

# The Electrostatic Potential of B-DNA

B. JAYARAM, KIM A. SHARP, and BARRY HONIG, *Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032*

## Synopsis

Electrostatic potentials around DNA are obtained by solving the nonlinear Poisson-Boltzmann (PB) equation. The detailed charge distribution of the DNA and the different polarizabilities of the macromolecule and solvent are included explicitly in the calculations. The PB equation is solved using extensions of a finite difference approach applied previously to proteins. Electrical potentials and ion concentrations are compared to those obtained with simpler models. It is found that the shape of the dielectric boundary between the macromolecule and solvent has significant effects on the calculated potentials near the surface, particularly in the grooves. Sequence-specific patterns are found, the most surprising result being the existence of positive regions of potential near the bases in both the major and minor grooves. The effect of solvent and ionic atmosphere screening of phosphate-phosphate repulsions is studied, and an effective dielectric function, appropriate for molecular mechanics simulations, is derived.

## INTRODUCTION

Nucleic acids are highly charged molecules and electrostatic interactions must play an important role in many aspects of their structure and function. It has become clear, for example, that electrical forces are an important component of protein-DNA and drug-DNA interactions.<sup>1-4</sup> A complete model of electrostatic effects in DNA should include an accurate description of the shape and charge distribution of the macromolecule and a rigorous treatment of DNA-solvent interactions. This has been difficult to achieve.

Theoretical studies of nucleic acids have to some extent been divided into two distinct areas. One area has focused almost entirely on electrostatic properties such as the potentials and ion atmosphere that surround the macromolecule. These have often been obtained from solutions to the Poisson-Boltzmann (PB) equation for simplified models of DNA. The other area has focused on details of DNA conformation and dynamics, for example, as in molecular mechanics simulations, but has tended either to ignore solvent and ionic strength effects, or to treat them in an ad hoc fashion. The present study presents a method for treating electrostatic interactions in DNA that combines some of the best features of both approaches. In particular, solutions to the PB equation are obtained numerically for a detailed atomic-level description of the DNA. This makes it possible to study electrostatic potentials close to the surface of DNA where simplified geometric models cannot be valid, while incorporating environmental effects (i.e., effects due to solvent and ion atmosphere) into the calculation of properties that depend upon the details of DNA conformation.

The approach introduced in this work is based upon finite difference solutions to the PB equation (the FDPB method) for molecules of arbitrary shape and charge distribution. An important feature of the FDPB method is that it can explicitly account for the different polarizabilities of the macromolecule and solvent, a factor shown to have important consequences in proteins where the method has been previously applied.<sup>5,6</sup> The application of the FDPB method to DNA has required a number of developments, including obtaining numerical solutions to the nonlinear PB equation and the introduction of periodic boundary conditions. These are described in the methods section below. The remainder of this section briefly summarizes previous work in the general area of DNA electrostatics and outlines its relationship to the current study.

The simplest electrostatic models of DNA treat the molecule either as a string of point charges or as a rod with a uniform charge density at its surface, around which counterions can condense above a critical charge density.<sup>7</sup> The concept of condensation has been very successful in explaining some colligative and thermodynamic properties of polyelectrolytes.<sup>7-9</sup> Zimm and LeBret<sup>10</sup> redefined the concept of counterion condensation and showed that it arises from solutions to the PB equation in cylindrical geometry. Other workers have also examined solutions to the PB equations for charged cylinders<sup>8,11,12</sup> and found an accumulation of counterions near the DNA even at low ionic strengths, in qualitative agreement with the condensation model.

The use of the PB equation for polyelectrolytes has been criticized principally because it neglects the finite size of the solvent ions and the spatial correlations between them.<sup>13-15</sup> However, quantitative comparisons have been made of cylindrical solutions to the PB equation with Monte Carlo simulations<sup>16-18</sup> and hypernetted chain (HNC) calculations<sup>19</sup> where discrete ion effects can be included explicitly. These studies have shown that for univalent ions with radii that are small compared with the thickness of the DNA strand, the nonlinear PB equation is a reasonable approximation. The PB equation underestimates the concentration of counterions very close to the DNA by about 12–18%<sup>18,19</sup> but otherwise agrees well with the more detailed solutions in spite of its continuum treatment of the solvent. On the other hand, Monte Carlo and HNC calculations have not been extended to account for the shape, charge distribution, and dielectric boundary of DNA in solution.

Calculations based on cylindrical models, however rigorously they are applied, are inherently limited when describing electrostatic features close to DNA, since they neglect the distribution of charge and the grooved structure of the molecule. Sequence-specific electrostatic phenomena, if they occur, must arise from the different partial charge distribution of the bases and the location of these charges with respect to each other and to the dielectric boundary. Some of the earliest calculations on electrostatic potentials around nucleic acids including detailed descriptions of partial charges were carried out by Pullman and co-workers.<sup>20</sup> The calculations were carried out using a vacuum dielectric constant and hence gave values of the potential near the DNA that were too high to be biologically relevant (hundreds of kilocalories per mole). Weiner et al.<sup>21</sup> performed similar calculations using coulombic potentials combined with a distance-dependent dielectric constant that is intended to account, in an ad hoc way, for the effects of solvent screening.

Neither of these models treats solvent and ion screening or the dielectric boundary satisfactorily, but both studies point to the possibility of base sequence specific electrostatic effects.

The first attempts to use the PB equation to obtain electrostatic potentials around DNA while including a detailed description of its charge distribution were carried out by Klein and Pack.<sup>22,23</sup> Their method involves an iterative solution to a combination of coulombic potentials from the fixed macromolecular charges and the distribution of mobile charges obtained from the Boltzmann equation. This is equivalent to a solution of the PB equation with a uniform dielectric constant throughout space. To investigate the effect of dielectric discontinuity, Troll et al.<sup>24</sup> used a clay model of DNA in an electrolyte tank to estimate the electrostatic interactions between charges. Although the methods used are very different from that of the present work, both studies attempt to investigate the effect of the dielectric discontinuity on the electrostatic potentials of DNA.

The only other attempt reported so far to incorporate the effects of the dielectric boundary in calculations on DNA was based on the Tanford-Kirkwood model modified to account for the solvent accessibility of the individual charges.<sup>25</sup> These calculations predict the highest concentrations of the mobile cations in the minor groove of DNA, in agreement with Klein and Pack,<sup>22,23</sup> and provide better agreement with ion distribution data obtained by charge transfer experiments than do PB solutions for a cylindrical, uniform dielectric model.<sup>26</sup> While this and related work<sup>27</sup> illustrates the importance of dielectric boundary effects, a spherical model is inherently limited when describing the detailed shape of a complex molecule such as DNA.

