The free energy difference between simple models of B- and Z-DNA: Computer simulation and theoretical predictions

J. C. Gil Montoro and J. L. F. Abascal
Departamento de Química-Física, Facultad de Ciencias Químicas, Universidad Complutense de Madrid, E-28040 Madrid, Spain

(Received 13 August 1996; accepted 5 February 1997)

A method recently proposed to calculate by computer simulation the relative free energy between two conformational states of a polyelectrolyte is used for the case of the salt induced B- to Z-DNA transition. In this method, the calculation of the free energy may be split in two steps, one corresponding to the setup of the *uncharged* conformer in solution while the other one is the *charging* process of such a structure. Following the description of the method, simulations are reported to compute the free energy difference between the above mentioned DNA conformers in presence of monovalent added salt. We use a simple DNA solution model—the DNA is represented by charged spheres at the canonical positions of the phosphate groups, water by a dielectric continuum of appropriate permittivity and counterions and coions are modeled as soft spheres of equal ionic radius—for which theoretical approximations have been proposed. It is seen that the *charging* term is much more important than the *setup* contribution at any of the investigated salt concentrations. The variation of the free energy of each conformer as a function of the added NaCl concentration has been calculated. Both the B and Z conformers increase noticeably their stabilities with higher salt concentrations but the effect is more pronounced for the latter. As a consequence, the relative population of B-DNA, which is clearly prevalent at moderate ionic strengths, decreases with the addition of salt. However, up to 4.3 M NaCl a B→Z transition is not predicted for this DNA solution model. Additionally, the theoretical calculations are checked for the first time against computer simulation results. In particular, we have tried to assess the foundations and predictive ability of (especially) the Soumpasis potential of mean force theory and, in a lesser extent, the counterion condensation theory of Manning and the polymer reference interaction site model theory of Hirata and Levy. [S0021-9606(97)51218-9]

I. INTRODUCTION

The conformational change between the B and Z forms of DNA was discovered in 1972 by Pohl and Jovin.[1] Although it is not a transition in the usual sense in physics, it is frequently referred to in molecular biology and biophysics as the B- to Z-DNA transition; in this work, we will respect such terminology. The transition involves dramatic interconversions at the molecular level. The double-helix twist has opposite senses in both forms—right-handed in B-DNA and left-handed in Z-DNA—and the base pairs are flipped over in one form with respect to the other relative to the sugar-phosphate backbone.[2] In solution, DNA is a negatively charged polyelectrolyte due to the complete ionization of the acidic phosphate groups.[3] As Z-DNA is thinner than B-DNA, its charged phosphates are closer to each other giving stronger repulsions among them so that B is, in principle, the most stable DNA form. The transition from B- to Z-DNA then requires a stabilization of the Z with respect to the B form by means of extrinsic effects. Among these, most of the experimental studies have been focused on the supercoiling[4,5] and salt effects[1,6] due to its biological relevance. It is well established that the Z form of purine-pyrimidine inserts in supercoiled plasmids (relatively short circular DNA molecules) is induced by increasing levels of topological stress (i.e., negative supercoiling).[5,7] On the other hand, the salt-induced transition is not as well understood. In particular, chemical nature, size and charge are ion attributes which modify its ability to induce the transition for a given DNA sequence.

Regarding this salt-induced conformational change, there is a wealth of experimental data available (see Ref. 8 for a general review and Ref. 9 for an extension of the data there provided). Among the well established qualitative aspects it is known that the DNA base sequence determines the salt concentration at which the transition between the B and Z forms takes place. As the conformational change is an equilibrium between two species with different properties, it is easy to determine the experimental point at which both forms are equally probable, the so-called transition midpoint. Depending on the DNA sequence, midpoint concentrations of added NaCl ranging from 0.7 M to 5.4 M have been reported.[8,10,11] Besides, an increase of the counterion (cation) charge results in a decrease of the salt concentration required to induce the transition.[1] Unfortunately, most workers in the area have not reported the variation of the free energy difference of the transition $\Delta G^{B\rightarrow Z}$ with salt concentration. In fact, the detailed variation of the relative stability of both DNA forms has only been reported for $d(C\text{-}G)_n \cdot d(C\text{-}G)_n$ oligomers in the high monovalent salt region (i.e., 2–5 M NaCl) by Pohl.[6] Thus, there is a need for more experimental data (variation of the free energy difference over a broad range of salt and for both oligomeric and polymeric DNA) in

order to better test the validity of the model and approximations made in both theoretical and computer simulation approaches. Pohl observed that the free energy per base pair, obtained using an all-or-none formalism, was linear in the logarithm of salt concentration and did not depend on the DNA length. Pohl accurately fitted his data to the expression

$$\beta \Delta G_{\mathrm{exp}}^{B\rightarrow Z} = -0.3 \log \frac{C}{2.25}, \tag{1}$$

where the free energy is given per DNA phosphate, C is the NaCl molar concentration, 2.25 M is the transition midpoint and $\beta^{-1}=k_B T$, k_B being the Boltzmann constant and T the absolute temperature.

Due to the tremendous complexity of the problem, few theories have been applied to the relative stability of the B and Z forms of DNA at high concentrations of added salt. Besides, they have been forced to use simplified models for the system components (polyion, water and ions) and the interactions between them. In particular, water is always modeled at the McMillan-Mayer level, i.e., its effect is included as a modification of the electrostatic interaction between charges through an effective dielectric constant, while B- and Z-DNA are modeled as simple representations of their canonical forms. One of the first polyelectrolyte theories available, the Poisson-Boltzmann (PB) equation,[12,13] was applied to study the B- to Z-DNA transition by using the homogeneously charged hard cylinder model of polyelectrolytes.[14–16] The results of this two-parameter theory—polyion radius and charge density—are strongly dependent on the cylinder radii. Besides, the free energy difference versus salt concentration curve is almost flat; its slope is about 50 times smaller than the experimental one. As other PB predictions for this polyelectrolyte model in presence of monovalent salt are satisfactory,[17,18] these results are expected to represent the correct behavior for the B→Z transition of the homogeneously charged model. It seems then unreasonable to use more complex theories with such simple DNA models.

A more sophisticated model was solved also within a PB scheme by Demaret and Guéron.[19] There, DNA was not modeled as a cylinder, but its grooved structure was introduced in the calculation in an averaged way, keeping the cylindrical symmetry. Proceeding this way no effective DNA radii are needed, the space available to the ions within the polyion being used instead. Although several aspects of the calculation—such as the location of the DNA charges and the inclusion of the low value of the dielectric constant inside the DNA—somewhat hinders the appraisal of the contribution of each part of the model to its relative success, the predicted high-salt slope is fairly satisfactory: almost half the experimental value. Nevertheless, yet another PB calculation of a toy representation of DNA as a planar charged surface is able to reproduce the experimental slope.[20] In this sense, one has to be careful as to what extent the apparent success of a model might be fortuitous. Recently, the electrostatic contribution to the transition free energy has been calculated using the non-linear Poisson-Boltzmann (NLPB) theory for *all-atom* DNA models with a two-dielectric constant (one for the DNA and another one for the solvent) model.[21] The predicted slope is about half the experimental one. This is a quite good result but it shows either that other contributions (solvent effects, intramolecular interactions, etc.) are important or that the NLPB theory is insufficient.

Counterion Condensation (CC) theory[22–24] has been extensively applied to polyelectrolyte systems due to its simplicity and predictive ability. The original formulation of the theory, which in essence considers the electrostatic free energy as a pairwise sum of Debye-Hückel potentials reduced by condensed counterions around the polyion charges located along an infinite line, erroneously predicts that the Z form is always more stable than the B form because of its lower axial charge density. Being the CC theory an exact double limiting law (infinite salt dilution and infinite polyion length[25]), it is fairly accurate at moderate salt concentrations[26] but its application to very high salt conditions (such as those present when the transition takes place) may not be justified. Recently, the application of the CC framework to more realistic models than the infinite line of charges has been reported[25] and the B to Z transition reexamined in this context.[27] In these papers, the CC equations are numerically evaluated for models in which the double-helical array of phosphates are explicitly considered. This approach, hereafter referred to as the discrete CC theory (for the constant-dielectric solvent model and polymeric DNA), predicts that the B form is more stable than the Z form at moderate monovalent salt concentrations, in qualitative accord with experimental data. The same holds for the progressive diminution of the free energy difference with increasing added monovalent salt concentration in the moderate to high salt end. Nevertheless, the slope of this variation is an order of magnitude *smaller* than the experimental one.

Another interesting calculation within a well founded approximation was that performed by Hirata and Levy with the polymer reference interaction site model (PRISM) theory.[28,29] They used a DNA model in which the phosphate groups were explicitly modeled as solution anions placed at the phosphates canonical positions, neglecting completely the sugar and base atoms. In this paper, such model will be referred to as the *empty* DNA model. Again, the theory predicts a stabilization of the Z form with increasing salt, giving a zero value of the free energy difference at about 3.6 M NaCl. However, the slope of the free energy curve is an order of magnitude *larger* than the experimental result.

