Inclusion of Ionic Interactions in Force Field Calculations of Charged Biomolecules – DNA Structural Transitions

REINHARD KLEMENT, DIKEOS M. SOUMPASIS, EBERHARD V. KITZING, and THOMAS M. JOVIN

Max-Planck Institut für Biophysikalische Chemie, Abteilung Molekulare Biologie, Postfach 2841, 3400 Göttingen, Federal Republic of Germany

SYNOPSIS

The potential of mean force (PMF) approach for treating polyion–diffuse ionic cloud interactions [D. M. Soumpasis (1984) *Proceedings of the National Academy of Sciences USA* **81**, 5116–5120] has been combined with the AMBER force field describing intramolecular interactions. The resultant generalized AMBER-PMF force field enables one to treat the conformational stabilities and structural transitions of charged biomolecules in aqueous electrolytes more realistically. For example, we have used it to calculate the relative stabilities of the B and Z conformations of $d(C\text{-}G)_6$, and the B and heteronomous (H) conformations of $dA_{12} \cdot dT_{12}$, as a function of salt concentration. In the case of $d(C\text{-}G)_6$, the predicted $B\text{–}Z_I$ transition occurs at $2.4M$ and is essentially driven by the phosphate–diffuse ionic cloud interactions alone as suggested by the results of earlier PMF calculations. The Z_{II} conformer is less stable than the B form under all conditions. It is found that the helical parameters of the refined B and Z structures change with salt concentration. For example, the helical rise of B-DNA increases about 10% and the twist angle decreases by the same amount above $1M$ NaCl.

In the range of $0.01\text{–}0.3M$ NaCl, the H form of $dA_{12} \cdot dT_{12}$ is found to be more stable than the B form and its stability increases with increasing salt concentration. The computed greater relative stability of the H conformation is likely due to noninclusion of the free energy contribution from the spine of hydration, a feature presumed to stabilize the B form of this sequence.

INTRODUCTION

The molecular mechanics or force field approach is currently the most widespread, computationally feasible methodology for calculating the intrinsic energy of large biomolecular structures (see Refs. 1 and 2 for recent reviews).

Typical force fields for proteins and nucleic acids[3–8] in use today include terms describing the interaction of bonded atoms (bond stretching and bending, torsional contributions) and so-called nonbonded interactions between atoms not connected via 1–3 consecutive covalent bonds (core repulsion, attractive dispersion forces, Coulomb interactions, and hydrogen bonds).

Using such a force field description, one can estimate the relative intrinsic energy E_{in} of a given molecular conformation, for example, derived from x-ray analysis, or determine conformations that minimize this energy, at least in a local sense. In addition, using normal mode analysis or molecular dynamics one can calculate the vibrational free energy contribution arising from fluctuations of the structure around a minimal energy conformation.[9,10] Since the evaluation of translational and rotational free energies of the molecule as a whole is relatively straightforward, one can obtain the intrinsic free energy of a molecular conformation *in the absence of solvent* at the cost of some programming effort and sufficient computer time.

However, it is clear that this free energy contribution alone does not suffice to describe the stability and structural transitions of biomolecular structures in solution, since in most, if not all, cases

© 1990 John Wiley & Sons, Inc.
CCC 0006-3525/90/051089-15 $04.00
Biopolymers, Vol. 29, 1089–1103 (1990)

of interest the ubiquitous biomolecule–water and biomolecule–ion interactions present in solution are at least as important as the intermolecular interactions accounted for in currently used force fields. This circumstance applies particularly to polyions such as DNA, which bears a large number of net charges (sugar–phosphate backbone) and electronegative base atoms. As judged from recent crystallographic studies, of the three major DNA conformations comprising the right-handed B-DNA,[11] A-DNA,[12] and left-handed Z-DNA,[13,14] the geometry of the DNA charged sites gives rise to the formation of well-defined, conformation- and sequence-dependent networks of water molecules,[15] which presumably also exist in solution. In addition to this kind of short-range structural hydration, hydrophobic base interactions and long-range electrostatic coupling to both bulk water and the diffuse cloud of mobile salt ions surrounding a DNA polyanion in typical solution environments give rise to additional important contributions to the overall conformational free energy.