As pointed out above, molecular mechanics simulations account for the detailed atomic structure of DNA but encounter difficulties in their treatment of the solvent. In fact, standard force fields will severely distort the DNA structure unless drastic assumptions are used to introduce the effects of solvent. For example, Levitt<sup>28</sup> set the phosphate charge to zero under the assumption that the solvent and counterions would totally screen their interactions. Tidor et al.<sup>29</sup> reduced the phosphate charge to -0.2 to account for ionic strength effects and used a distance-dependent dielectric constant to represent the effects of solvent screening. Prabhakaran and Harvey<sup>30</sup> and Rao and Kollman<sup>31</sup> carried out molecular mechanics studies of the energetics associated with the formation of intercalation sites, again using a distance-dependent dielectric constant to treat electrostatic interactions. These empirical approximations were necessary components in early efforts to investigate DNA structure and dynamics, but while computationally convenient, they are not really satisfactory. The distance-dependent dielectric constant, for example, tends to underestimate solvent screening of charges near the surface of a macromolecule,<sup>32</sup> while the use of scaled phosphate charges precludes the study of salt effects on DNA structure and dynamics.

Calculations have been carried out on small DNA fragments including water and counterions explicitly<sup>33</sup>; however, the computational demands involved in including many waters and ions in DNA simulations limits the range of the applicability of this approach. Moreover, the ability of current water models to account for dielectric behavior and ion solvation is still unclear.<sup>34</sup> For this reason, it would appear to be of considerable value to use

the PB equation as a basis for obtaining effective dielectric constants that can be incorporated into molecular mechanics simulations.

As outlined above, in this work we use an approach (the FDPB method) that treats the macromolecule in atomic detail, and accounts for the effects of the dielectric boundary, solvent screening, and salt concentration through numerical solutions to the PB equation. Our previous work on proteins has shown that this approach provides an intuitively understandable picture of electrostatic phenomena while providing the basis for accurate quantitative calculations.<sup>6,35,36</sup> Specific questions addressed in this study on DNA include the effects of shape and dielectric boundary on potentials around and at the surface of DNA, the salt concentration dependence of counter ion distribution, the effect of the dielectric boundary and salt concentration on phosphate-phosphate interactions, and the effect of base sequence on electrostatic potentials.

## METHODS

### Model

The DNA molecule is described in terms of its three-dimensional structure as determined by the coordinates of all its atoms. The coordinates of the idealized form of B-DNA were generated from the local coordinates of Arnott and Hukins.<sup>37</sup> The decamer sequence chosen here contains the Eco RI endonuclease recognition site and the netrospin binding site, and is the same as that studied experimentally by Marky and Breslauer.<sup>38</sup>

The details of the model have been described previously for proteins.<sup>6,36</sup> Charges are assigned to the center of each atom and are considered to be embedded in a low dielectric medium consisting of the volume enclosed by the solvent-accessible surface of the nucleic acid. In this work, the values of the charges were taken from the AMBER force field.<sup>39</sup> The surrounding solvent is treated as a continuum of dielectric constant 80 with an electrolyte behaving according to the PB equation. In order to solve the PB equation, the DNA and a region of surrounding solvent are mapped onto a  $65 \times 65 \times 65$  point cubic lattice. The midpoint of each grid line joining two lattice points is assigned a dielectric constant of value 80 if it lies in the solvent, and of value 2 if it lies within the solvent accessible volume of the DNA molecule. Lattice points lying further than one ion exclusion radius outside the molecular surface are assigned a Debye-Huckel parameter corresponding to the required salt concentration, while all other lattice points are assigned a value of zero. The ion exclusion radius is taken as 2.0 Å (corresponding approximately to the hydrated radius of a sodium ion) and the salt is assumed to be a 1:1 electrolyte. Each atomic point charge, which generally does not lie on a lattice point, is distributed over the eight closest lattice points in such a way as to preserve the monopole and dipole moments of the charge distribution.<sup>6</sup>

All calculations are performed with an explicit representation of one complete turn of DNA within the lattice. This corresponds to a spacing between lattice points of about 0.53 Å. In cases where potentials at distances greater than 16 Å are of interest, three turns of DNA are used, giving a lattice spacing of 1.6 Å. Calculations were performed with the DelPhi program<sup>6,36</sup> running on

a Convex C1-XP computer and/or on a Star ST-100 array processor hosted by a Digital microvax.

### The Nonlinear PB Equation

The calculation of the potential is performed as described previously,<sup>6,36</sup> except that the PB equation is not linearized due to the highly charged nature of DNA:

$$\nabla \cdot (\epsilon(\mathbf{x}) \nabla \phi(\mathbf{x})) - \bar{\kappa}^2(\mathbf{x}) \sinh[\phi(\mathbf{x})] + 4\pi\rho_f(\mathbf{x}) = 0 \quad (1)$$

where  $\phi$  is the potential expressed in units of  $kT/e$ ,  $k$  being Boltzmann's factor,  $T$  the absolute temperature, and  $e$  the unit electron charge;  $\epsilon$  is the dielectric constant; and  $\rho_f$  is the fixed charge density.  $\bar{\kappa}$  is the modified Debye-Hückel parameter (which does not include the dielectric constant), given by

$$\bar{\kappa} = \epsilon^{1/2} \kappa \quad (2)$$

where  $1/\kappa$  is the Debye length.  $\phi$ ,  $\epsilon$ ,  $\kappa$ , and  $\rho_f$  are all functions of the vector coordinate  $\mathbf{x}$ . To obtain a finite difference representation of Eq. (1),  $\sinh(\phi)$  is expanded in a power series:

$$\sinh(\phi) = \sum \phi^{2n+1} / (2n + 1)! \quad (3)$$

For each lattice point, a finite difference equation is constructed that relates the potential at that point to the potential at its six neighboring points. This results in 65<sup>3</sup> simultaneous equations of the form

$$\phi'_0 = \frac{(\sum \epsilon_i \phi_i + 4\pi q_0/h)}{(\sum \epsilon_i + \bar{\kappa}^2 h^2 [1 + \phi_0^2/3! + \phi_0^4/5! \cdots + \phi_0^{2n}/(2n + 1)!])} \quad (4)$$

Except for points at the edge of the lattice, where some value for the potential must be assumed (see below), these equations are solved by iterative relaxation. An initial guess is made for the potentials at each point (usually zero for convenience), and a new value for the potential at each point,  $\phi'_0$ , is obtained from the previous value at that point,  $\phi_0$ , the previous values of potential at the neighboring points,  $\phi_i$  ( $i = 1-6$ ), the dielectric constants associated with the grid lines connecting the point to its six neighbors,  $\epsilon_i$ , and the Debye-Hückel parameter for that point,  $\bar{\kappa}$ . This iteration process is continued until convergence is obtained (a mean change in potential per lattice point of less than  $10^{-5} kT/e$  was used here).

The terms in powers of  $\phi$  in the denominator on the right-hand side of Eq. (4) represent successive corrections to the linear PB equation. With the charge densities and salt concentrations used in modeling DNA in this work, it was found that inclusion of three terms of the power series reduced the average error in final potential due to truncation to less than 3%. Inclusion of the nonlinear terms in Eq. (4) results in an implicit equation, with  $\phi_0$  occurring on both sides of the equation. In some cases, particularly at high salt concentrations (large  $\kappa$ ), iteration of these equations can become numerically unstable,

the potential values oscillating or diverging to infinity. In these cases, stability can be recovered by underrelaxation, i.e., by multiplying the computed change in potential obtained from Eq. (4),  $\Delta\phi_0 = (\phi'_0 - \phi_0)$ , by a relaxation parameter, typically 50–80%, before updating the potentials. The stability of the iteration is indicated by smooth convergence, and independence of the final potential values on the initial guess (further details of finite difference and relaxation methods can be found in any standard numerical methods book, see, e.g., Ref. 40).