The most successful theory for the B→Z transition up till now is the Soumpasis potential of mean force (SPMF) approach.[30] The theory is also based on the empty DNA model. The many body correlations are decoupled by using the Kirkwood superposition approximation (KSA) (Ref. 31) and the ionic interactions replaced by the potentials of mean force (PMF) between the anions in a simple electrolyte solution at the same concentration as the added salt in the DNA system. The anion-anion PMFs are evaluated by means of integral equation theories, in particular, the hypernetted chain (HNC) theory and the exponential mean spherical approximation. Using the restricted primitive model (RPM) of electrolytes, in which the ions are considered as charged hard

spheres of equal diameter in a continuum medium of dielectric constant equal to that of the solvent, Soumpasis was able to reproduce quite accurately the full experimental $\Delta G^{B\to Z}$ curve by fitting the only adjustable parameter of the theory, namely, the ionic diameter σ, which was set to 4.9 Å.[30] Proceeding this way, Soumpasis considers that the free energy difference between the Z and B DNA conformers is only due to diffuse ionic cloud contributions, neglecting other effects such as the dependence on the DNA sequence, the intramolecular and vibrational contributions and those arising from the molecular nature of water. Although García and Soumpasis[32] have estimated the value of the intramolecular and vibrational free energies to almost cancel each other, adding up a mere $-0.1\ k_BT$ per phosphate for the alternating guanine-citosine sequence, some caution is needed. The hydration contribution may also favor the Z conformation of DNA at high salt (i.e., low water activity) because of the more ''economical'' or efficient hydration of this form with respect to the B form.[33]

The last paragraph brings out the important question of the theoretical limits due to the insufficiencies of the models. Schematically, the contributions neglected in simplified models fall into two categories, namely, those independent of added salt concentration and those dependent on it. The former leads to a constant contribution to the free energy of the conformers, and thus, they modify the precise transition midpoint while preserving the slope of the relative free energy curve as a function of salt concentration. It remains the question of which contributions are dependent on the salt concentration and which are not. Now, from a theoretical point of view, one is forced to speculate. It is likely that several effects (especially, the intramolecular and vibrational contributions) are not essentially salt concentration dependent. In the case of the DNA sequence, its influence on the transition slope is difficult to ascertain. Regarding the hydration contribution, it is expected to have some influence but this would clearly be a refinement over the continuum (implicit) solvent models. Fortunately, simulation results on a given model enable one to check (without ambiguity) the validity of the theoretical framework irrespective of the underlying model limitations. Besides, it is our belief that, at the current stage of development, and because of the difficulties mentioned, the assessment of the model predictions should be mostly concerned with the question of the slope rather than with the prediction of transition midpoints. Of course, this does not invalidate the use of any theory—in particular, the SPMF approach—as a semiempirical approximation which can be fitted to provide excellent predictions of the transition salt concentration.

So, what can be learned from previous theoretical work on the B to Z transition? The only thing more or less clear is that the homogeneously charged model seems somewhat unable to predict the experimental trends at the salt concentration range investigated in this work. Beyond that, very little is known. Some theories applied to the empty DNA model give promising predictions of the transition midpoint though different theories provide completely divergent results for the B-Z free energy slope. That from the discrete CC theory as applied to (roughly) the empty model[27] is an order of magnitude smaller than the SPMF result. The latter compares favorably with the experimental slope. On the other hand, the experimental slope is an order of magnitude smaller than the PRISM prediction. Thus, there is a clear need of an ''exact'' result which can only be provided by computer simulation results on the empty model. In this way, it should be possible to separate two different questions: the ability of a chosen model to account for the main features of the transition and the validity of the theoretical assumptions for that model. This typical use of computer simulation to close the triangle whose other vertices are theory and experiment has here an important drawback, namely, none of the methods already proposed to compute the B-Z free energy difference is applicable to our problem. So the first task is to develop a thermodynamic pathway which allows the computation of the free energy of a polyion immersed in a simple electrolyte solution from a reference state.

Recently,[34] we have reported a method to calculate, by computer simulation, the free energy difference in conformational changes of polyelectrolytes. The procedure uses a common reference state for both DNA conformers consisting in a collection of uncharged spheres immersed in a electrolyte solution. In a subsequent step, the spheres are taken to the appropriate positions to build the uncharged B and Z forms which, in a final stage, are charged up to give the DNA polyelectrolyte. In this way, the contributions can be ''exactly'' evaluated by computer simulation. This is especially true for the major contribution coming from the charging process of the neutral forms. In Ref. 34 we presented only terse ideas of the method and its preliminary implementation to the calculation of the relative free energy between the B and Z forms of DNA. The purpose of this paper is to report a deeper discussion of the theoretical scheme and the details of its application to the empty DNA model. On the other hand, we compare the results at different added monovalent salt concentrations with both the predictions of the various theories and the experimental data.

In Section II we discuss the methodology which will enable the evaluation of the free energy difference by computer simulation. The model used for B- and Z-DNA together with the details of the simulations are described in Section III. In Section IV we present the results and compare them with the theoretical approximations. The main conclusions are summarized in Section V.

II. THE THERMODYNAMIC PATH FOR THE COMPUTATION OF THE FREE ENERGY DIFFERENCE

For the computation of the free energy difference between two states, a reversible path linking them may be established; the free energy change along that path is the desired difference. If both states are *very similar* the free energy change can be calculated via a free energy perturbation (FEP) scheme[35] by means of the expression

$$\beta\Delta G_{\mathrm{FEP}} = -\log\langle \exp(-\beta\Delta U)\rangle, \tag{2}$$

where ΔU is the difference between the potential energy describing both states. The average, denoted by angle brackets in Eq. (2) is taken over the equilibrium ensemble for any of the states as long as they are supposed to be very close. If the states are similar but not quite so, a number of closely spaced intermediate states can be constructed and the total free energy evaluated as a sum over them. The FEP approach has been employed by Kollman's group in the estimation of the change in $\Delta G^{B\to Z}$ caused by certain modifications of the DNA sequence.[36–38] However, the actual free energy change of the B- to Z-DNA transition cannot be computed by this method due to the dissimilarity of the conformers which results in insurmountable sampling problems. Even with an extremely simplified DNA model, a direct estimation of $\Delta G^{B\to Z}$ through the FEP method would require the simulation of a huge number of intermediate models, which is exceedingly expensive in terms of computer time.

An alternative route might be provided by the Widom method of estimating the chemical potential.[39] The direct insertion of the full DNA molecule into the electrolyte would conceptually give the desired quantity but is, of course, out of the question. Other insertion methods for the estimation of the chemical potential of chain molecules such as the gradual insertion scheme,[40] the Rosenbluth sampling[41,42] and the recursive sampling of Grassberger[43,44] attempt to explore the available conformational space and, thus, cannot be applied here because of the fixed geometry of each of the DNA forms. Besides, the application of insertion methods to charged particles must solve the problems associated with the preservation of the box electroneutrality which makes them complex and rather specific.[45,46]

The SPMF theory uses another approach. The method can be visualized as if a number of anions from a simple homogeneous electrolyte, at the concentration under consideration, were moved and kept fixed at the positions of the DNA phosphates which builds up the DNA model. The free energy of this change is $\Delta G^{\Omega}_{\mathrm{SPMF}}$—the excess free energy of the corresponding Ω-DNA form ($\Omega=B,Z$) *with respect to the electrolyte*— from which $\Delta G^{B\to Z}_{\mathrm{SPMF}}=\Delta G^{Z}_{\mathrm{SPMF}}-\Delta G^{B}_{\mathrm{SPMF}}$ is readily obtained. The SPMF theory simplifies the many body problem by decoupling the contribution of the second and successive phosphates to the chemical potential of a given one by using the Kirkwood superposition approximation together with the HNC correlation functions. To assume uncorrelated the ionic distribution around consecutive charged phosphates despite the long-range interactions between them and the mobile ions is a weak point of the SPMF theory. The use of simulation results instead of HCN calculations within this scheme would not overcome the problem. As the HNC correlation functions and PMFs compare very well with simulation for monovalent ions at the high salt concentration range relevant to this study, the transition free energy obtained using simulation results within Soumpasis' scheme would be essentially the same as that of the original SPMF theory.

The situation may be different if the building up of the charged structures from the solution anions is split in several steps. The key idea is to first set up an uncharged Ω-DNA form by moving uncharged groups from the electrolyte solution to their canonical DNA positions ($\Delta G^{\Omega}_{\mathrm{setup}}$) which is followed by a further step corresponding to the charging process of the DNA structures ($\Delta G^{\Omega}_{\mathrm{charging}}$). The advantage of this method, which we call *setup and charge* (SUCH) is that the leading contribution to the free energy would then come from the $\Delta G^{\Omega}_{\mathrm{charging}}$ term, i.e., from the electrostatic interactions at the polyion charged groups. Interestingly, this term can be ''exactly'' calculated by computer simulation through a charging process starting at the uncharged DNA. Besides, the correlations between uncharged particles immersed in a medium of high ionic strength are small provided they are separated by a distance larger than their repulsive radii (the statement has proven to be virtually exact in uncharged media[47]). As the minimum interphosphate distance in B- or Z-DNA is about one and a half times the value of the effective ionic diameter (typically 4.2 Å), $\Delta G^{\Omega}_{\mathrm{setup}}$ for the empty model must be small as well. Moreover, in this case, three body (and higher) contributions to this term are expected to be negligible, i.e., the KSA is then virtually exact. Thus, by separating the contributions to the free energy in two components—that coming from the setup of uncharged DNA structures from neutral groups and that of the charged DNA relative to the neutral one—we are able to evaluate the leading one (a purely electrostatic term) as accurately as needed while the minor contribution can be treated by means of a superposition approximation whose application is now accurate enough.

The thermodynamic path of our route for the computation of the free energy difference is shown in Fig. 1, where the stages of the process are depicted. The SPMF route is also drawn as a dashed line. Starting with the homogeneous simple electrolyte at the desired concentration (stage 1), a number of ions are uncharged (stage 2). This step is not necessary if one is interested exclusively in the transition free energy difference because the ''uncharging'' term is the same for both DNA conformers and, thus, it gives no net contribution to the free energy of the transition. As a matter of fact, we will forget about this term in most of the paper. Next, the uncharged phosphates are placed at their positions in the corresponding DNA form (stage 3). This is the *setup* step. The last step (stage 4) is the *charging* process, where the phosphates are charged up back to its full charge, giving the desired complete DNA form in the salt solution. Given this pathway, the excess free energy of each DNA form with respect to the starting electrolyte solution is

$$\Delta G^{\Omega}=\Delta G^{\Omega}_{\mathrm{uncharging}}+\Delta G^{\Omega}_{\mathrm{setup}}+\Delta G^{\Omega}_{\mathrm{charging}}, \tag{3}$$

and, thus, the ''total'' free energy difference between the DNA conformers is

$$\Delta G^{B\to Z}=\Delta G^{B}_{\mathrm{setup}}+\Delta G^{B}_{\mathrm{charging}}-\Delta G^{Z}_{\mathrm{setup}}-\Delta G^{Z}_{\mathrm{charging}}, \tag{4}$$

where the subindices *uncharging*, *setup* and *charging* refer to the corresponding steps. We have quoted the word ''total'' to emphasize that some contributions (hydration, intramolecular interactions, etc.) have not been considered in the method. Although the latter contributions could be incorpo-