A quantitative and simultaneous treatment of all these solvent effects is not possible at present. However, the contribution arising from the coupling of the hydrated DNA charge distribution to bulk water and the diffuse cloud of mobile hydrated ions can be reasonably well estimated using the approach based on the potential of mean force (PMF).[16–18] Thus, one can treat the influence of nonbinding ions, such as the alkali halides on DNA conformation. This formalism has been used successfully to describe the salt-induced B-Z[16,17] and B-A[17] transitions, and their dependence on counterion valence[16] and counterion size.[19] As discussed further below, using PMFs one can construct a new generation of substantially more realistic force fields according to which the effective interactions of charged sites in solution are represented by solvent averaged statistical PMFs. The range of applicability of force field calculations can be thereby expanded and the salt effects on biomolecular stability and structural transitions treated in a quantitative way. In addition, this approach has been used to study the harmonic dynamics of DNA oligomers in a wide range of ionic conditions. [36]

THEORETICAL METHODOLOGY

In this work we use the well-known AMBER force field elaborated by Kollmann and colleagues,[6,7] which has been updated and refined by one of us (E.v.K.).[8,34] The original AMBER intrinsic energy E_{in} of a given biomolecular conformation (e.g., of an oligomeric DNA) has the form

$$E_{in} = \sum_{\text{bonds}} K_d(d - d_{eq})^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_{eq})^2 + \sum_{\text{dihedrals}} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)] + \sum_{\text{van der Waals}} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right] + \sum_{\text{H bonds}} \left[\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}}\right] + \sum_{\text{Coulomb}} \frac{q_i q_j}{\epsilon r_{ij}} \quad (1)$$

The first three terms describe the contributions of bond stretching, bond angle bending, and bond torsion; the last three terms describe the van der Waals interactions, hydrogen bonding, and Coulomb interactions of the atoms (or pseudoatoms) comprising the structure. K_d, K_θ, V_n, A_{ij}, B_{ij}, C_{ij}, and D_{ij} are energy parameters; d_{eq} and θ_{eq} equilibrium bond lengths and bond angles, respectively; and r_{ij} the distance between atoms i and j. The partial atomic charges q_i, q_j of atoms i, j are obtained from separate quantum chemical calculations, and ϵ is an effective dielectric constant ($2 \leq \epsilon \leq 8$) assumed to describe the weak dielectric screening due to molecular polarizability. More detailed descriptions can be found in Refs. 6–8.

An immediate problem arising with E_{in} as described by Eq. (1) is that both the strength and the range of the electrostatic interactions of sites bearing net charge (i.e., the phosphates in the case of DNA in solution) are hopelessly overestimated, since they are modeled as if these charges were interacting quasi in vacuo. In reality, both the strength and the range of their interactions are greatly reduced through (a) dielectric screening provided by water, and (b) ionic screening by the statistical cloud of mobile ions surrounding the structure in solution. Explicit consideration of these screening effects is absolutely necessary since otherwise the force field approach cannot be applied to the analysis of DNA conformations in a meaningful way. For instance, in a molecular dynamics study[3] using a force field similar to AMBER, the unscreened Coulomb repulsions of the phosphates were so strong that the calculated DNA double helix simply fell apart. Several groups using the

AMBER program or similar force fields have tried to cope with the Coulomb problem using various heuristic recipes such as (a) simply setting the phosphate charges equal to zero,[3] (b) placing neutralizing counterions (e.g., Na^+) on the phosphates,[21] (c) reducing the phosphate charges according to Manning's counterion condensation hypothesis,[22,23] (d) assuming distance-dependent dielectric constants,[21-23] and/or (e) simply ignoring Coulomb interactions beyond some arbitrary cutoff distance of the order 10 Å.[21,22] In these computational recipes, the electrostatic contribution remains totally independent of the prevailing environmental conditions such as salt concentration and temperature, and therefore salt-induced structural variations and transitions cannot be treated.

