

Mol, C. D., A. K. Muir, M. Cygler, J. S. Lee, and W. F. Anderson. 1994a. Structure of an immunoglobulin Fab fragment specific for triple-stranded DNA. *J. Biol. Chem.* 269:3615–3622.

Mol, C. D., A. K. Muir, J. S. Lee, and W. F. Anderson. 1994b. Structure of an immunoglobulin Fab fragment specific for poly(dG) · poly(dC). *J. Biol. Chem.* 269:3605–3614.

Nelson, H. C. M., and S. T. Sauer. 1985. Lambda repressor mutations that increase the affinity and specificity of operator binding. *Cell.* 42: 549–558.

Nicholls, A., and B. Honig. 1991. A rapid finite difference algorithm, utilizing successive over-relaxation to solve the Poisson-Boltzmann equation. *J. Comp. Chem.* 12:435–445.

Nicholls, A., K. A. Sharp, and B. Honig. 1991. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.* 11:281–296.

Nilsson, L., and M. Karplus. 1984. Energy functions for energy minimization and dynamics of nucleic acids. *J. Comp. Chem.* 1:591–616.

Overman, L. B., and T. M. Lohman. 1994. Linkage of pH, anion and cation effects in protein-nucleic acid equilibria. *Escherichia coli* SSB protein-single stranded nucleic acid interactions. *J. Mol. Biol.* 236:165–178.

Phillips, K., and S. E. V. Phillips. 1994. Electrostatic activation of the *Escherichia coli* methionine repressor. *Structure.* 2:209–316.

Phillips, S. E. V., I. Manfield, I. Parsons, B. E. Davidson, J. B. Rafferty, W. S. Somers, D. Margarita, G. N. Cohen, I. Saint-Girons, and P. G. Stockley. 1989. Cooperative tandem binding of *met* repressor of *Escherichia coli. Nature.* 341:711–720.

Pickett, S. D., and M. J. E. Sternberg. 1993. Empirical scale of side-chain conformational entropy in protein folding. *J. Mol. Biol.* 231:825–839.

Pranata, J., S. G. Wierschke, W. L. Jorgensen. 1991. OPLS potential functions for nucleotide bases. Relative association constants of hydrogen-bonded base pairs in chloroform. *J. Am. Chem. Soc.* 113: 2810–2819.

Record, M. T., C. F. Anderson, and T. M. Lohman. 1978. Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening, and ion effects on water activity. *Q. Rev. Biophys.* 11: 103–178.

Record, M. T., J.-H. Ha, and M. A. Fisher. 1991. Analysis of equilibrium and kinetic measurements to determine thermodynamic origins of stability and specificity and mechanism of formation of site-specific complexes between proteins and helical DNA. *Methods Enzymol.* 208: 291–343.

Sarai, A., and Y. Takeda. 1989. λ repressor recognizes the approximately 2-fold symmetric half-operator sequences asymmetrically. *Proc. Natl. Acad. Sci. USA.* 86:6513–6517.

Sauer, R. T., S. R. Jordan, and C. O. Pabo. 1990. Lambda repressor: a model system for understanding protein-DNA interactions and protein stability. *Adv. Protein Chem.* 40:1–61.

Schultz, S. C., G. C. Shields, and T. A. Steitz. 1991. Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees. *Science.* 253: 1001–1007.

Senear, D. F., and G. K. Ackers. 1990. Proton-linked contributions to site-specific interactions of lambda cI repressor and OR. *Biochemistry.* 29:6568–6577.

Senear, D. F., and R. Batey. 1991. Comparison of operator-specific and nonspecific DNA binding of the lambda cI repressor: [KCl] and pH effects. *Biochemistry.* 30:6677–6688.

Shakked, Z., G. Guzikevich-Guerstein, F. Frolow, D. Rabinovich, A. Joachimiak, and P. B. Sigler. 1994. Determinants of repressor/operator recognition from the structure of the trp operator binding site. *Nature.* 368:469–473.

Sharp, K. A. 1995. Polyelectrolyte electrostatics: salt dependence, entropic, and enthalpic contributions to free energy in the nonlinear Poisson-Boltzmann model. *Biopolymers.* 36:227–243.

Sharp, K. A., and B. Honig. 1990a. Calculating total electrostatic energies with the nonlinear Poisson-Boltzmann equation. *J. Phys. Chem.* 94: 7684–7692.

Sharp, K. A., and B. Honig. 1990b. Electrostatic interactions in macromolecules: theory and applications. *Annu. Rev. Biophys. Biophys. Chem.* 19:301–332.

Sharp, K. A., A. Nicholls, R. F. Fine, and B. Honig. 1991. Reconciling the magnitude of the microscopic and macroscopic hydrophobic effects. *Science.* 252:106–109.

Sitkoff, D., K. A. Sharp, and B. Honig. 1994. Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.* 98:1978–1988.

Spolar, R. S., and M. T. J. Record. 1994. Coupling of local folding to site-specific binding of proteins to DNA. *Science.* 263:777–784.

Tanford, C. 1970. Protein denaturation, part C. *Adv. Protein Chem.* 24: 1–95.

Tidor, B., K. K. Irikura, B. R. Brooks, and M. Karplus. 1983. Dynamics of DNA oligomers. *J. Biomol. Struct. Dyn.* 1:231–252.

Warwicker, J., B. P. Engelman, and T. A. Steitz. 1987. Electrostatic calculations and model-building suggest that DNA bound to CAP is sharply bent. *Proteins.* 2:283–289.

Warwicker, J., D. Ollis, F. M. Richards, and T. A. Steitz. 1985. Electrostatic field of the large fragment of *Escherichia coli* DNA polymerase I. *J. Mol. Biol.* 186:645–649.

Weiner, S. J., P. A. Kollman, D. T. Nguyen, and D. A. Case. 1986. An all atom force field for simulations of proteins and nucleic acids. *J. Comput. Chem.* 7:230–252.

Wyman, J., and S. J. Gill. 1990. Binding and Linkage. Functional Chemistry of Biological Macromolecules. University Science Books, Mill Valley, CA.

Yadav, P. N., J. S. Yadav, E. Arnold, and M. J. Modak. 1994. A computer-assisted analysis of conserved residues in the three-dimensional structures of the polymerase domains of *Escherichia coli* DNA polymerase I and HIV-1 reverse transcriptase. *J. Biol. Chem.* 269:716–720.

Yadav, P. N., J. S. Yadav, and M. J. Modak. 1992. Binding of DNA to large fragment of DNA polymerase. I. Identification of strong and weak electrostatic forces and their biological implications. *J. Biomol. Struct. Dyn.* 10:311–316.

Yang, A.-S., M. R. Gunner, R. Sampogna, K. Sharp, and B. Honig. 1993. On the calculation of pK_as in proteins. *Proteins.* 15:252–265.

Yang, A.-S., and B. Honig. 1993. On the pH dependence of protein stability. *J. Mol. Biol.* 231:459–474.

Yang, L., S. Weerasinghe, P. E. Smith, and B. M. Pettitt. 1995. Dielectric response of triplex DNA in ionic solution from simulations. *Biophys. J.* 69:1519–1527.

Zacharias, M., B. A. Luty, M. E. Davis, and J. A. McCammon. 1992. Poisson-Boltzmann analysis of the λ repressor-operator interaction. *Biophys. J.* 3:1280–1285.