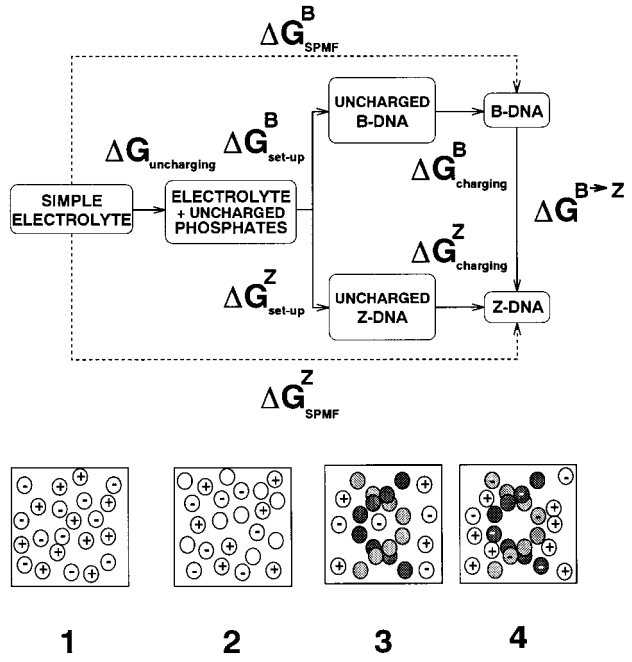

FIG. 1. The thermodynamic path for the computation of the free energy difference (see the text for details). Top: the free energy differences involved in each of the steps along the thermodynamic path. Bottom: schematic representation of the actual system in every stage. From left to right, starting with a simple electrolyte solution, some particles are uncharged; then, the uncharged particles are placed at the positions of a DNA form, and finally, the uncharged DNA structure is charged-up.

rated into the formalism, we prefer to limit the method to the terms appearing in almost any model, even in the simplest ones. Notice that for some models (as the one used in this paper) the above expression is exact. For the sake of generality, in the remainder of this section, we deal with a polyion which presents two different conformers (called B and Z) in solution. Both conformers are modeled with the same number of sites, N_s, from which N_q are charged. Besides, the free energy will be divided by a number N, which is related to the number of repeating units that build the polyion. Usual choices for DNA systems are the number of phosphates (as in this paper) or the number of base pairs.

The setup contribution could be computed via the free energy perturbation method introduced above. As the polyelectrolyte groups are discharged in this case, the correlations among them are much weaker than for charged sites and the required number of intermediate steps of the FEP process should be much smaller since there is no ionic condensation around the uncharged macromolecules. Nevertheless, such calculation seems a waste of computing resources as long as approximate methods should be accurate enough in these conditions. In fact, it has been recently shown that the KSA predicts good results in inhomogeneous uncharged media.[48] Besides, Zhou and Stell[47] have shown that the equation of state of a n-particle linear chain fluid may be derived from an approximation closely related to the superposition approximation. In addition, they use the same general framework to investigate solvent effects on ionic association. Thus, if one writes the setup free energy in terms of the multiparticle PMF among the uncharged particles, which depends on the set of positions of those particles in the Ω form

$$\Delta G^{\Omega}_{\text{setup}} = \frac{1}{N} W^{(N_s)}_{\text{uncharged}}(\mathbf{r}^{\Omega}_1, \ldots, \mathbf{r}^{\Omega}_{N_s}), \tag{5}$$

it is then expected that the multiparticle PMF may be safely expressed in terms of the PMFs between pairs (the KSA approximation)

$$W^{(N_s)}_{\text{uncharged}}(\mathbf{r}^{\Omega}_1, \ldots, \mathbf{r}^{\Omega}_{N_s}) = \frac{1}{2}\sum_{i=1}^{N_s}\sum_{j=1}^{N_s} W^{(2)}_{\text{uncharged}}(r^{\Omega}_{ij}), \tag{6}$$

so that

$$\Delta G^{\Omega}_{\text{setup}} = \frac{1}{2N}\sum_{i=1}^{N_s}\sum_{j=1}^{N_s} W^{(2)}_{\text{uncharged}}(r^{\Omega}_{ij}). \tag{7}$$

In the above expressions, r^{Ω}_{ij} is the distance between the i and j sites and $W^{(2)}_{\text{uncharged}}$ is the PMF between the i and j uncharged particles dissolved in the homogeneous electrolyte solution which, in general, is given by[46]

$$W^{(2)}_{\text{uncharged}}(r_{ij}) = \mu_{\text{uncharged}}(r_{ij}) - \mu_{\text{uncharged}}(\infty), \tag{8}$$

where $\mu_{\text{uncharged}}(\infty)$ is the chemical potential of the i uncharged particle in a bulk electrolyte (i.e., with the j particle infinitely apart) and $\mu_{\text{uncharged}}(r_{ij})$ is the same quantity when the j site is present at a distance r_{ij}.

As for the calculation of the charging term, Verwey and Overbeek[49] introduced two charging routes appropriate for calculating electrostatic free energies of polyelectrolytes. In the first one, the polyelectrolyte is charged up in an isothermal process, the ionic atmosphere being in equilibrium at all stages; in the second, the energetic cost of assembling the ionic atmosphere around the already charged polyion is computed. There is a parallelism between these routes and the Debye and Guntelberg charging processes of simple electrolytes, respectively, although in the latter the central ion is the entity charged against the field produced by its ionic atmosphere.[50] Nevertheless, the Guntelberg approach is unphysical in the canonical ensemble, where the number of ions is fixed, because electroneutrality is violated during the process.[15] Thus, we will use a Debye-like charging process where the polyelectrolyte sites are charged up while a number of counterions exactly compensate its charge at each stage. The total charging free energy is then given by the usual integral,

$$\Delta G^{\Omega}_{\text{charging}} = \frac{1}{N}\int_0^1 \sum_{i=1}^{N_q} q_i \varphi^{\Omega}_i(\lambda)\, d\lambda, \tag{9}$$

where the sum extends over the polyion sites i with charges q_i and $\varphi^{\Omega_i}(\lambda)$ is the total electrostatic potential at those points in the Ω-DNA form[51] in the charging state λ. The above expression was incorrectly written in Ref. 34 and the preliminary results there presented differ slightly with respect to those given in this paper.

TABLE I. Parameters of the B- and Z-DNA helices.

	B-DNA	Z-DNA
Helix sense	Right-handed	Left-handed
Repeating unit	Base pair	Base pair dimer
Repeating units per turn	10 (20 phosphates)	6 (24 phosphates)
Rise per repeating unit (Δz^{Ω}) (Å)	3.38	7.43
Rotation per repeating unit ($\Delta \phi^{\Omega}$) (degrees)	+36	−60
Helix pitch (Å)	33.80	44.58
Reduced axial charge density ξ	4.23	3.85

TABLE II. Generating phosphorus coordinates parameters for each DNA form.

DNA form	s[a]	p[b]	ρ_0^p (Å)	$\phi_0^{s,p}$ (degrees)	$z_0^{s,p}$ (Å)
B	1	1	8.91	0.0	0.00
	2	1	8.91	154.4	0.78
Z	1	1	7.31	0.0	0.00
	2	1	7.31	164.0	0.03
	1	2	6.27	115.4	1.70
	2	2	6.27	348.6	5.76

[a]Strand.
[b]Phosphate type.

III. IMPLEMENTATION FOR "EMPTY" B- AND Z-DNA MODELS

A. Models and potential parameters

As mentioned in the Introduction, we deal with empty B- and Z-DNA models, i.e., only the phosphate groups are explicitly considered. There are two canonical Z-DNAs, called Z_I and Z_{II},[52] Z_I being more stable at the high salt concentrations relevant for the salt-induced transition;[53,54] in this work we always refer to the Z_I form as Z-DNA. As is well known, both B- and Z-DNA consist of two strands ($s=1,2$). All the phosphates are equivalent in the B form but the phosphates along each strand of the Z form are alternatively located at two different distances from the helix axis, thus being two non-equivalent phosphate types. We will refer to these phosphates as Z^1 and Z^2. The alternate residue conformation is manifested as a zig-zag course of the sugar-phosphate backbone which is the origin of the Z-DNA name.[2] As a consequence, the repeating unit in Z-DNA is a base pair dimer instead of a single base pair as in B-DNA (see Table I where some relevant structural parameters of both DNA forms are listed). The phosphate coordinates for the B and Z_I forms have been taken from the phosphorus atoms positions reported by Arnott and Hukins[55] and Wang *et al.*,[52] respectively. For a complete DNA turn, they may be written down in a convenient compact form in a cylindrical coordinate system as

$$\begin{aligned} \rho_i^{\Omega,p} &= \rho_0^{\Omega,p}, \\ \phi_i^{\Omega,s,p} &= \phi_0^{\Omega,s,p} + \Delta\phi^{\Omega} \cdot i, \\ z_i^{\Omega,s,p} &= z_0^{\Omega,s,p} + \Delta z^{\Omega} \cdot i, \end{aligned} \tag{10}$$

where $s=1,2$ represents each of the strands, $p=1$ and i ranges from 0 to 9 for B-DNA, and $p=1,2$ and $i=0$ to 5 for Z-DNA. In the above expression Δz and $\Delta\phi$ are helical parameters, the rise and rotation per repeating unit, respectively, (see Table I), and $\rho_0^{\Omega,p}$, $\phi_0^{\Omega,s,p}$, and $z_0^{\Omega,s,p}$ are the coordinates of the minimum set of phosphorus atoms which enables to generate the positions of the remainder ones (Table II). From these parameters it is possible to calculate the reduced axial charge density of each DNA conformer. It is defined as

$$\xi = \frac{e^2\beta}{4\pi\epsilon\epsilon_0 b}, \tag{11}$$

where ϵ is the (relative) dielectric constant of the pure solvent (78.358 for water at 25 °C, the temperature of our study), ϵ_0 is the vacuum permittivity, e the magnitude of the electronic charge, and b the separation between charges along the polyelectrolyte axis, i.e., Δz^{Ω} divided by the number of phosphate groups in a repeating unit which are 2 for B- and 4 for Z-DNA. The ξ values for B- and Z-DNA are 4.23 and 3.85, respectively.