Here we treat the Coulomb problem within the statistical mechanical framework introduced by Soumpasis.[16-18] We retain all interactions present in the usual AMBER form of the intrinsic energy E_{in} described by Eq. (1), *but replace the electrostatic interactions* for all atoms with a distance r_{ij} larger than a given distance of closest approach σ through effective interactions (PMF's) obtained from statistical mechanics. The contribution of these interactions to the total free energy balance, denoted F_1, is approximated in the form

$$F_1 = \sum_{i>j} W_{11}(r_{ij}) \qquad (2)$$

where W_{11} is the anion–anion potential of mean force in a homogeneous electrolyte solution, considered here to be a fully dissociated aqueous 1 : 1 electrolyte of specified composition and thermodynamic state; r_{ij} is the separation of charged sites i and j in the conformation envisaged. A first-order model description of an aqueous ionic system is provided by the well-known restricted primitive model (RPM) picturing all hydrated ions as charged hard spheres of the same average diameter σ and interacting in a dielectric continuum of bulk dielectric constant ϵ_w.

Adopting this model description, calculation of the statistical interaction $W_{11}(r)$ can be performed using approximations of varying accuracy, computational cost, and analytic character. In previous work on salt-induced DNA structural transitions both a semianalytic approximation[16] and very accurate numerical PMFs obtained from solution of the hypernetted chain (HNC) integral equation[17] have been used.

In view of the diversity of force field applications, which also include energy minimization, harmonic analysis, and molecular dynamics, it is more convenient to sacrifice some of the accuracy obtained using totally numeric approaches such as the HNC equation in favor of more analytic, computationally flexible PMFs such as those obtained through the so-called exponential mean spherical approximation (EXP-MSA).[24,25] As discussed in the appendix, using this approximation for an 1 : 1 electrolyte, W_{11} can be decomposed into a hard-sphere contribution W_0 and a charge dependent contribution W_e

$$W_{11}(\tilde{r}) = W_0(\tilde{r}) + W_e(\tilde{r})$$

$$= -k_B T \ln g^0(\tilde{r}) - k_B T h_D(\tilde{r}) \quad \text{for } \tilde{r} \geq 1 \qquad (3)$$

where $\tilde{r}$ is the reduced distance r/σ, σ is the effective counterion diameter or average distance of closest approach of an anion—cation pair, and $g^0(\tilde{r})$ is the essentially exact *uncharged* hard-spheres pair correlation function obtained through Monte Carlo simulations,[26] or at lower densities, through the solution of the Percus–Yevick equation.[27,28] The function $h_D(\tilde{r})$ is a quantity obtained in the MSA treatment.[29,30] More details are presented in the appendix.

Computational Implementation

As a first application, we use the formalism to calculate the stabilities and structural variations of Z-DNA, heteronomous DNA, and B-DNA as a function of monovalent salt (NaCl) concentration.

The starting dodecamer structures $d(C\text{-}G)_6 \cdot (C\text{-}G)_6$ used in the computations were generated from the fiber diffraction data for B-DNA[31] and the two idealized left-handed crystallographic conformers Z_I- and Z_{II}-DNA.[32] Each DNA structure was optimized at a bulk salt concentration of 0.3*M* using the Bremermann[33] method (E. v. Kitzing, to be published) in order to obtain stable structures. We then refined the latter further by applying the AMBER-PMF force field (see below) at 12 different salt concentrations in the range of 0.01–5.0*M*. The helical parameters of the optimized structures were calculated from cartesian coordinates using the methodology described in Ref. 34.