### Boundary Conditions

Lattice points on the boundary of the grid lack one or more neighboring points and must be treated separately. A method for introducing periodic boundary conditions was developed for this work in order to be able to represent the field of an infinite piece of DNA. The DNA was mapped into the grid so that the helical axis was aligned along the Z axis of the grid, with exactly one or three turns (10 or 30 base pairs for B-DNA) inside the grid. To iterate the potentials at the boundary points on the Z-axis edges of the grid, values for the potential at the missing neighbors were taken from the corresponding points on the opposite edge of the grid. On the X and Y axis edges of the grid, the potential  $\phi_i$  at the  $i$ th grid point was calculated analytically using the Debye–Hückel equation<sup>6</sup>:

$$\phi_i = \sum (q_j e^{-\kappa r_{ij}}) / 80r_{ij} \quad (5)$$

where the sum is over  $j$ , for all the fixed charges,  $q_j$ , and  $r_{ij}$  is the distance of the  $j$ th charge from the  $i$ th boundary point. This expression approximates the analytical solution far from the molecule.

### Tests of Accuracy

The numerical algorithm used here has been tested previously<sup>36</sup> on spherical models where analytical solutions to the linear PB equation are known. The numerical results agree with the analytical solutions over a range of conditions to within a few percent, except within one to two grid spacings of a charge or a dielectric boundary, where the errors range from 5–15%. We have also tested the method on uniformly charged cylindrical models using the nonlinear PB equation, making comparisons to numerical solutions obtained by the Runge–Kutta method.<sup>9,11</sup> The results are accurate everywhere to within 1%, except near the dielectric boundary, where errors up to 7% are encountered. Thus, our numerical method succeeds in reproducing the results of previous studies on simplified models of DNA.

### Ion Distribution

The net charge density,  $\rho(\mathbf{x})$ , and the counterion concentration,  $\rho_+(\mathbf{x})$ , at any point in solution were calculated from the potential,  $\phi$ , and the bulk salt concentration,  $C_0$ , using the Boltzmann equations

$$\rho = 2C_0 \cdot \sinh(\phi(\mathbf{x})) \quad (6)$$

$$\rho_+ = C_0 \exp(-\phi(\mathbf{x})) \quad (7)$$

The radially averaged charge density,  $\rho(r)$ , was calculated from

$$\rho(r) = \langle 2C_0 \cdot \sinh(\phi(r, z, \theta)) \rangle_{z, \theta} \quad (8)$$

where  $r, z, \theta$  are cylindrical coordinates defined relative to the helical axis. From the averaged charge density, the radial distance within which any fraction of the total counterion charge is located can be calculated. For example, with our representation of DNA the radius containing 76% of all the counterion charge,  $R_m$ , (where 0.76 is the fraction of condensed charge obtained with the Manning model<sup>7</sup> for the linear charge density corresponding to B-DNA), can be compared with the so-called Manning radius in the cylindrical PB model of Zimm and LeBret<sup>10</sup> and the uniform dielectric model of Pack and co-workers.<sup>22,23</sup>

### Effective Dielectric constants

The effective dielectric constant, which is perhaps better termed a screening parameter,  $\epsilon^e$ , for the interaction between any two charged atoms  $i$  and  $j$ , separated by a distance  $r_{ij}$ , is commonly defined as the ratio of the strength of interaction between the two charges (obtained either from experimental data or some form of calculation), compared to that which would occur in vacuum,  $\phi_v$  (i.e., using a dielectric constant of 1 in Coulomb's law). In this work, the calculated interaction is obtained by removing all fixed charges except at one of the atoms (the "source" charge) and calculating the potential  $\phi$  at the other atom(s). Thus,

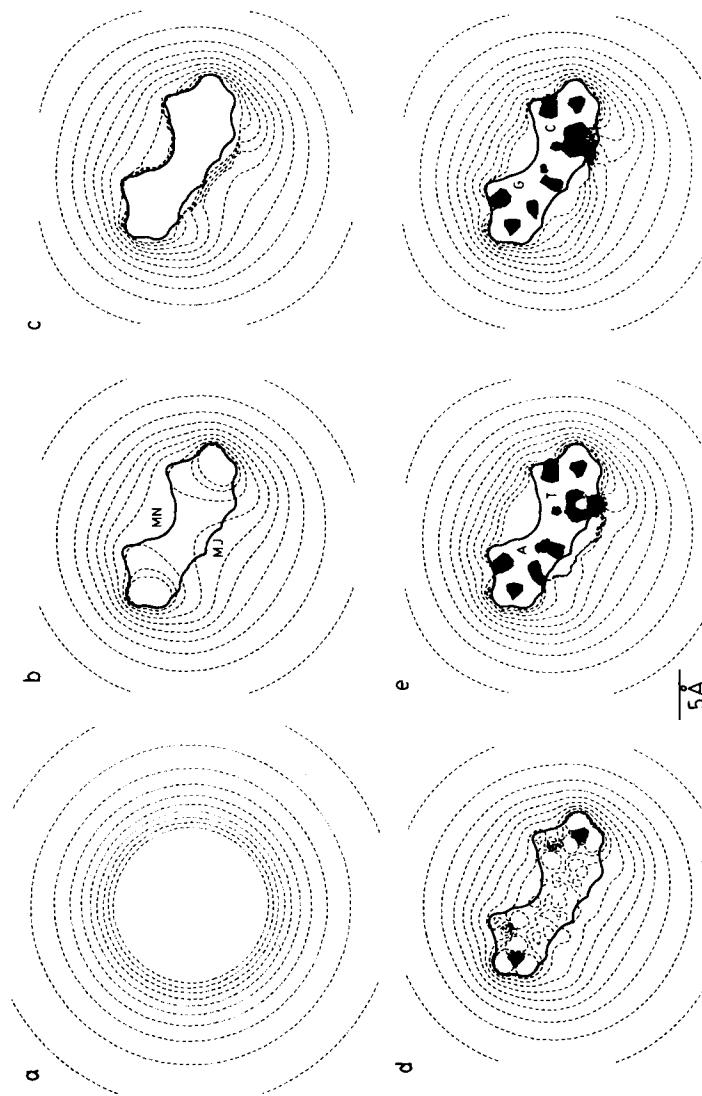
$$\epsilon_{ij}^e = \phi_v/\phi = q/\phi r_{ij} \quad (9)$$

where  $q$  is the charge on the source atom. When ions are present in the solvent, very effective screening can occur, resulting in screening constants of several hundred. It is convenient in these cases to redefine the screening constant as

$$\epsilon_{ij}^e = \phi_d/\phi = q \cdot \exp(-\kappa r_{ij})/\phi r_{ij} \quad (10)$$

which is the ratio of the interaction between two Debye-Hückel-type point ions in ionic solution,  $\phi_d$ , and that calculated from our solutions to the nonlinear PB equation. Thus in cases where no salt is present, deviations of  $\epsilon^e$  from the value of 80 (expected for two charges interacting in water) represent the effect on the screening of the shape and position of boundary between water and the less polarizable DNA. When salt is present, deviations from a value of 80 will also reflect the effect of exclusion of mobile ions from DNA, and alteration in their distribution due to the low dielectric region.