The ion-ion (and phosphate-ion, the phosphates are just immobile anions) potential is of the form[18,56–59]

$$U_{ij}(r_{ij}) = \frac{A_M e^2 |z_i z_j|}{4\pi\epsilon_0 N_c n} (r_i + r_j)^{(n-1)} \frac{1}{r_{ij}^n} + \frac{z_i z_j e^2}{4\pi\epsilon\epsilon_0} \frac{1}{r_{ij}}, \tag{12}$$

where z_i is the valency of ion i ($\pm$ 1 in this study), r_{ij} the distance between ions, $A_M=1.7476$ the Madelung's constant of the NaCl structure, $N_c=6$ the coordination number of the same structure, $n=9$ the steepness parameter of the repulsive interaction, and r_i the nominal radius of ion i. The structure of the homogeneous electrolyte with $r_i=1.4214$ Å (irrespective of the sign of the ionic charge) roughly corresponds to that of a RPM with a hard-sphere diameter of 4.2 Å.[57]

B. Calculation of the free energy terms

To implement the method, the terms in Eq. (4) must be evaluated through the relations (7) and (9). In the DNA model defined in the previous subsection all the polyelectrolyte sites are coincident with the phosphate positions and carry one negative charge. In this way, if the free energy is referred to one phosphate group, $N=N_s=N_q$. As there are only p structurally distinct sites, the sum over j in Eq. (7) has N/p identical terms and the setup free energy per phosphate takes the form (for an infinite polyion)

$$\beta\Delta G_{\text{setup}}^{\Omega} = \frac{\beta}{2p} \sum_p \sum_i^{\infty} W_{\text{uncharged}}^{(2)}(r_{ip}^{\Omega}). \tag{13}$$

The sums extend over nonequivalent phosphates and polyion sites, respectively. Besides, r_{ip}^{Ω} is the distance of the i phosphate to the reference one of type p. Analogously, the sum over i in Eq. (9) also has N/p identical terms. As every site carries a charge $-e$, the charging free energy per phosphate for an infinite polyion is given by

$$\beta\Delta G^{\Omega}_{\text{charging}} = \int_0^1 \left(\frac{-1}{p}\right) \sum_p \phi_p^{\Omega}(\lambda) d\lambda, \tag{14}$$

where ϕ_p^{Ω} is the reduced electrostatic potential at the charged polyion site of type p, $\phi_p^{\Omega} = \varphi_p^{\Omega} \beta e$.

The setup PMFs between uncharged particles has been computed by Widom's method.[39] For each desired added salt concentration, the corresponding electrolyte solutions at the same salt concentrations (but with a neutral particle added) has been simulated. Ghost uncharged particle insertions are attempted at distances from the reference one covering the whole set of interparticle separations between the polyion phosphates in the B and Z forms. In this way, all the necessary $\mu_{\text{uncharged}}(r_{ij}^{\Omega})$ are readily obtained, the asymptotic behaviour at large distances providing $\mu_{\text{uncharged}}(\infty)$. By using Eq. (8), the PMF can be evaluated from which the setup term—Eq. (13)—is computed. Notice that, as $\mu_{\text{uncharged}}(\infty)$ appears in the contribution to each of the DNA conformers, it exactly cancels when calculating the transition free energy.

Both the intramolecular (due to other phosphates) and the ionic cloud contributions must be considered in the computation of the electrostatic potential at each phosphate[51]

$$\phi_p^{\Omega}(\lambda) = \frac{1}{4\pi\epsilon\epsilon_0}\left[\sum_{i\neq p}^{\infty} \frac{-\lambda e}{r_{ip}} + \int \frac{q^{\lambda}(\tau)}{r_{p\tau}} d\tau\right], \tag{15}$$

where $q^{\lambda}(\tau)$ is the net ionic charge at τ due to the ionic distribution around the Ω-DNA form in charging state λ; the integration takes place over all the space. Since the electrically neutral bulk electrolyte far from the polyion does not contribute to the potential at the polyion sites, the integration along the radial direction has a natural limit defined by the simulation cell boundary (or the internal region of the Modulated Bulk as a Fuzzy Boundary method, see below). As for the axial direction, the intramolecular and ionic cloud contributions tend to compensate so that its net effect vanishes at certain distance to the considered phosphate. We have used ten DNA turns above and below the central one (21 total turns) obtained replicating the one explicitly simulated (or the average when the simulation cell contains more than one helix turn). Computations considering a larger system in the axial direction gave identical results.

In the Debye-type charging process the overall electroneutrality of the system is required at all stages. As the explicit part of the system in the axial direction must contain an exact number of DNA turns (to preserve the helical symmetry), and as the total polyelectrolyte charge (in electron units) must be an integer number (identical to the number of counterions), only certain values for the fractional charge at the phosphate sites are allowed. Happily, the number of phosphates per turn in both B- and Z-DNA (20 and 24, see Table I) are highly divisible numbers. In particular, both are multiple of four, so that we can use λ values of 0, 0.25, 0.5, 0.75 and 1.0.

C. Simulation of the polyelectrolyte solutions

The experimentally reported NaCl midpoint for the B $\rightarrow$Z transition of different DNA sequences vary from 0.7 M for poly(dGm^{5}dC) (Ref. 10) to 5.4 M for poly(dAm^{5}dC $\cdot$dGdT).[11] Monte Carlo simulations have been performed for the B and Z forms of the empty model of DNA at bulk nominal salt concentrations of 0.5, 1.0, 2.5 and 4.3 M, approximately covering the experimental range and centered on that corresponding to the original poly(dGdC) which is 2.5 M.[1] For a given salt concentration, each value of the charging parameter implies a different polyelectrolyte solution simulation. The simulated states (24 in total) are given in Table III. The minimum number of λ values is 0, 0.5 and 1 for the system at 2.5 M. The other three salt concentrations also include $\lambda = 0.75$ and, additionally, at 1 M the value $\lambda = 0.25$ has been simulated. The simulations with λ lower than unity have λ times the number of DNA charge compensating counterions of the $\lambda = 1$ case (Table III). Two DNA turns are explicitly considered for the two lowest concentrations and just one for two more concentrated cases. Notice that, as the counterions condense on the polyion surface, it is not possible to exactly set the desired simulated bulk salt concentration, especially when $\lambda = 1$. Moreover, different values are obtained in the simulation of each of the DNA forms. The actual bulk salt concentrations are presented in Table III. Nevertheless, for simplicity, concentrations will often be referred by their nominal values.

TABLE III. The simulations (only the conditions for states with $\lambda = 1$ are given). For each DNA conformer, additional simulations with $\lambda = 0.5$ have been performed at all the salt concentrations. Other states with $\lambda = 0.25$ and 0.75 have been also simulated (24 simulations in total). States with λ lower than unity are similar to those at full charge but contain λ times the number of compensating DNA charge counterions.

DNA form	Nominal salt concentration (M)	Number of ions[a]	Box apotheme (Å)[b]	Bulk salt concentration (M)[c]
B	0.5	40+155	48.36	0.512
	1.0	40+195	39.20	0.988
	2.5	20+120	27.20	2.58
	4.3	20+200	26.63	4.34
Z	0.5	48+205	48.43	0.503
	1.0	48+256	39.11	0.972
	2.5	24+160	27.35	2.53
	4.3	24+265	26.69	4.31

[a]DNA sites / compensating cations + added salt ion pairs.
[b]The simulation box is a hexagonal prism. Its height is the pitch of the corresponding DNA form (Table I) times the number of helix turns in each case (two for the two lowest salt concentrations and just one for the other two).
[c]Estimated uncertainties affect the last figure.

The simulations of the polyelectrolyte solutions are somewhat complex. The discussion of the general method can be found elsewhere.[58–60] Next, we give a brief summary of the technical details and the particular conditions used in this work. The simulations have been performed in the canonical ensemble using the standard Metropolis algorithm. In the axial direction, standard periodic boundary conditions are used. An exact formula is used to calculate the interaction between a mobile ion and the arrays of DNA charges.[60] Regarding the radial direction, we use the Modulated Bulk as a Fuzzy Boundary method[58] in which the polyelectrolyte is surrounded by bulk electrolyte. In this way, a state of infinite

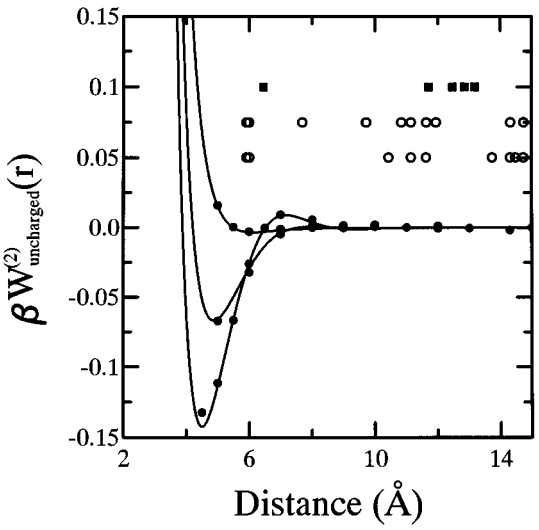

FIG. 2. The potential of mean force (in k_BT units) between two uncharged particles within an electrolyte solution as a function of the separation between them at three salt concentrations (from top to bottom 0.5, 2.5, and 4.3 M). Solid circles are computer simulation results (Widom insertion) and solid lines represent the HNC integral equation calculations. Solid squares at the top of the figure show the distances between different phosphorus atoms in B-DNA while hollow circles are the distances in Z-DNA (recall that there are two different phosphorus atoms in the Z form; the upper symbols stand for the inner atoms, Z^2, and, the others, for the outer ones, Z^1).

dilution for DNA is implicitly employed. As the salt concentration in the bulk is of the same order as that in the vicinity of the polyion, there is no need of special sampling schemes.[61] Typically, 15000 cycles (of as many single-particle moves as mobile ions) were used for the complete equilibration period and 20000 cycles were additionally run for the collection of structural and thermodynamic properties.

IV. RESULTS

A. The setup term

The setup term to the free energy is obtained from the sum of the PMFs between uncharged particles at the corresponding interphosphate distances in the polyelectrolyte, Eq. (13). The variation of $W^{(2)}_{\text{uncharged}}(r)$ with distance is shown in Fig. 2 for selected salt concentrations. Except for the small distances region (up to about 8 Å) the curves are essentially flat. To give a better idea of the contributions to the sum, the set of interphosphate distances for each of the DNA conformers is also plotted in the figure (notice that Z-DNA has two nonequivalent phosphates and, thus, one different set for each phosphate type). As the separations between phosphates mostly lie in the flat region of the curves, the PMFs and, hence, the contributions to the setup free energy are almost negligible but for the very few first neighbouring phosphates. A good estimate of three-body contributions to the multiparticle PMF of Eq. (5) can be obtained by having a look at the distance between the two closest neighbouring phosphates to a third one. This distance is 12.5 Å for B-DNA (the next and previous phosphates of the reference one in its same strand) and 9.7 Å for Z-DNA (the same situation as for B-DNA but when the reference phosphate is of type Z^2; this value for Z^1 phosphates is 10.4 Å). Even the direct PMF between those is negligible at such distances, so that its indirect combined effect over the reference phosphate must be almost meaningless. Thus, the KSA is quite acceptable for this particular application.