In the AMBER-PMF program, the nonelectrostatic interactions are treated within the framework of the standard AMBER force field.[7] Electrostatic interactions are calculated in the following

atom-dependent manner:

1. Due to their net charge and solvent accessibility, the anionic phosphate oxygens are the DNA sites that exhibit the strongest coupling to the diffuse ionic cloud. We use the full PMF W_{11} of Eq. (3) to describe their effective mutual interactions.
2. All other electrostatic interactions are treated as follows:
   (a) For distances greater than σ, we replace the standard Coulomb AMBER term with the charge-dependent part W_e of the PMF [see Eq. (3)].
   (b) For distances less than σ, the Coulomb form is retained with $\epsilon = 4$. A smooth transition to W_e is achieved by means of a fifth-order polynomial.[8]

It is computationally convenient to represent all electrostatic terms by means of cubic splines of sufficient resolution (we use 300 intervals over a distance of 25 Å). The advantage of this method lies in the fast and accurate calculation of the complicated functions and derivatives required in minimization algorithms, and for harmonic analysis and molecular dynamics.

All calculations were performed on an IBM 3090-200 mainframe computer using the current set of AMBER parameters. The average CPU time for optimization of a structure at one salt concentration was about 70 min.

RESULTS AND DISCUSSION

The EXP-MSA PMF

The full EXP-MSA PMF and the constituent electrostatic and hard-sphere contributions are shown in Figs. 1, 2(a), and 2(b), respectively, as functions of the separation r of the charges and the bulk salt concentration. In the calculations, it is convenient to use the dimensionless reduced ionic number density $\tilde{\rho}$, which for a $1:P$ salt MX_P (M, cation; X, monovalent anion) is given by

$$\tilde{\rho} = 6.023 \cdot 10^{-4}(P + 1)c\sigma^3 \qquad (4)$$

where c is the molar salt concentration and σ is in ångstroms. The temperature T in all calculations reported here is 300 K, the solvent dielectric constant is that of water ($\epsilon_w = 78.4$), and the value of σ for NaCl is $\sigma = 4.9$ Å as in previous work.[16-19]

At higher salt concentrations, the behavior of the EXP-MSA PMF is dominated by the many-bodied hard-sphere effects, which together with the oscillatory charge-dependent terms [see Eqs. (3), (A5)–(A6)] give rise to a pronounced local structure [Figs. 1 and 2(b)]. In the case of 1 : 1 electrolytes of $\sigma \approx 5$ Å (e.g., NaCl), hard-sphere contributions become progressively negligible below $0.1M$ and the PMF approaches that given by the Debye–Hückel approximation.

A measure of the importance of the hard sphere relative to the screening contribution is the fractional quantity $f_0 = |W_0|/(|W_e| + |W_0|)$, as depicted in Fig. 2(c) for all distances and salt concentrations considered.

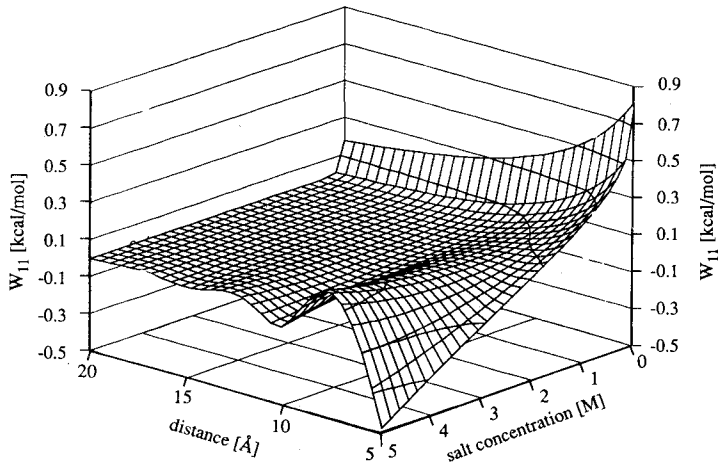

Figure 1. Dependence of the anion–anion EXP-MSA PMF W_{11} on distance and monovalent salt concentration ($\sigma = 4.9$ Å). Due to screening effects the anion–anion repulsion is reduced with increasing salt concentrations.