Each of the 20 phosphates in the middle turn of a three-turn piece of DNA was charged separately and the potential due to that charge at each other phosphate was evaluated. The potentials thus obtained for equivalent pairs of phosphates were averaged over all 20 phosphates, and the effective dielectric constants were determined from Eq. (9) at zero salt concentration and from Eq. (10) at physiological salt concentration (0.15M). Since each phosphate is positioned slightly differently with respect to the grid, this averaging greatly improves the precision of the finite difference calculation.<sup>36</sup>



**Fig. 1.** Two-dimensional isopotential contours in a plane perpendicular to the helix for various DNA models at physiological salt concentration. Except for panel (a), the plane lies between the middle two bases of a 10 base turn of helix. Contours (dotted lines) are at  $0.5 kT/e$  intervals from  $-0.5$  to  $-5.0 kT/e$ . Solid shading indicates regions of positive potential. The cross section of the solvent-accessible surface of DNA is shown as a bold solid line. (a) Linear lattice of negative charges placed 1.7 Å apart on the axis of a 9.5 Å radius cylinder, uniform dielectric of 80. (b) B-form DNA, surrounded by a 2.0 Å ion exclusion zone, with only the phosphates, a dielectric constant of 80 everywhere. (c) As for (b), but using a dielectric of 2 for DNA. (d) Same as (b) except that all partial charges corresponding to a poly(dA) · poly(dT) sequence are included. (e) As for (d) but using a dielectric of 2 for DNA. (f) As for (e), but using partial charges corresponding to a poly(dG) · poly(dC) sequence.

## RESULTS

### Electrical Potentials and Ion Concentrations

Electrostatic potentials around DNA at physiological salt concentration obtained from various models are shown in Fig. 1. The figure shows two-dimensional potential contours in a midplane slice perpendicular to the helix axis. Contours are drawn at intervals of  $0.5 \text{ kT/e}$  from  $-0.5$  to  $-5 \text{ kT/e}$ . Regions of positive potential are shown by solid shading. The Debye length is  $8 \text{ \AA}$  in all cases. Figure 1(a) is obtained for a linear lattice of charge. This model, and the model of a cylinder with a uniform charge density on the surface, studied by Klein, Anderson, and Record<sup>9</sup> and by Gueron and Weisbuch,<sup>11</sup> give identical potentials in solution up to and including the surface. Figure 1 (b and d) are obtained by solving the PB equation using coulombic potentials and therefore assume a dielectric constant of 80 both inside and outside the macromolecule. This is essentially the model assumed by Klein and Pack,<sup>22</sup> although their method of solving the PB equation differs from that used here. Figure 1 (c, e, and f) display the potentials obtained from the PB equation where the DNA and solvent are assigned dielectric constants of 2 and 80, respectively. Since Fig. 1(b) and Fig. 1(c) are obtained from identical charge distributions [as is the case for Fig. 1(d) and Fig. 1(e)], the differences in the potentials are due to the existence of a dielectric boundary in fig. 1(c) and Fig. 1(e).

The  $-0.5 \text{ kT/e}$  contour (the outermost) is essentially circular in all models. The contours become increasingly more elliptical as the distance from the DNA decreases. In Figure 2 the potentials of Fig. 1(c) along radii passing

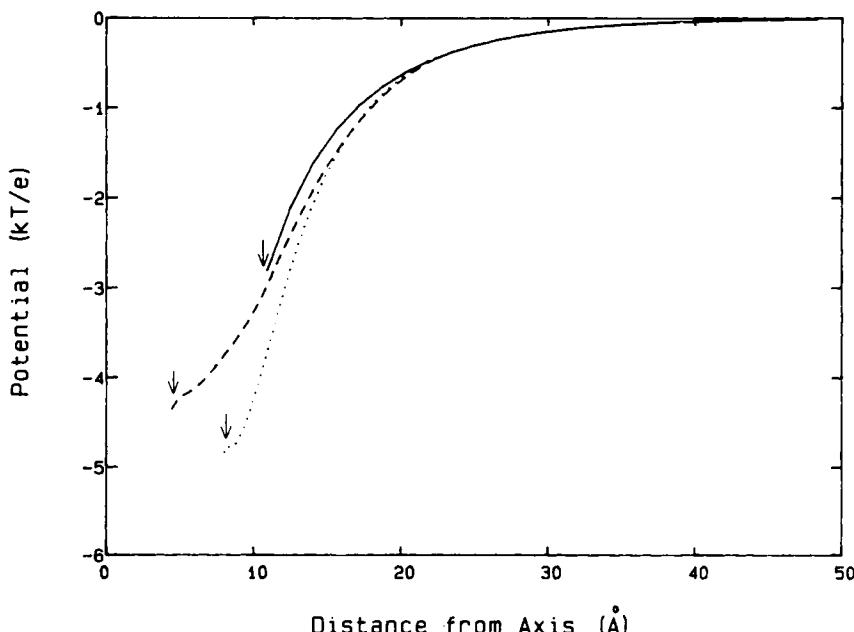


Fig. 2. Radial profiles of potential along a line perpendicular to the helix axis passing through the center of the minor groove (dashed line) and through a phosphate group (dotted line) of the DNA model in Fig. 1(c) and for the cylindrical model of Fig. 1(a) (solid line).

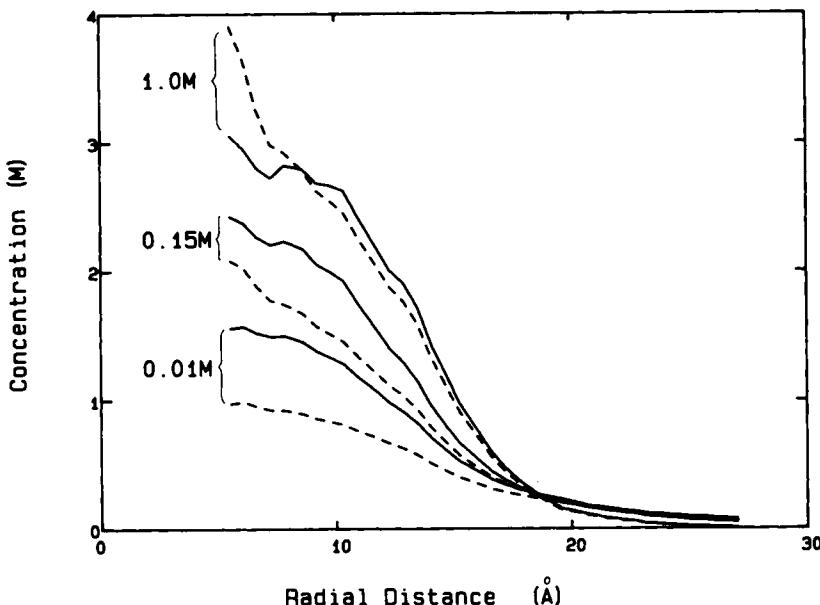


Fig. 3. Radially averaged net charge density in solvent around DNA as a function of distance from the helical axis for uniform dielectric of 80 (dashed lines) and for the two dielectric model (solid lines). Salt concentrations are indicated on the curves.

through the center of the minor groove and through a phosphate group are plotted along with potentials obtained from the cylindrical model of Fig. 1(a). The location of the surface is denoted by arrows. It is clear from Fig. 2 that the cylindrical model is accurate beyond about 20 Å, although it does fairly well even at shorter distances. On the other hand, the 9.5 Å radius typically used in cylindrical models does cut off the negative potential at values that are too small. On the basis of Fig. 2, a 5 or 6 Å cylindrical radius might yield more realistic potentials close to the surface of the DNA in the grooves. It is also worth pointing out that the potential 2 Å and farther from the surface of the minor groove (which is very close to the helical axis) is more negative than the potentials at the same distance from the surface of the phosphate group, even though the potential on a line going through the phosphates is more negative for a given distance from the helix axis.