TABLE IV. Setup contribution to the free energy (HNC integral equation calculations) in k_BT units per phosphate.

Nominal salt concentration (M)	DNA form	$\beta\Delta G^{\Omega}_{\text{setup}}$	$\beta\Delta G^{B\to Z}_{\text{setup}}$
0.5	B	−0.003	−0.001
	Z	−0.004	
1.0	B	−0.008	−0.005
	Z	−0.013	
2.5	B	−0.014	−0.017
	Z	−0.031	
4.3	B	−0.002	−0.026
	Z	−0.028	

In Fig. 2, the predictions of the HNC theory are also depicted. This quantity cannot be obtained in usual HNC calculations as the system involved is an homogeneous electrolyte solution with uncharged particles at infinite dilution. This can be solved by treating a mixture of anions, cations (both species at molar fraction $0.5-x$) and uncharged particles (molar fraction $2x$). A vanishing x is required. Our experience shows that the results do not change provided that x is small enough, of the order of 10^{-5} or 10^{-6}. Once the radial distribution function between uncharged particles $g_{\text{uncharged}}(r)$ is obtained, the PMF is simply

$$\beta W^{(2)}_{\text{uncharged}}(r) = -\ln g_{\text{uncharged}}(r). \tag{16}$$

The differences between the simulation results by Widom's method and theoretical calculations are of the same order as the statistical uncertainty in the former. Since the error in the simulations are randomly distributed, they may add when subtracting the contributions for both DNA forms to compute the desired free energy balance. Contrarily, the integral equation errors are essentially systematic so it should mostly cancel in the calculation of the setup contribution to the transition free energy. Given the small difference between integral equation and simulation results, and considering the above arguments, the HNC calculations seem more reliable and will be used in the remainder of this paper. The terms $\Delta G^{\Omega}_{\text{setup}}$ have been computed at salt concentrations of 0.508, 0.980, 2.555 and 4.33 M, the averages of the measured bulk salt concentrations for the B and Z simulations at each nominal concentration (see Table III).

The results for the free energy (per phosphate) of the setup step are presented in Table IV. As anticipated, the setup free energies are always very small; in fact, they are negligible at the lower salt concentrations. The values for both DNA forms are always negative meaning that the building of the uncharged polyelectrolyte is favoured with respect to the state where the uncharged particles are free in the solution. As for the difference between the B and Z forms,

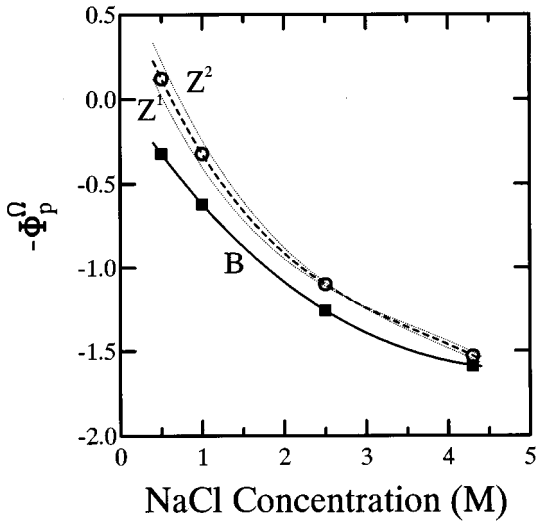

FIG. 3. The negative of the reduced electrostatic potential at the phosphates as a function of the added-salt concentration. The symbols represent simulation results (lines are a guide to the eye) for the fully charged polyelectrolyte solutions, i.e., with $\lambda=1$. For Z-DNA, the hollow circles are the mean value of both types of phosphorus atoms, Z^1 and Z^2, while the dotted lines show the electrostatic potential for each kind of site.

these are also negative (favoring the Z form, presumably due to its higher compactness), their absolute values greatly increasing with salt concentration.

B. The charging term

The evaluation of the charging term requires the knowledge of the reduced electrostatic potentials at the phosphate positions, ϕ_p^Ω. The variation of $-\phi_p^\Omega$ (notice that $\Delta G_{\text{charging}}$ depends on $-\phi$) with the added salt concentration is displayed in Fig. 3 for the fully charged ($\lambda=1$) DNA conformers. The intramolecular interaction of Eq. (15) is positive due to the repulsion between the charged phosphate groups. Besides it is a constant term, independent of salt concentration. The ionic cloud contribution is negative due to the accumulation of counterions near the phosphates. The latter term increases (in absolute value) with salt concentration because the counterions penetrate more efficiently into the DNA structure, screening the interphosphate repulsion. Due to the smaller distances between phosphates in the Z form, its intramolecular electrostatic energy is more positive than that of B-DNA. This effect dominates in the whole salt concentration range studied. As can be seen in Fig. 3, the screening is more effective for Z-DNA than for the B form resulting in a diminution of the difference between the potential at the phosphate groups of both DNA conformers with increasing salt concentration.

These facts may be explained in terms of the set of interphosphate distances displayed in Fig. 2. In general, the phosphates in Z-DNA are closer to each other than in the B form, but what affects distinctly the behaviour of each conformer is the existence in Z-DNA of several interphosphate separations in the range $7-11$ Å which are absent in B-DNA (see Fig. 2). These distances are of the order of two ionic diameters thus allowing the presence of a counterion between two phosphates which is considerably stabilized. As the probability for this to occur increases with the counterion concentration, the decay of the $-\phi_p^\Omega$ curve for Z-DNA is steeper than that of B-DNA. This is more clearly seen when the curves for each phosphate type of Z-DNA are drawn separately (Fig. 3). Notice that Z^1 phosphates, which have its neighbours at similar distances (but always a bit closer) than those of B-DNA lead to a electrostatic potential curve with a not too different slope. On the other hand, Z^2 sites have neighbouring phosphates at distances suitable for trimer formation (at 7.7 and 9.7 Å) and, thus, its $-\phi_p^\Omega$ curve is more affected by the counterion concentration.

The negative of the average reduced electrostatic potential at the phosphates (the integrand of the charging integral) for the λ values used in the charging integrals, the integrated $\beta\Delta G^\Omega_{\text{charging}}$, and the total $\beta\Delta G^{B\to Z}_{\text{charging}}$ can be found in Table V at the salt concentrations investigated. A representation of

TABLE V. The negative of the reduced electrostatic potential ($\phi_p^\Omega=\varphi_p^\Omega\beta e$) at the phosphates and charging free energy differences (in k_BT units per phosphate) obtained from Monte Carlo simulations.

Nominal salt concentration	DNA form	$\frac{-1}{p}\Sigma_p\phi_p(\lambda)$ $\lambda=0.25$	$\lambda=0.5$	$\lambda=0.75$	$\lambda=1$	$\beta\Delta G^\Omega_{\text{charging}}$	$\beta\Delta G^{B\to Z}_{\text{charging}}$
	B		−0.014	−0.068	−0.139	−0.036	
0.5		—					+0.227
	Z		0.238	0.255	0.231	0.191	
	B	−0.176	−0.318	−0.457	−0.620	−0.316(-0.315 [a])	
1.0							+0.159(+0.161 [a])
	Z	−0.100	−0.150	−0.217	−0.321	−0.157(−0.154 [a])	
	B		−0.692		−1.257	−0.671	
2.5		—		—			+0.077
	Z		−0.616		−1.101	−0.594	
	B		−0.920	−1.260	−1.587	−0.867	
4.3		—					+0.058
	Z		−0.840	−1.189	−1.530	−0.809	

[a]Values obtained integrating over the points $\lambda=0$, 0.5 and 1 only.

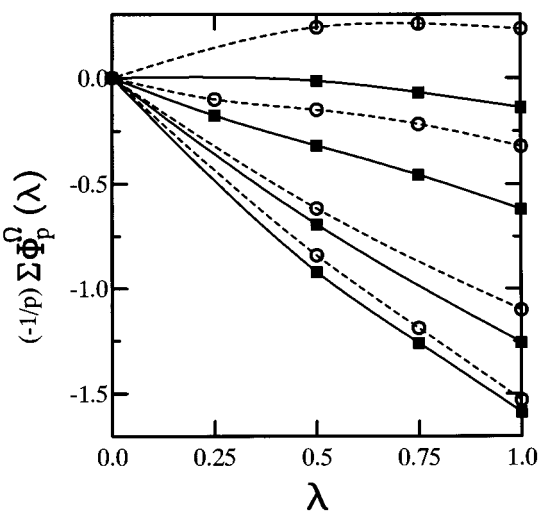

FIG. 4. The negative of the reduced electrostatic potential at the phosphates as a function of the charging parameter λ. The symbols are simulation results (filled squares for any B-DNA phosphate and hollow circles for the mean value of the Z^1 and Z^2 phosphates of Z-DNA) and the lines represent the actual functionality assumed by the integration method. The data for each Ω-DNA correspond, from top to bottom, to the nominal added salt concentrations 0.5, 1, 2.5, and 4.3 M, respectively.