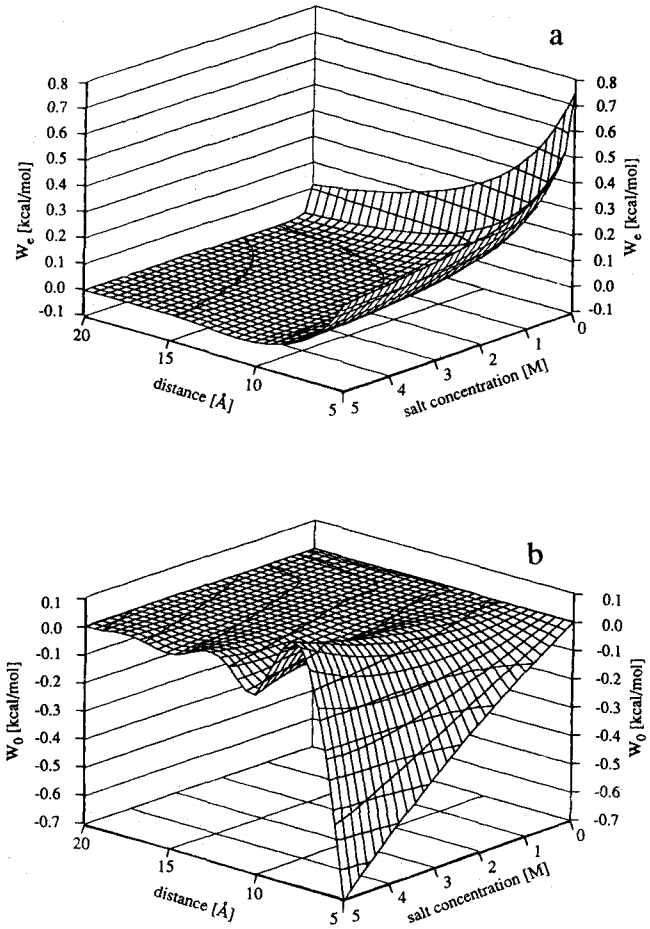

Figure 2. EXP-MSA potential energy surface of an anion–anion pair. (a) Charge-dependent contribution W_e, (b) hard-sphere contribution W_0, and (c) fractional hard-sphere contribution f_0. Parameters and dimensions as in Fig. 1. For ionic strength $< 0.5M$ the hard-sphere contribution to the PMF is negligible compared to the electrostatic contribution.

The B–Z Transition

The $B–Z_I$ and $B–Z_{II}$ energy differences of the $d(C–G)_6 \cdot d(C–G)_6$ dodecamers calculated by means of our generalized AMBER-PMF force field are shown in Fig. 3. In the case of the B and Z_I conformations, we obtain a high salt right-to-left transition midpoint at $2.4M$ NaCl, in correspondence with experimental findings[20,35] and first quantitatively described considering just the phosphate–phosphate interactions within the PMF framework.[16,17]

The present calculations show that in the case of the $B–Z_I$ equilibrium of the canonical $d(C–G)_6 \cdot d(C–G)_6$ sequence, the contribution of nonelectrostatic intramolecular interactions is relatively unimportant. The transition is driven primarily by solvent-averaged phosphate–phosphate interactions, the basis for the success of the PMF treatment considering only phosphate–phosphate interactions. That is, although the other energy contributions are much larger than the PMF parts in each conformation, they cancel when one builds energy differences, the quantities of relevance for conformational equilibria. This feature is evident from inspection of Table I, in which the primary salt dependence enters via the PMF terms in the electrostatic contribution. The same behavior is also observed even when one includes the vibrational free energy contributions in addition to the energies discussed here.[36]

As briefly discussed earlier,[16] the main reason for the stabilization of the Z conformation of the canonical $d(C–G) \cdot d(C–G)$ sequence at high salt