Comparing Fig 1(b) and Fig. 1(c), the effect of representing the interior of DNA as a low dielectric cavity does not change the potentials due to the phosphates at distances more than 5–10 Å from the surface of the molecule. However, the potentials inside the molecule, and outside it close to the surface, are greatly increased in the two dielectric model, particularly in groove regions. The field is also increased in the groove at the molecular boundary, as can be seen from the more closely spaced contour lines in Fig. 1(c) compared to Fig. 1(b).

The effects of the dielectric discontinuity can also be seen in Fig. 3, which compares the radially averaged distribution of net charge density obtained with a uniform dielectric of 80 everywhere with that using a dielectric of 2 for the DNA. Results are shown for the B-form of poly(dA) · poly(dT), with only

TABLE I  
Calculated Manning Radii as a Function of Salt Concentration

Salt Concentration ( <i>M</i> )	<i>R</i> <sub>m</sub> (Å)
0.01	19.5 ± 3
0.1	15.4 ± 2
0.15	15.4 ± 2
1.0	12.9 ± 2

the phosphates charged, at salt concentrations of 0.01, 0.1, and 1.0*M*. The distributions are very similar at distances more than 12 Å from the axis, indicating that the Manning radius is insensitive to the dielectric discontinuity. However at small distances (< 8 Å) corresponding to regions in the groove, the two-dielectric model gives as much as 50% higher counterion concentrations at the two lower salt concentrations, since the lower dielectric constant of the DNA increases the potential in the grooves, resulting in greater ion accumulation. At the highest salt concentration, where the potentials everywhere are smaller, the concentration of ions in the groove is lower with the two dielectric model. This is due to a repulsion of the mobile charge from the low dielectric, weakly charged molecular surface in the grooves (that is, there is an increase in self-energy of the ions in the grooves as observed by Troll et al.<sup>24</sup>). The ions accumulate more around the phosphates, where the potential is enhanced by the low dielectric molecular interior, giving rise to a higher net charge concentration around 10 Å in the two dielectric model (top two curves in Fig. 3). This figure also shows that the local concentrations exceed 1*M* for all salt concentrations at distances less than 10 Å, which is in agreement with results reported from several other models discussed above.

The Manning radius as a function of salt concentration is shown in Table I. The radius increases with decreasing salt concentration, and agrees closely with results from other models.<sup>9,18,22</sup>

### Base Sequence Effects

The results shown in Fig. 1 indicate that the partial charges of the bases can significantly modify the electrostatic potential, particularly in regions close to the molecular surface. The partial atomic charges assumed in this study<sup>39</sup> are collected in Table II. In both the uniform [1(d)] and two dielectric models [1(e)], the bases are found to increase the negative potential in the grooves [compare 1(d) and 1(e) to 1(b) and 1(c), respectively] but the effect is much more significant in the two-dielectric model. In the latter case, inspection of Fig. 1 (c and e) reveals that the effect of the partial charges extends to about the -2 *kT/e* contour level, beyond which point the potentials are due almost entirely to the phosphates.

A second effect of the partial charges that is particularly pronounced in the two-dielectric model is the appearance of several areas of positive potential inside the DNA [see Fig. 1 (e and f)], some of which extend to the surface of the molecule in the grooves. The effect of base sequence on surface potential is shown in Fig. 4, which compares the potentials near the surface for poly(dA) ·

TABLE II  
Partial Atomic Charges<sup>a</sup> on Nucleic Acid Bases (in Atomic Units)

Atom	Adenine	Guanine	Thymine	Cytosine
N1	-0.760	-0.746	-0.739	-0.572
HN1	—	0.340	—	—
C2	0.571	0.842	1.113	0.938
N2	—	-0.758	—	—
H1N2	—	0.324	—	—
H2N2	—	0.333	—	—
O2	—	—	-0.529	-0.518
N3	-0.717	-0.702	-1.012	-0.791
HN3	—	—	0.370	—
C4	0.695	0.490	0.980	0.630
O4	—	—	-0.472	—
N4	—	—	—	-0.743
H1N4	—	—	—	0.338
H2N4	—	—	—	0.335
C5	-0.151	-0.088	-0.595	-0.230
CH3(C5)	—	—	0.097	—
C6	0.813	0.714	0.551	0.377
N6	-0.793	—	—	—
H1N6	0.339	—	—	—
H2N6	0.335	—	—	—
O6	—	-0.472	—	—
N7	-0.599	-0.575	—	—
C8	0.488	0.428	—	—
N9	-0.457	-0.479	—	—

<sup>a</sup>From Ref. 39.

poly(dT), poly(dG) · poly(dC), and the decamer containing the Eco RI recognition sequence and netropsin binding site. The minor groove of both homopolymer sequences is completely negative, but the poly(dA) · poly(dT) sequence is more negative by about 1.5  $kT/e$  [see also Fig. 1 (e and f)]. The major groove in the AT sequence is relatively neutral, or weakly negative. In contrast, the GC sequence shows a polarity within the major groove; positive patches of potential appear on the C side, with a more negative G side. This can also be seen in Fig. 1(f), where the contour lines at -2.5  $kT/e$  and lower balloon out from the DNA surface on the G side of the major groove (lower left-hand side), while on the C side of the groove, the contour lines are much closer to the surface. The decamer shows a more complex pattern of surface potentials, with various patches of neutral or positive potential, particularly near the dyad axis.