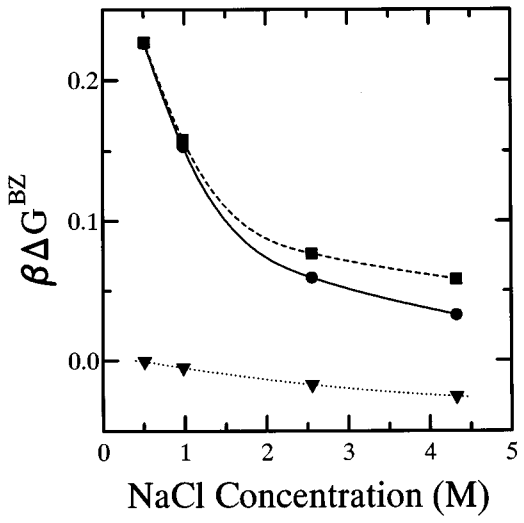

FIG. 5. Contribution (in k_BT units per phosphate) of the $\Delta G^{B\to Z}_{\text{setup}}$ (dotted line, inverted triangles) and $\Delta G^{B\to Z}_{\text{charging}}$ (dashed line, squares) terms to the ''total'' free energy difference $\Delta G^{B\to Z}$ (solid line, circles) as a function of the added monovalent salt concentration.

the dependence of the negative average electrostatic potential at the phosphates on λ is also given in Fig. 4 for both DNA forms and all salt concentrations considered. Notice that the potential (and thus the integrand) at the uncharged ($\lambda=0$) phosphates is zero due to the symmetry of the supporting electrolyte. The charging integral has been solved by Simpson's integration rule. To assess the influence of the number of points in the charging integral, we have evaluated it for the pair of systems at 1 M bulk salt concentration using both the five simulated λ points or only the three values at $\lambda=0$, 0.5 and 1. With $\Delta\lambda=0.25$ we obtain essentially the same value of the difference $\beta\Delta G^{B\to Z}_{\text{charging}}$ as with $\Delta\lambda=0.5$ (Table V). Analogous conclusions are reached at other salt concentrations for which the integrals include the points at $\lambda=0$, 0.5, 0.75 and 1. The success of Simpson's formula lies in the smooth shape of the curves of Fig. 4 so the error in each of the $\beta\Delta G^{\Omega}_{\text{charging}}$ is already quite small. Besides, we benefit from the cancellation of errors due to the computation of a free energy difference as the shape of the curves for both conformers is similar. The charging free energy for the B conformer is always negative, its absolute value strongly increasing with the salt concentration (Table V). For Z-DNA this term is positive at 0.5 M, which means that this form is unstable with respect to the uncharged structure due to the strong interphosphate repulsions. Nevertheless, its stabilization with increasing salt concentrations is even stronger than for the B form so that at high salt concentrations the free energy values are similar for both the conformers. Consequently, the charging step contribution to the ''total'' free energy difference is always positive, but decreases strongly as the added salt concentration increases. In other words, the charging term clearly favors the B form at moderate ionic strength but the differences are not significant in highly concentrated ionic solutions.

The setup, charging, and ''total'' free energy differences (in k_BT units per phosphate) between Z- and B-DNA are shown in Fig. 5 as a function of the added monovalent salt concentration. The results for the ''total'' free energy difference show that the B conformer is much more stable at low to moderately high concentrations of added salt. The slope of the charging and setup terms being negative, both contributions increase the relative stability of the Z form in more concentrated ionic solutions. The relative importance of each of the terms is manifested by the fact that from the change in the free energy difference in going from 0.5 M to 4.3 M (0.194 k_BT per phosphate), the charging term accounts for about 87% (0.169 k_BT); only the remainder 13% (0.025 k_BT) is due to the setup term (cf. Tables IV and V). This is even more glaring for each of the individual $\beta\Delta G^{\Omega}$ changes, where the setup term scarcely contributes a 2% to the ''total'' free energy change for Z-DNA in the salt concentration range considered and even less for the B form. As we are dealing with a conformational equilibrium, a small value of $\Delta G^{B\to Z}$ (negative or positive) simply indicates that the relative population of the corresponding form is marginally larger. The main effect of the setup term is to modify the salt concentration at which the population of both forms is the same, the transition midpoint. Even after the inclusion of the setup term, the empty model does not predict the transition, at least up to 4.3 M NaCl. It is likely that other contributions not included in the model (especially chemical bonding and hydration) are needed to reproduce the experimental data. However, the model investigated in this paper is a crude representation of DNA. Recently, we have shown the important effect of a proper account of the DNA shape on the ionic distribution around the polyion.[59] Our preliminary studies with a simple grooved model confirm the importance of an

adequate treatment of the DNA shape in order to give more satisfactory predictions for the B→Z transition. The same conclusion has been reached by Misra and Honig[21] in their study with the NLPB theory.

C. Comparison with experiment and theoretical calculations

In this section we compare the SUCH method results with the experimental measurements and with theoretical predictions. The comparison between theory and computer simulation requires that the underlying model is the same. All the theories below are based on empty DNA models, in some cases with some differences to that used in this paper. The previous reported results from the SPMF theory are for a dodecamer and hard-sphere potential,[30] while our method considers an infinite polyelectrolyte with soft repulsive potentials. Besides, the ionic radii is used by Soumpasis as a semiempirical parameter in order to bring the results close to the experimental values and, thus, do not agree with our values. On the other hand, the discrete CC theory[27] was shown to give identical electrostatic $\beta\Delta G^{B\to Z}$ curves irrespective of the DNA length above 0.1 M NaCl. Nevertheless, it does not include the short-range repulsive interaction between particles so those results are comparable with ours with some caution. Finally, the procedure of Hirata and Levy is to use the PRISM theory previously developed which is based on the actual polyelectrolyte solution properties. The PRISM calculations were done for a slightly different potential model, a 12-6 Lennard-Jones potential with $\sigma=4.0$ Å plus a Coulombic interaction term.

Within the SPMF framework, the change in free energy per phosphate in going from an oligomer in the B form to the Z form is

$$\beta\Delta G^{\Omega}_{\text{oligomer,SPMF}}=\frac{\beta}{2N}\sum_{i=1}^{N}\sum_{j=1}^{N} W^{(2)}_{--}(r^{\Omega}_{ij}), \tag{17}$$

where $\beta W^{(2)}_{--}(r)=-\ln g_{--}(r)$ is the PMF between anions of a homogeneous electrolyte at the given concentration, $g_{--}(r)$ being the anion-anion pair correlation function. These are evaluated by means of integral equation theories, the exponential mean spherical approximation or the HNC theory in particular. The SPMF formulation can be very easily extended to the calculation of the free energy difference between the polymeric forms of B- and Z-DNA rather than the oligomer of Eq. (17). The relevant expression is

$$\beta\Delta G^{\Omega}_{\text{SPMF}}=\frac{\beta}{2p}\sum_{p}\sum_{i}^{\infty} W^{(2)}_{--}(r^{\Omega}_{ip}), \tag{18}$$

almost identical to that of the setup step of our method, Eq. (13), with the important difference that now $W^{(2)}_{--}(r)$ is the PMF between fully charged anions. Although Soumpasis set the hard diameter of the ions to 4.9 Å to reproduce the experimental data,[30] we are interested in checking the theoretical predictions and, thus, the SPMF hard sphere results presented below correspond to $\sigma=4.2$ Å, the hard diameter ''equivalent'' to the soft repulsive interactions used in our computer simulations. Fig. 6 shows the effects of the DNA length and the steepness of the repulsive potential on the free energy difference as given by the SPMF theory, using the HNC potentials of mean force. Three DNA sizes are used, a hexamer (6 base pairs, the minimum oligoelectrolyte size used in the original Pohl paper), a dodecamer (12 base pairs, the length used by Soumpasis) and the infinite polyelectrolyte. The results for a short oligoelectrolyte (the hexamer) are already relatively close to the infinite polyelectrolyte, in agreement with Pohl's experimental findings. It is also to be noticed that the effect of the polyelectrolyte length is not the same for hard ions as for soft ones (see also below). The discrete CC theory predicts a strong dependence on the DNA length for NaCl concentrations lower than 0.1 M but this dependence also vanishes in more concentrated salt solutions, at least for the studied fragments of 12 base pairs or longer.[27]

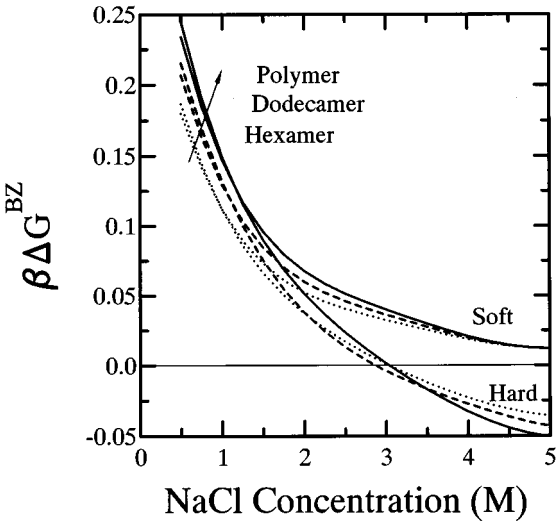

FIG. 6. Effects of the DNA length and the repulsive potential [see Eq. (12)] on the free energy difference as given by the SPMF theory for a hexamer (dotted lines), a dodecamer (dashed lines) and the infinite polyelectrolyte (solid lines). Upper curves are for a soft repulsive potential and lower curves are for hard repulsive interactions. The free energies are given in k_BT units per phosphate.

As for the effect of the repulsive interaction, the SPMF results are independent on this factor at moderate ionic strengths, but the curves for the hard and soft repulsive potentials diverge when the salt concentration is increased. The soft model results do not cross the $\Delta G^{B\to Z}_{\text{SPMF}}=0$ curve and are independent of the polyelectrolyte length at the highest salt concentrations considered while the hard repulsive potential predicts the transition for the empty model at about 3 M NaCl (almost independent of the polyelectrolyte length) but the curves do not remain together at any other salt concentration. This significant qualitative dependence on the repulsive interaction considered can be explained with the aid of the HNC PMFs between anions $W^{(2)}_{--}(r)$ in Fig. 7. At moderate concentrations (0.5 M, upper curves) the PMF does not depend on the interaction model because the low particle density makes the system scarcely sensitive to the repulsive

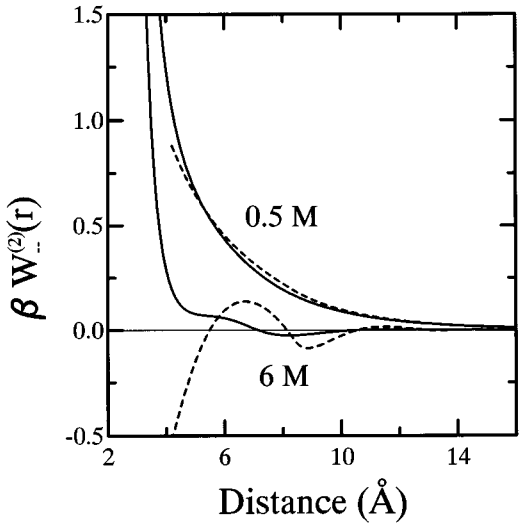

FIG. 7. The potential of mean force (in k_BT units) between two anions within an electrolyte solution as a function of the separation between them at two salt concentrations (two upper curves 0.5 M, two bottom curves 6.0 M) as given by the HNC theory. The solid lines correspond to a soft repulsive potential and the dashed lines are for the ''equivalent'' hard-sphere interactions.