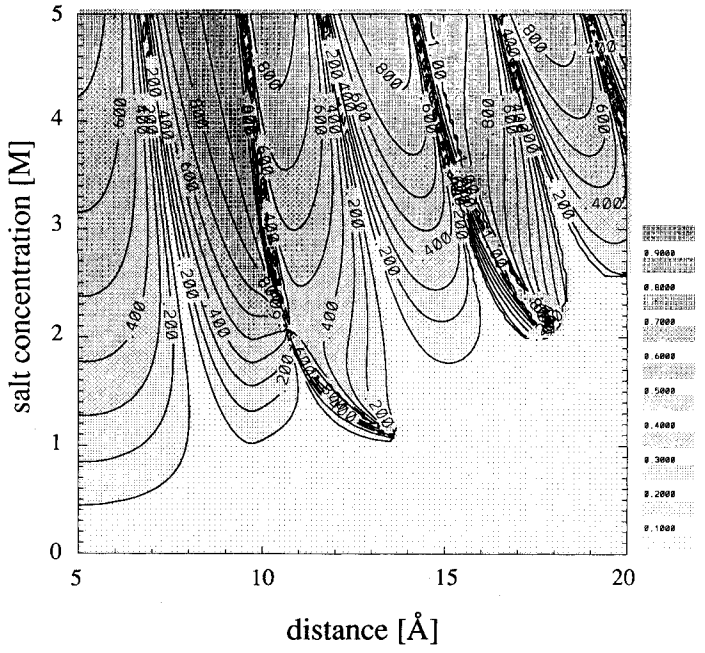

Figure 2c. (*Continued from the previous page.*)

concentrations is the dramatic change occurring in W_{11} (and therefore in the PMF hypersurface F_1) as the salt concentration increases above $1M$. In contrast to the situation at low salt, the hypersurface F_1 becomes highly structured [many maxima and minima, Figs. 2(a) and 2(c)] and the conformation that has a geometry corresponding to a lower F_1 (i.e., Z in the case at hand) acquires greater stability. This behavior can be seen in Figs. 4(a) and 4(b), in which the phosphate–phosphate distance distributions of the B and Z_I conformations are depicted together with the distance dependence of the high salt PMF.

Although the PMF contributions are always present, it is very likely that in other cases (e.g., B–Z transitions of other sequences and/or other structural equilibria), intramolecular contributions as well as vibrational free energy terms and structural hydration will be at least as important or will even dominate the free energy balance. For example, as seen in Fig. 3 in the case of the B and Z_{II} conformations, the present calculation does not

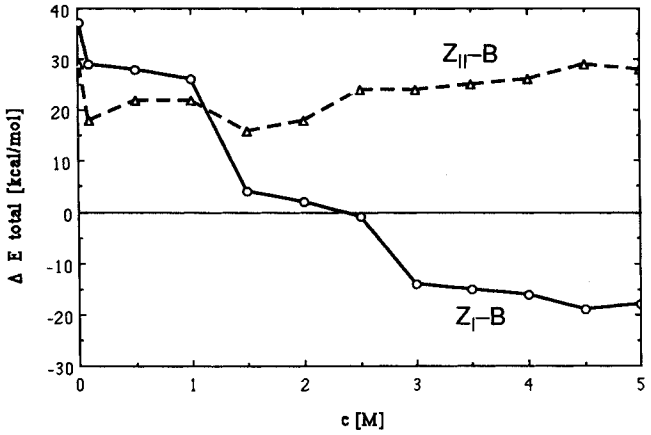

Figure 3. Energy difference $E_{Z_I}-E_B$(–) and $E_{Z_{II}}-E_B$(---) of the $d(C-G)_6 \cdot d(C-G)_6$ structures. The midpoint of $B-Z_I$ transition occurs at $2.4M$ NaCl concentration (experimental $2.5M$[20]).