The sign and location of the partial charges that are exposed across the floor of the grooves account, in part, for the observed pattern of potentials. In an AT pair, the minor groove in B-DNA contains N3 (A) and O2 (T) atoms, which are both negative. In a GC pair the corresponding atoms, N3 (G) and O2 (C), are separated by the positive 2-amino group of guanine. Thus the potentials in the minor groove of an AT pair are expected to be more negative than for a GC pair. The major groove of an AT pair contains the negative N7 (A), positive 6-amino group of adenine, and then the negative O4 (T). In GC pairs the negative N7 (G) and O6 (G) are followed by the positive 4-amino

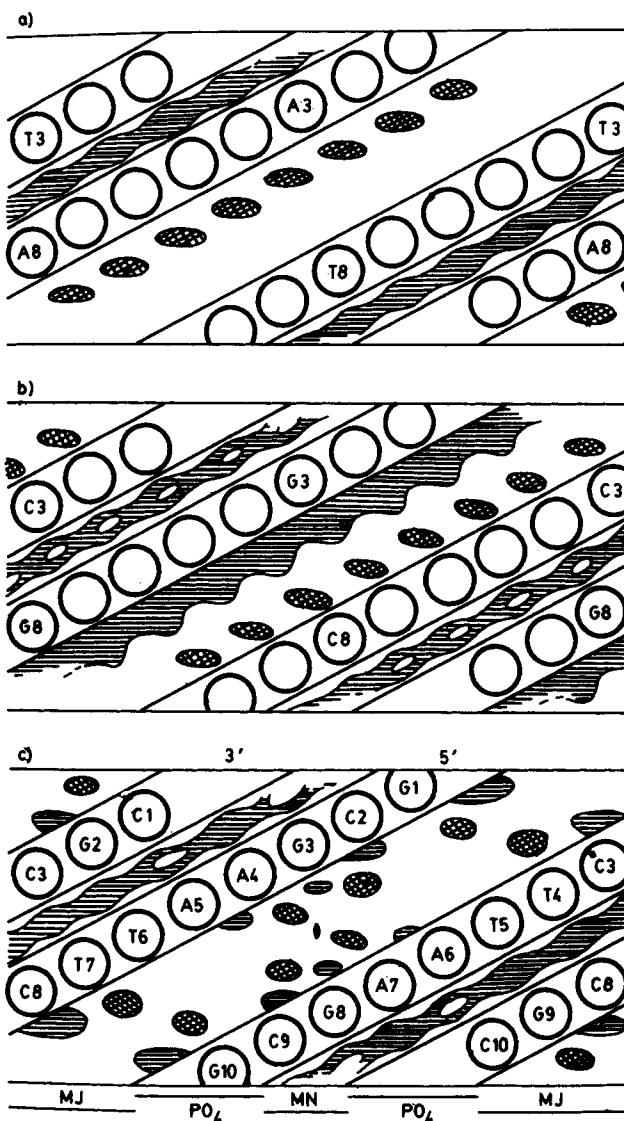


Fig. 4. Schematic representation of the potentials on the surface of B-DNA for different base sequences at physiological salt concentration. The DNA surface was mapped onto a cylinder that was cut parallel to the axis and unrolled. The helical axis is vertical, with one turn represented. The approximate positions of the phosphate groups are indicated by circles and the grooves are delimited by diagonal lines. MJ: major groove, MN: minor groove, PO<sub>4</sub>: phosphate region. Regions of the groove surface that are weakly negative or positive ( $> -0.5 \text{ } kT/e$ ) are crosshatched. Regions that are strongly negative ( $< -3 \text{ } kT/e$ ) are shaded. Regions of intermediate negative potential are unshaded. Base sequences are (a) poly(dA) · poly(dT), (b) poly(dG) · poly(dC), and (c) the decamer d(GCGAATTTCGC); the dyad axis is indicated in the center of the major groove.

group of cytosine. Thus, the GC pair, in contrast to the AT pair, is expected to show a greater polarity within the major groove, as observed in Fig. 4. Of course, these patterns are well known and follow the sequence specificity associated with hydrogen-bonding donor and acceptor groups.<sup>41</sup> What we have shown in this work is that the partial charges of the bases produce well-defined sequence-specific regions of positive and negative potential along the entire groove. Moreover, these regions are not simply functions of the sign of the partial charges since they do not appear on the surface of the DNA when a uniform dielectric model is used [see Fig. 1(d)].

### Effective Dielectric Constants

Effective dielectric constants for the interaction between any two phosphates were calculated for the two dielectric model at 0 and 0.15M salt concentration as described in the methods section. The results for poly(dA) · poly(dT) in the B-conformation are shown in Fig. 5 as a function of the angular distance between phosphates with respect to the helical axis. The results show that the effective dielectric constant at zero salt concentration can be as high as 105 if the two charges are located on the opposite sides of the macromolecule, and as low as 70 if they are on the same side. This variation is a result of the dielectric discontinuity, which effectively excludes field lines from the low dielectric region. This effect cannot be obtained in any model with coulombic or other pairwise potentials, since for these types of

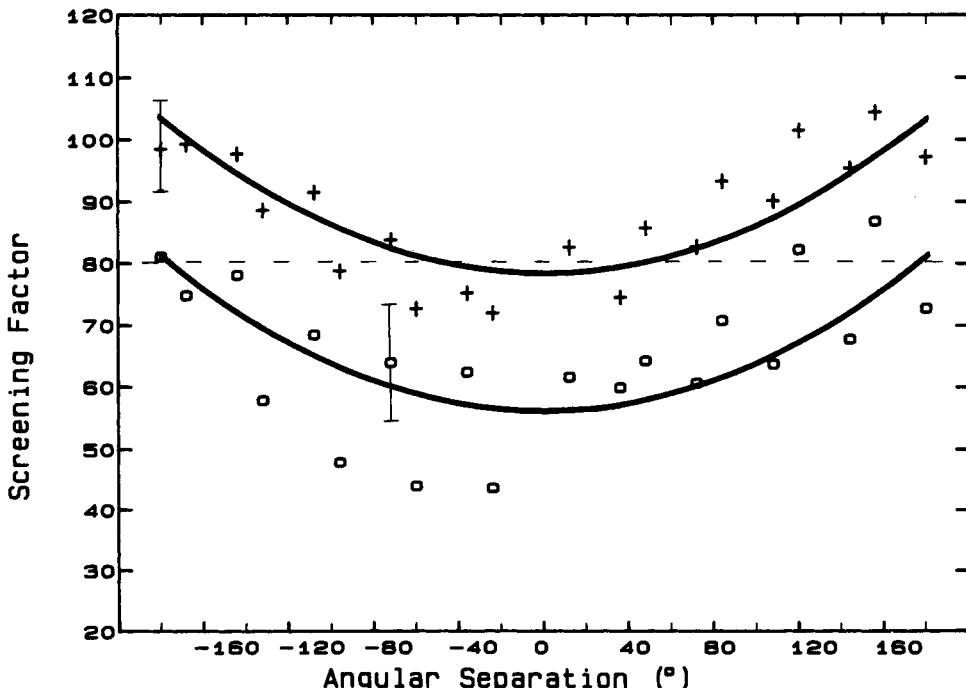


Fig. 5. Effective dielectric constants for the interaction between two phosphates, shown as a function of the angle defined by the planes containing the helical axis and each phosphate. Crosses indicate data for 0M salt concentration and circles for 0.15M salt concentration. Solid lines indicate the best fit to a harmonic function of the angle.

potential functions the field lines emanating from a charge are always spherically symmetric. Over the distance of one to two turns, the effect of distance along the helical axis is very small, the predominant effect being the angle dependence. At larger distances however, the dielectric tends to 80, as in the uniform dielectric case (results not shown). The curve can be fitted to a harmonic function (upper continuous curve in Fig. 5) in angle,

$$\epsilon^e = A\theta^2 + B$$

with coefficients  $A = 7.7 \pm 2 \times 10^{-4} \text{ deg}^{-2}$  and  $B = 78.6 \pm 3$ . These results are quantitatively very similar to those obtained by Troll et al.,<sup>24</sup> who found more effective screening for interactions between locations on the opposite side of DNA than for locations on the same side.