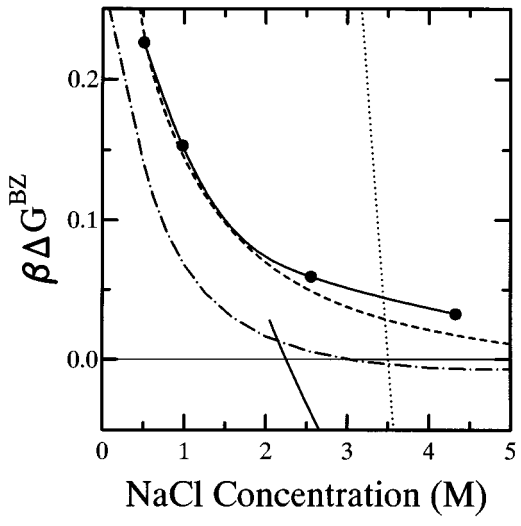

FIG. 8. Variation of the ''total'' free energy (in k_BT units per phosphate) of the B- to Z-DNA transition with monovalent added salt concentration for the empty model of polymeric DNA using soft repulsive potentials. Solid line with solid circles, *setup and charge* (SUCH) results; dashed line, SPMF theory; dashed and dotted line, discrete CC theory; dotted line, PRISM theory. The experimental curve is also drawn for comparison (solid line).

component of the potential. Monovalent electrolytes at these moderate salt concentrations are dominated by electrostatic mean field effects.[56,62] The curve is monotonic with high positive values at short distances consistent with a purely repulsive potential between like-charged particles. At higher salt concentrations the soft repulsion keeps the PMF positive at short distances, but the hard potential gives a negative PMF at contact. In the example of the figure, at 6 M, the contact value of the anion-anion radial distribution function is as high as 1.7; this is not an artifact of the HNC theory, a simulation of the same system gives a value of 1.5.

In Fig. 8 we present the dependence of the ''total'' free energy for the B- to Z-DNA transition on the added salt concentration for the empty model with soft repulsive potentials calculated with the SUCH method along with theoretical predictions. These include our SPMF calculations and the discrete CC (Ref. 27) and PRISM (Ref. 29) estimates. The experimental curve given by Eq. (1) is also drawn. The comparison between the SUCH results for the empty DNA and the experimental measurements allow us to determine whether the model is able to reproduce the DNA behaviour. The figure reveals that the model *qualitatively* describes the experimental data, with a strong stabilization of the Z with respect to the B form when the salt concentration increases as discussed above. The order of the free energy differences involved is also correct, always below 0.25 k_BT units per phosphate (in absolute value), i.e., very small. The empty model does not predict a B to Z transition up to 4.3 M NaCl but this is not significant as the model does not take into account all the contributions to the free energy balance. Adding the García and Soumpasis' estimation for the intramolecular and vibrational terms for the alternating guanine-citosine sequence $\Delta G^{B\to Z}_{\text{intra,vibr}} \approx -0.1\ k_BT$ per phosphate,[32] and being the hydration contribution still missing, the model would give the transition at about 1.5 M NaCl. Nevertheless, the slope of the curve at concentrations above 2 M is too small. A fit to the form of Eq. (1) between 2 and 5 M gives a slope of -0.051 to be compared to the experimental value of -0.3, about 6 times larger. This discrepancy is not excessive for an oversimplified model as the empty DNA, and more if one remembers that homogeneously charged cylindrical models give slopes of, at best, 50 times smaller than the experimental one.[14,15]

As for the comparison of the theories with the simulation, the SPMF approach gives excellent results. The difference between the free energies from the SPMF and the SUCH methods increases with salt concentration but the maximum discrepancy at 4.3 M is only 0.015 k_BT per phosphate (both set of values are listed in Table VI). The SPMF slope at the higher salt concentrations is -0.063, 24% larger than the SUCH value. The PRISM theory gives poor results. It also predicts a negative slope of the free energy versus the logarithm of the salt concentration but its absolute value is far too large, about 50 times the SUCH value. As mentioned above, the results reported for this theory are for a somewhat different model than that used in this paper (a Lennard-Jones potential with $\sigma = 4.0$ Å instead of a purely repulsive term with $\sigma = 4.2$ Å). Nevertheless, the departures are so important that it is not likely that the difference in the models would be responsible for the poor behaviour of the PRISM theory. Some problems of the PRISM approach have been pointed out when applied to charged systems, and are reviewed in Ref. 63. The last theoretical curve in Fig. 8 is that of the discrete CC theory. Its prediction for $\Delta G^{B\to Z}$ is always lower than the SUCH computed values but its overall shape is correct. The CC free energy slope is -0.027, about

TABLE VI. Comparison of the free energy for each Ω ($\Omega = B, Z$) DNA form with respect to the electrolyte and the ''total'' free energy difference via the SPMF theory. The free energies are given in $k_B T$ units per phosphate.

Nominal salt concentration (M)	DNA form	$\beta\Delta G^{\Omega}_{\text{uncharging}}$	$\beta\Delta G^{\Omega}$	$\beta\Delta G^{B\to Z}$	$\beta\Delta G^{\Omega}_{\text{SPMF}}$	$\beta\Delta G^{B\to Z}_{\text{SPMF}}$
	B		0.445		0.490	
0.5		0.484		0.226		0.236
	Z		0.671		0.726	
	B		0.261		0.286	
1.0		0.585		0.164		0.148
	Z		0.415		0.434	
	B		0.065		0.112	
2.5		0.750		0.060		0.050
	Z		0.125		0.162	
	B		−0.007		0.058	
4.3		0.862		0.032		0.017
	Z		0.025		0.075	

half the SUCH value. These are surprisingly good results for a theory which strictly speaking is rigorous only in the zero salt concentration limit, even if this limiting behaviour is known to be often retained up to monovalent salt concentrations of the order of 0.1 M or so.[25,64] Also, keep in mind that the CC theory cannot be applied to exactly the same simulated model since it is based on the Debye-Hückel theory which considers point-like polyelectrolyte sites and ions. In this sense, it does not include the repulsive potential effect which, as seen in Fig. 6, may affect the results at high added salt concentrations.

A last word remains to be said about the excess free energy of each DNA form with respect to the electrolyte, ΔG^{Ω}. The only remaining term to compute this quantity from Eq. (3) is the free energy of the uncharging step. This contribution corresponds, per particle, to the chemical potential of the uncharged particle in the electrolyte minus the chemical potential of an ion in the actual electrolyte, i.e.,

$$\Delta G_{\text{uncharging}} = \mu_{\text{uncharged}}(\infty) - \mu. \tag{19}$$

We could compute this difference through simulation[65] but because of its limited importance we have just made an estimation via the HNC integral equation. The resulting contribution and $\beta\Delta G^{\Omega}$ values are compared with the SPMF results in Table VI. The uncharging contribution to the ''total'' free energy is positive and increases with salt concentration, which simply means that an ion dissolved in an electrolyte solution is stabilized with respect to the uncharged sphere due to the ionic cloud of unlike charges surrounding the former. The ''total'' $\beta\Delta G^{\Omega}$'s are also always positive due to the distortion of the electrolyte structure induced by the polyelectrolyte but now they decrease with an increasing ionic force because of the increasing role of packing effects. Finally, the SPMF estimation for $\beta\Delta G^{\Omega}$ are worse than for the difference $\beta\Delta G^{B\to Z}$. This might be due to an intrinsic error of the SPMF theory which is canceled when the free energy difference is computed. Nevertheless, we cannot emphasize this statement as the method we have used to compute the uncharging term is based in the HNC equation which is closer to the SPMF philosophy than to the SUCH one (the latter being essentially based in computer simulation results).

V. CONCLUSIONS AND DISCUSSION

We have computed the free energy of the transition between the B and Z conformers of DNA through a thermodynamic route which we call *setup and charge* (SUCH) method. The method divides the ''total'' free energy in the electrostatic and nonelectrostatic contributions. It is shown that the former is considerably more important than the latter one. For this reason, whereas the correct evaluation of the electrostatic contribution requires simulation results, the calculation of the nonelectrostatic term may be accurately done through the use of approximate theories. The chosen *empty* DNA model is extremely simple and corresponds approximately to that used in previous theoretical estimates of the free energy difference between both DNA forms. The SUCH results indicate that the model is able to describe *qualitatively* the main features of the transition. We have paid especial attention to the variation of the free energy difference with the salt concentration as we believe that other contributions not included in the model are less salt concentration dependent and, thus, they may affect the transition midpoint but not the slope of the curve. An advantage of the method is that it enables, in principle, to incorporate systematic refinements of the model. It is thus interesting to discuss which are the possible improvements to the empty model. PB calculations for homogeneously charged models lead to almost flat curves whereas the recent PB results reported by Misra and Honig[21] for an all-atom DNA predict a slope relatively close to the experimental one. For this reason we believe that a simple refinement to the empty model is to incorporate the overall DNA shape (including the steric effects due to the bases and sugars) in a manner as simple as possible. Preliminary calculations for the grooved primitive model described in Ref. 59 confirm the importance of the DNA shape and its effect on the slope of the $B\to Z$ free energy difference curve. The use of the method for complex DNA models such that used in the Misra and Honig paper is more tricky. Although nothing in the SUCH method disables its application to those models, the use of the KSA to evaluate the setup term is more questionable because of the small separation between the atoms. In this way, the calculation of a minor contribu-

tion as the setup term may require an enormous computational effort. Apart of a more adequate definition of the DNA shape, the next refinement to the model would likely be the explicit inclusion of the solvent. Recent experiments show that this cannot be done in a simplified way as the changes observed in the transition do not correlate well with the dielectric constant nor the molecular shape of the solvent.[66] It seems thus that the treatment of the solvent must be done at the molecular level.

Another important issue of the paper is that it allows for the first time an unambiguous test of the theories in the context of the $B \rightarrow Z$ DNA transition. We show that, despite its tremendous simplicity, the SPMF theory quantitatively describes the free energy of the empty model. The PRISM theory gives discordant results and the discrete CC approach leads to qualitatively valid results. It seems that the quality of the SPMF theory relies on the use of the KSA. The studies on associative systems and, in particular, their application to pearl-neck polymer models[47] provide some theoretical basis but new studies are required to establish the validity of the KSA for these type of models. The objections to the SPMF theory derive from its simplicity. First, KSA deviations from the MC results are systematic. This does not affect the quality of the results in which a free energy *difference* is required. In other cases, the errors may be more important. In particular, the KSA deviations lead to an incorrect description of the charge of the ionic envelope of the macromolecule and, thus, the application of the KSA-SPMF theory to problems in which the correct treatment of the electrostatic forces is critical should be avoided. But the more important objection to the theory just comes from its success. Given the simplicity of the SPMF approach one is desiring to apply it to more complex systems but the same simplicity makes it very difficult to extend the theory to such models.