Table I Energies of the B–, Z_I–, and Z_{II}–DNA Structures $d(C–G)_6 \cdot d(C–G)_6$ Kilocalories per Mole at Different Bulk Salt Concentrations

c [M]	E_T[a]	$E_{\text{bond angle}}$[b]	E_{dihedral}[c]	E_{VDW}[d]	E_{estat}[e]
B–DNA					
0.01	−1033.	211.4	348.7	−405.6	293.0
0.1	−1052.	211.2	350.0	−404.4	271.2
0.5	−1068.	210.7	350.8	−402.3	252.0
1.0	−1074.	210.7	351.3	−401.8	245.1
1.5	−1082.	213.0	348.8	−406.6	239.5
2.0	−1085.	213.3	347.7	−406.1	236.1
2.5	−1094.	213.5	347.1	−405.2	225.9
3.0	−1104.	225.7	340.9	−409.0	213.5
3.5	−1105.	225.6	341.0	−408.8	212.7
4.0	−1104.	225.8	339.9	−409.5	214.4
4.5	−1105.	225.0	341.4	−408.5	212.2
5.0	−1104.	225.4	341.1	−408.8	213.2
Z_I–DNA					
0.01	−996.	217.2	408.8	−430.0	286.1
0.1	−1023.	212.9	410.4	−422.6	253.2
0.5	−1040.	212.3	411.5	−421.6	234.3
1.0	−1048.	212.5	412.5	−422.7	226.6
1.5	−1078.	210.9	405.7	−417.9	200.5
2.0	−1083.	212.0	406.1	−420.3	195.7
2.5	−1095.	210.0	406.4	−416.0	181.2
3.0	−1118.	215.0	396.5	−421.5	169.6
3.5	−1120.	212.0	399.4	−423.6	168.1
4.0	−1120.	213.4	397.7	−422.9	168.1
4.5	−1124.	210.1	401.4	−424.0	164.1
5.0	−1122.	211.7	399.4	−423.2	165.9
Z_{II}–DNA					
0.01	−1004.	204.6	419.4	−464.3	314.8
0.1	−1034.	204.3	419.9	−465.9	285.3
0.5	−1046.	204.6	421.4	−463.3	270.3
1.0	−1052.	204.4	422.0	−463.3	263.2
1.5	−1066.	208.3	421.8	−466.1	247.7
2.0	−1067.	208.4	421.9	−466.0	246.1
2.5	−1070.	211.3	421.6	−468.5	243.5
3.0	−1080.	208.1	428.2	−468.3	231.7
3.5	−1080.	208.1	428.5	−468.6	232.6
4.0	−1078.	208.2	428.2	−468.5	234.6
4.5	−1076.	208.3	428.2	−468.6	236.2
5.0	−1076.	208.1	429.2	−469.1	236.3

[a] Total energy (internal plus PMF terms).
[b] Bond angle energy contribution.
[c] Dihedral angle energy contribution.
[d] van der Waals energy contribution.
[e] Electrostatic energy contribution including PMF terms.

yield a right-to-left transition, whereas judging from the PMF contribution alone a transition should occur at $0.2M$.[17] However, for salt concentrations lower than $1.2M$, we predict that the Z_{II} structure is energetically more favorable than the Z_I structure. This seems to be in agreement with experimental data suggesting that the low salt conditions favor Z_{II} relative to Z_I.[37,38]

Variation of the B and Z Structures Due to Salt

Some helical parameters of the optimized B and Z structures are compiled in Table II. It is found that all these structures display systematic structural variations as a function of the salt concentration. In Fig. 5 we depict the dependence of the helical twist angle and helical rise on NaCl concentration for the B, Z_I, and Z_{II} conformations of