At physiological salt concentration (where the clay model system cannot be simply applied), the data can also be fit to a harmonic function with the same curvature, but shifted down,  $B = 56 \pm 3$  (lower curve in Fig. 5). Notice that  $\epsilon^e$  at physiological salt concentration is actually lower than at zero salt concentration. It should be recalled from Eq. (10), however, that  $\epsilon^e$  is multiplied by an exponential term in  $\kappa$ . The interpretation of  $\epsilon^e$ , as defined by Eq. (10), is that for values less than 80, the charges interact more strongly than would two Debye-Hückel-type ions in water, and for values greater than 80, this interaction is weaker.

## DISCUSSION

A theoretical and computational approach has been described that makes it possible to treat a variety of phenomena associated with electrostatic interactions in nucleic acids under physiological conditions. Electrostatic potentials are obtained from numerical solutions to the nonlinear PB equation (the FDPB method) based on a model of DNA that includes a detailed description of its charge distribution and the shape of the dielectric boundary it forms with the surrounding solvent. It is likely that the calculations reported here provide a reasonable description of solvent and electrolyte screening on the potentials. This expectation is based on the success of the FDPB method in recent studies of proteins (see, e.g., Refs. 6 and 35) and on the ability of PB calculations on cylindrical models of DNA to reproduce, with reasonable accuracy, the results obtained from Monte Carlo and HNC calculations on identical systems containing simple 1-1 electrolytes.<sup>16-19</sup> Judging by these comparisons, the approximations introduced into the PB model by treating the solvent and ions as a continuum are probably around 15%, in return for the ability for detailed shape and dielectric effects to be treated. For electrolytes containing large or multivalent ions, neglect of ion size and correlation effects would introduce larger errors, although the qualitative differences between cylindrical or uniform dielectric models and the current model would be expected to remain the same.

The relative ease with which it is now possible to describe the electrical potentials and ion atmosphere around DNA suggests a variety of applications. For example, it will be of interest to attempt to interpret the results of <sup>23</sup>Na-nmr measurements of ion concentrations<sup>42</sup> based on the results of FDPB

calculations. In this study, we have emphasized comparisons to simpler models for both DNA structure and DNA-solvent interactions. In particular, we have studied the validity of cylindrical descriptions of DNA, determined the effect of the dielectric boundary on the nature of the calculated potentials, and described a phosphate-phosphate screening function.

With regard to the effects of molecular shape, it is found that cylindrical models are accurate at distances greater than about 20 Å from the helical axis and remain qualitatively reasonable at distances as close as about 12 Å, i.e., an ion exclusion radius away from the phosphates, in a radial direction. However, close to the surface (< 10 Å) and particularly in the grooves, the potentials are sensitive both to the shape and polarizability of the macromolecule.

The overall effect of the dielectric boundary is to increase both the potentials and fields in the grooves, which should enhance the tendency of cations and dipolar groups to bind. Both negative and positive potentials occur on the van der Waals surface at physiological salt concentration, but negative potentials dominate in moving radially outward from the helical axis. The potentials near the surface are on the order of a few  $kT/e$  and appear to be physically reasonable. Previous calculations, which did not include the effects of solvent screening, produced surface potentials of many hundred  $kT$ .<sup>20</sup> The electrostatic potentials exhibit a strong sequence dependence, which can be qualitatively understood in terms of the locations of hydrogen-bonding donor and acceptors. However, the magnitudes of the potentials are strongly affected by the existence of the dielectric boundary.

A particularly striking effect of the dielectric boundary may be seen by comparing the potentials in the minor groove in Fig. 1(d) and Fig. 1(e). In the two-dielectric model [Fig. 1(e)], the -5 to -4  $kT/e$  contours balloon out more from the cleft of the minor groove compared to the uniform dielectric case [Fig. 1(d)]. This difference is due to the "focusing" of the electric field of the bases into the high dielectric solvent and away from the low dielectric walls of the cleft. The effect is reminiscent of the focusing of electric fields in the active site cleft of the protein superoxide dismutase.<sup>6</sup>

It is of interest to relate the effects of the dielectric discontinuity found in this work to those found by Troll et al.<sup>24</sup> In that study it was concluded that the increase both in the self-energy and ion-ion repulsions should significantly diminish the concentration of ions in the groove region, relative to the expectation based on Coulomb's law alone. In contrast, at low salt concentrations we find more ions in the groove due to the enhanced (relative to Coulombic) attractive phosphate potential. This effect dominates at low salt concentrations (including physiological), but at higher salt concentrations the phosphate attraction is effectively screened and it is at this point that the self-energy terms diminish the concentration in the grooves as predicted by Troll et al.<sup>24</sup>

In principle there is no numerical or mathematical reason to expect significant differences between the two studies since in both cases the potential is governed by Poisson's equation. Possible sources of discrepancy are as follows: (a) In the present study charges are placed on atomic centers that are inside the van der Waals envelope while the electrodes are kept outside of the clay model in contact with the high dielectric medium (Although Troll et al. point out the possibility of excavating a cavity in their clay model, inserting an

electrode, then refilling the cavity with a low dielectric material); (b) We approximate the electrical potential at the boundaries of our grid with a superposition of Debye-Hückel potentials (or coulombic potentials at zero salt) whereas the grounding of the tank in the study of Troll et al.<sup>24</sup> effectively clamps these potentials at zero. Both of these differences should reduce the phosphate potentials in the grooves as measured in the clay tank relative to those calculated in this study.

The results summarized in Fig. 5 suggest that solvent screening of phosphate-phosphate interactions is very effective, even for groups that are adjacent on a single strand. The effective dielectric constants that are obtained can provide a basis for incorporating a more realistic screening function in molecular mechanics simulations than those currently in use. For example, using an average value for  $\epsilon^e$  of 70 at physiological salt concentration, the function  $(q_i q_j / 70 r_{ij}) \exp(-\kappa r_{ij})$  should provide a reasonable approximation for the pairwise interactions of charges  $q_i$  and  $q_j$  between phosphate groups. This function should not be valid for interactions between atoms on the bases, which will in general be less effectively screened by solvent. The magnitude of base-base interactions will be considered in a future publication.

This screening function should be more appropriate for molecular mechanics simulations than those currently being used, since it includes the effect of dielectric screening and provides a means of including the effects of salt concentration explicitly. It is of interest to compare the screening predicted by the equation given in the previous paragraph to the reduction in phosphate-phosphate interactions predicted by a distant-dependent dielectric constant commonly used in simulations. For adjacent phosphates 7 Å apart, the screening function suggested above predicts an interaction energy of 0.67 kcal/mole at zero salt concentrations and 0.27 kcal/mole at 0.15M salt. The distant-dependent dielectric constant predicts an interaction energy of 6.8 kcal/mole at zero ionic strength and 0.27 kcal/mole if each phosphate charge is reduced to 0.2 to account for salt effects. The agreement at 0.15M is striking, but also highly fortuitous. It is due to the fact that the distance-dependent model underestimates dielectric screening by about an order of magnitude while the reduction of phosphate charges to -0.2 overestimates salt effects, also by an order of magnitude. It should be emphasized in this regard that the procedure of reducing the phosphate charge to account for salt effects is not supported by our results.