In this work we have restricted our study to the empty DNA model in presence of monovalent salt at moderate to high concentrations. This is simply because these are the conditions for which more theoretical and experimental results have been reported. It has been mentioned above that the SUCH method is not limited to these conditions. It seems then interesting to use the method in conditions where there are doubts about the feasibility of the theories and/or experiments. Some experimental data indicate that there is a maximum in the stability of the B form at low concentrations of monovalent added salt.[67–69] It has been conjectured that the experimental findings might be the effect of traces of divalent cations present in the samples. Theoretical reports also contradict each other. The former PB calculations[14] predicted the stabilization of the Z form at very low salt concentrations but other PB calculations showed that the B form is always more stable.[15] To complicate the situation even more, recent careful experiments seem to confirm the maximum stability of the B form at low salt concentrations.[70] Perhaps the SUCH method might throw some light on this question.

ACKNOWLEDGMENTS

This work was partially supported by Grant SEUI No. PB93-0085 provided by the Dirección General de Investigación Científica y Tecnológica of Spain. The HNC calculations in this work have been done with the program HNCR,[71] extended by F. Bresme to consider multiple species and the soft repulsive potential of Eq. (12).

[1] F. M. Pohl and T. M. Jovin, J. Mol. Biol. **67**, 375 (1972).
[2] A. H.-J. Wang, G. J. Quigley, F. J. Kolpak, J. L Crawford, J. H. van Boom, G. van der Marel, and A. Rich, Nature **282**, 680 (1979).
[3] W. Saenger, *Principles of Nucleic Acid Structure* (Springer, New York, 1984).
[4] C. K. Singleton, J. Klysik, S. M. Stirdivant, and R. D. Wells, Nature **299**, 312 (1982).
[5] L. J. Peck, A. Nordheim, A. Rich, and J. C. Wang, Proc. Natl. Acad. Sci. USA **79**, 4560 (1982).
[6] F. M. Pohl, Cold Spring Harbor Symp. Quant. Biol. **47**, 113 (1983).
[7] A. R. Rahmouni and R. D. Wells, Science **248**, 359 (1989).
[8] D. M. Soumpasis and T. M. Jovin, in *Nucleic Acids and Molecular Biology*, edited by F. Eckstein and D. M. J. Lilley (Springer, Berlin, 1987), p. 85.
[9] T. M. Jovin, D. M. Soumpasis, and L. P. McIntosh, Annu. Rev. Phys. Chem. **38**, 521 (1987).
[10] M. Behe and G. Felsenfeld, Proc. Natl. Acad. Sci. USA **78**, 1619 (1981).
[11] T. M. Jovin, L. P. McIntosh, D. J. Arndt-Jovin, D. A. Zarling, M. Robert-Nicoud, J. H. van de Sande, K. F. Jorgenson, and F. Eckstein, J. Biomol. Struct. Dynam. **1**, 21 (1983).
[12] M. Fuoss, A. Katchalsky, and S. Lifson, Proc. Natl. Acad. Sci. USA **37**, 579 (1951).
[13] R. A. Marcus, J. Chem. Phys. **23**, 1057 (1955).
[14] M. D. Frank-Kamenetskii, A. V. Lukashin, and V. V. Anshelevich, J. Biomol. Struct. Dynam. **3**, 35 (1985).
[15] D. M. Soumpasis, J. Biomol. Struct. Dynam. **6**, 563 (1988).
[16] A. V. Lukashin, D. B. Beglov, and M. D. Frank-Kamenetskii, J. Biomol. Struct. Dynam. **8**, 1113 (1991).
[17] M. Le Bret and B. H. Zimm, Biopolymers **23**, 271 (1984).
[18] C. S. Murthy, R. J. Bacquet, and P. J. Rossky, J. Phys. Chem. **89**, 701 (1985).
[19] J. P. Demaret and M. Guéron, Biophys. J. **65**, 1700 (1993).
[20] M. Guéron and J. P. Demaret, Proc. Natl. Acad. Sci. USA **89**, 5740 (1992).
[21] V. K. Misra and B. Honig, Biochemistry **35**, 1115 (1996).
[22] G. S. Manning, J. Chem. Phys. **51**, 924 (1969).
[23] G. S. Manning, J. Chem. Phys. **51**, 934 (1969).
[24] G. S. Manning, J. Chem. Phys. **51**, 3249 (1969).
[25] M. O. Fenley, G. S. Manning, and W. K. Olson, Biopolymers **30**, 1191 (1990).
[26] G. S. Manning, in *Theoretical Biochemistry and Molecular Biophysics*, edited by D. L. Beveridge and R. Lavery (Adenine, New York, 1990), Vol. 1, p. 191.
[27] M. O. Fenley, G. S. Manning, and W. K. Olson, Biopolymers **30**, 1205 (1990).
[28] F. Hirata and R. M. Levy, Chem. Phys. Lett. **136**, 267 (1987).
[29] F. Hirata and R. M. Levy, J. Phys. Chem. **93**, 479 (1989).
[30] D. M. Soumpasis, Proc. Natl. Acad. Sci. USA **81**, 5116 (1984).
[31] J. Kirkwood, J. Chem. Phys. **3**, 300 (1930).
[32] A. E. García and D. M. Soumpasis, Proc. Natl. Acad. Sci. USA **86**, 3160 (1989).
[33] W. Saenger, W. N. Hunter, and O. Kennard, Nature **324**, 385 (1986).
[34] J. C. Gil Montoro and J. L. F. Abascal, Europhys. Lett. **34**, 471 (1996).
[35] M. Mezei and D. L. Beveridge, Ann. NY Acad. Sci. **482**, 1 (1986).
[36] W. Ross, C. Hardin, I. Tinoco, S. W. Rao, D. A. Pearlman, and P. A. Kollman, Biopolymers **28**, 1939 (1989).
[37] L. X. Dang, D. A. Pearlman, and P. A. Kollman, Proc. Natl. Acad. Sci. USA **87**, 4630 (1990).
[38] D. A. Pearlman and P. A. Kollman, Biopolymers **29**, 1193 (1990).
[39] B. Widom, J. Chem. Phys. **39**, 2808 (1963).
[40] K. K. Mon and R. B. Griffiths, Phys. Rev. A **31**, 965 (1985).
[41] J. Harris and S. A. Rice, J. Chem. Phys. **88**, 1298 (1988).
[42] J. I. Siepmann, Mol. Phys. **70**, 1145 (1990).
[43] P. Grassberger, J. Phys. A **26**, 1023 (1993).
[44] P. Grassberger, J. Phys. A **26**, 2769 (1993).
[45] T. S. Sorensen, J. Chem. Soc. Faraday Trans. **87**, 479 (1991).

[46] B. Svensson, T. Akesson, and C. E. Woodward, J. Chem. Phys. **95**, 2717 (1991).
[47] Y. Zhou and G. Stell, J. Chem. Phys. **96**, 1507 (1992).
[48] D. Beglov and B. Roux, J. Chem. Phys. **103**, 360 (1995).
[49] E. J. W. Verwey and J. G. Overbeek, *Theory of the Stability of Lyophobic Colloids* (Elsevier, Amsterdam, 1948).
[50] J. O'M. Bockris and A. K. N Reddy, *Modern Electrochemistry* (Plenum, New York, 1970).
[51] M. Fixman, J. Chem. Phys. **70**, 4995 (1979).
[52] A. H.-J. Wang, G. J. Quigley, F. J. Kolpak, G. van der Marel, J. H. van Boom, and A. Rich, Science **211**, 171 (1981).
[53] C. K. Mitra, M. H. Sarma, and R. H. Sarma, J. Am. Chem. Soc. **103**, 6727 (1981).
[54] V. Sasisekharan, M. Bansal, S. K. Bramachari, and G. Gupta, in *Biomolecular Stereodynamics*, edited by R. H. Sarma (Adenine, New York, 1981), p. 123.
[55] S. Arnott and D. W. Hukins, Biochem. Biophys. Res. Commun. **47**, 1504 (1972).
[56] J. L. F. Abascal, F. Bresme, and P. Turq, Mol. Phys. **81**, 143 (1994).
[57] P. J. Rossky, J. B. Dudowick, B. L. Tembe, and H. L. Friedman, J. Chem. Phys. **73**, 3372 (1980).
[58] J. C. Gil Montoro and J. L. F. Abascal, Mol. Simul. **14**, 313 (1995).
[59] J. C. Gil Montoro and J. L. F. Abascal, J. Chem. Phys. **103**, 8273 (1995).
[60] J. C. Gil Montoro and J. L. F. Abascal, Mol. Phys. **89**, 1071 (1996).
[61] H. L. Gordon and S. Goldman, Mol. Simul. **3**, 213 (1989).
[62] J. L. F. Abascal and P. Turq, Chem. Phys. **153**, 79 (1991).
[63] L. R. Pratt, G. Hummer, and A. E. García, Biophys. Chem. **51**, 147 (1994).
[64] G. S. Manning, Biophys. Chem. **7**, 95 (1977).
[65] G. Hummer, L. R. Pratt, and A. E. García, J. Chem. Phys. **100**, 1206 (1996).
[66] R. S. Preisler, H. H. Chen, M. F. Colombo, Y. Choe, B. J. Short, Jr., and D. C. Rau, Biochemistry **34**, 14 400 (1995).
[67] P. K. Latha and S. K. Brahmachari, FEBS Lett. **182**, 315 (1985).
[68] W. C. Krueger and M. D. Prairie, Biopolymers **24**, 905 (1985).
[69] G. Feuerstein, L. J. Marton, M. A. Keniry, D. L. Wade, and R. H. Shafer, Nucleic Acids Res. **13** (1985).
[70] V. I. Ivanov, A. T. Karatepetyan, E. E. Minyat, and J. Sagi, Mol. Biol. **27**, 713 (1993).
[71] E. Lomba, Comput. Phys. Commun. **69**, 420 (1992).