The major anticipated application of the FDPB method to molecular mechanics simulations is based on the possibility it offers of incorporating solvent screening effects that are based on a well-defined physical model. The PB model itself is of course based on assumptions, but these are known, the errors they introduce can be estimated, and in some cases they are subject to improvement. In any case, the PB equation should constitute a major improvement over the ad hoc assumptions currently being used. Moreover, the PB equation provides a means of accounting for changes in salt concentration and for changes in solvent screening as the conformation changes. For example, the screening constants calculated here would be expected to vary with the shape of the molecule and the effect can easily be studied with further calculations on other forms of DNA. Changes in screening would also be expected to occur during the course of a molecular mechanics simulation, and

while it is not known how large this effect is likely to be, current potential functions have no provision for dealing with the problem. However, FDPB calculations can at the very least reveal the magnitude of the effect and, moreover, it may be possible to devise a means of recalculating screening parameters during the course of a simulation. However, this lies beyond the scope of the present work.

In summary, the results of this study have shown that the detailed shape of the DNA as well as the existence of a dielectric boundary have important effects on the magnitude of the electrostatic potentials around the macromolecule. The approach that has been introduced should make it possible to make improved calculations of the magnitude of electrostatic interactions in different forms of DNA at different salt concentrations. Extensions of the FDPB method to include electrostatic charge-solvent interactions have been described for proteins<sup>43</sup> and will also provide a means of incorporating solvation energies, including the effects of the ion atmosphere, in calculations of total conformational energies of DNA, as well as the binding energies of charged ligands.

We are grateful to Drs. Hillary Rodman Gilson and Michael Gilson for helpful discussions on the nonlinear PB equation. We have also benefited greatly from discussions with Professor Peter Rossky on the validity of the PB equation. This work was supported by grants from the ONR (N00014-86-K-0483) and NIH (GM30518).

### References

1. McClarlin, J. A., Frederick, C. A., Wang, B. C., Greene, P., Boyer, H. W., Grable, J. & Rosenberg, J. M. (1986) *Science* **234**, 1526-1541.
2. Warwicker, J., Ollis, D., Richards, F. M. & Steitz, T. (1985) *J. Mol. Biol.* **186**, 645-649.
3. Lohmann, T. M. (1986) *CRC Crit. Rev. Biochem.* **19**, 191-245.
4. Zimmer, C. & Wahnert, U. (1986) *Prog. Biophys. Mol. Biol.* **47**, 31-112.
5. Warwicker, J. & Watson, H. C. (1982) *J. Mol. Biol.* **157**, 671-679.
6. Klapper, I., Hagstrom, R., Fine, R., Sharp, K. & Honig, B. (1986) *Proteins* **1**, 47-59.
7. Manning, G. S. (1978) *Q. Rev. Biophys.* **2**, 179-246.
8. Anderson, C. F. & Record, M. T., Jr. (1982) *Ann. Rev. Phys. Chem.* **33**, 191-222.
9. Klein, B. K., Anderson, C. F. & Record, M. T., Jr. (1981) *Biopolymers* **20**, 2263-2280.
10. Zimm, B. H. & LeBret, M. (1983) *J. Biomol. Struct. Dynam.* **1**, 461-471.
11. Gueron, M. & Weisbuch, G. (1980) *Biopolymers* **19**, 353-382.
12. Record, M. T., Jr., Anderson, C. F. & Lohman, T. M. (1978) *Quart. Rev. Biophys.* **2**, 103-178.
13. Rice, S. A. & Nagasawa, M. (1961) *Polyelectrolyte Solutions*, Academic Press, New York, chap. 7.
14. Fixman, M. (1979) *J. Chem. Phys.* **70**, 4995-5005.
15. Soumpasis, D. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5116-5120.
16. Mills, P., Anderson, C. F. & Record, M. T., Jr. (1985) *J. Phys. Chem.* **89**, 3984-3994.
17. LeBret, M. & Zimm, B. H. (1984) *Biopolymers* **23**, 271-285.
18. Murthy, C. S., Bacquet, R. J. & Rossky, P. J. (1985) *J. Phys. Chem.* **89**, 701-710.
19. Bacquet, R. & Rossky, P. J. (1984) *J. Phys. Chem.* **88**, 2660-2669.
20. Zakrzewska, K., Lavery, R. & Pullman, B. (1983) *Biomol. Stereodynam.* **1**, 163-183.
21. Weiner, P. K., Langridge, R., Blaney, J. M., Schaefer, R. & Kollman, P. A. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 3754-3758.
22. Klein, B. J. & Pack, G. R. (1983) *Biopolymers* **22**, 2331-2352.
23. Pack, G. R. & Klein, B. J. (1984) *Biopolymers* **23**, 2801-2823.
24. Troll, M., Roitman, D., Conrad, J. & Zimm, B. H. (1986) *Macromolecules* **19**, 1186-1194.

25. Matthew, J. B. & Richards, F. M. (1984) *Biopolymers* **23**, 2743–2759.
26. Wensel, T. G., Meares, C. F., Vlachy, V. & Matthew, J. B. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 3267–3271.
27. Gilson, M., Rashin, A., Fine, R. & Honig, B. (1985) *J. Mol. Biol.* **183**, 503–516.
28. Levitt, M. (1982) *Cold Spring Harbor Symp. Quant. Biol.* **47**, 251–275.
29. Tidor, B., Irikura, K., Brooks, B. & Karplus, M. (1983) *J. Biomol. Struct. Dynam.* **1**, 231–252.
30. Prabhakaran, M. & Harvey, S. (1985) *J. Phys. Chem.* **89**, 5767–5769.
31. Rao, S. & Kollman, P. (1985) *J. Am. Chem. Soc.* **107**, 1611–1617.
32. Gilson, M. & Honig, B. (1988) *Proteins* **3**, 32–52.
33. Seibel, G., Singh, U. & Kollman, P. (1986) *Proc. Natl. Acad. Sci. USA* **82**, 6537–6540.
34. Neumann, M. (1985) *J. Chem. Phys.* **82**, 5663–5672.
35. Gilson, M. & Honig, B. (1987) *Nature* **330**, 84–86.
36. Gilson, M., Sharp, K. & Honig, B. (1988) *J. Comp. Chem.*, **9**, 327–335.
37. Arnott, S. & Hukins, D. W. L. (1972) *Biochem. Biophys. Res. Commun.* **47**, 1504–1509.
38. Marky, L. A. & Breslauer, K. J. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 4359–4363.
39. Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., Jr. & Weiner, P. (1984) *J. Am. Chem. Soc.* **106**, 765–784.
40. Burden, R. L. & Faires, J. D. (1985) *Numerical Analysis*, Prindle, Weber & Schmidt, Boston.
41. Seeman, N. C., Rosenberg, J. M. & Rich, A. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 804–808.
42. Bleam, M. L., Anderson, C. F. & Record, M. T., Jr. (1983) *Biochemistry* **22**, 5418–5425.
43. Gilson, M. & Honig, B. (1988) *Proteins*, **3**, 32–52.

Received April 7, 1988

Accepted August 18, 1988