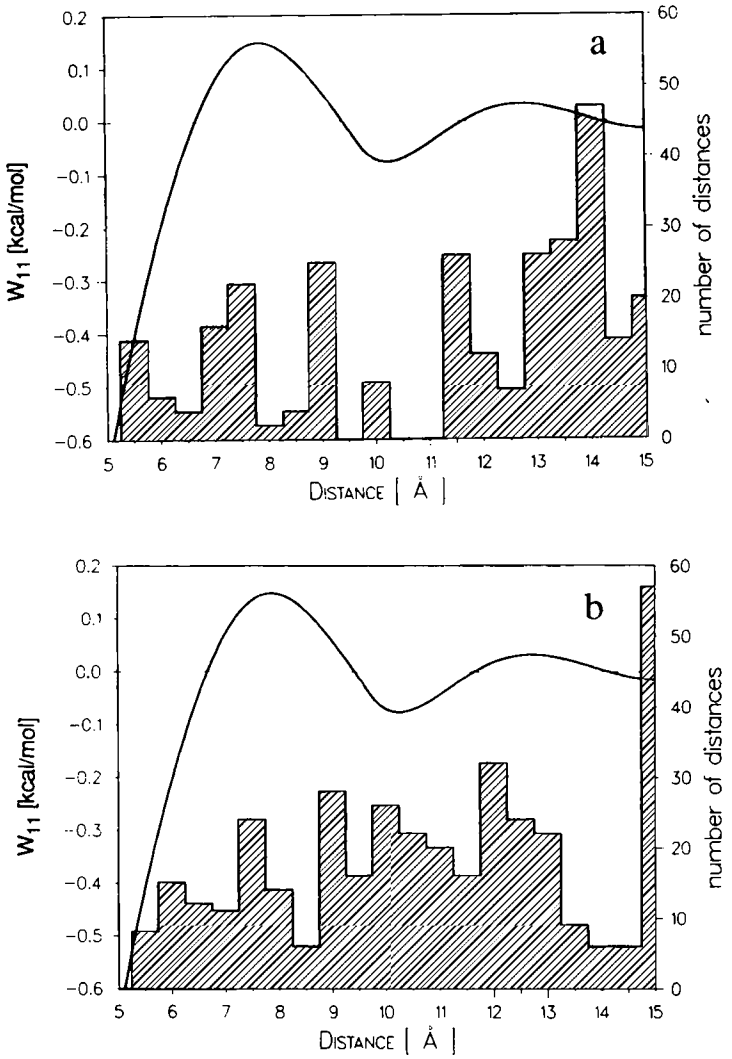

Figure 4. Distribution of the anionic oxygen distances (histogram; right scale) in the B-DNA form (a) and Z_I-DNA form (b) of $d(C-G)_6 \cdot d(C-G)_6$. The EXP-MSA PMF (solid curve; left scale) at $5M$ salt concentration is also depicted. The histograms show the occurrences of oxygen–oxygen distances within an interval of 1 Å. In the B form most of the short-range distances coincide with the first maximum of the potential. The average interphosphate distance in Z_I-DNA is shorter than in B-DNA. In particular, some of the short distances in Z_I-DNA (b) yield an energetically favorable contribution to the total internal energy.

$d(C-G)_6 \cdot d(C-G)_6$. The significant features are (a) an unwinding of the B form and a winding of the left-handed Z form above $1M$ NaCl; and (b) corresponding increases in the helical rises of the B and Z_I forms, which attain their canonical values in the range of $1.5-2.5M$ NaCl.

Heteronomous Poly(dA) · Poly(dT)

In order to estimate the relative stability of heteronomous DNA relative to B-DNA, we have optimized the structure of the $dA_{12} \cdot dT_{12}$ dodecamer in both conformations and for various NaCl concentrations below $0.5M$ using AMBER-PMF. In this range of salt concentrations, the hard-sphere contributions are negligible and the EXP-MSA PMF is dominated by the screened phosphate–phosphate Coulomb interactions. The resulting optimized structures are displayed in Figs. 6(a) and 6(b) and their structural parameters in Table III.

DNA helices containing dA · dT stretches are known to have unusual properties. For instance,

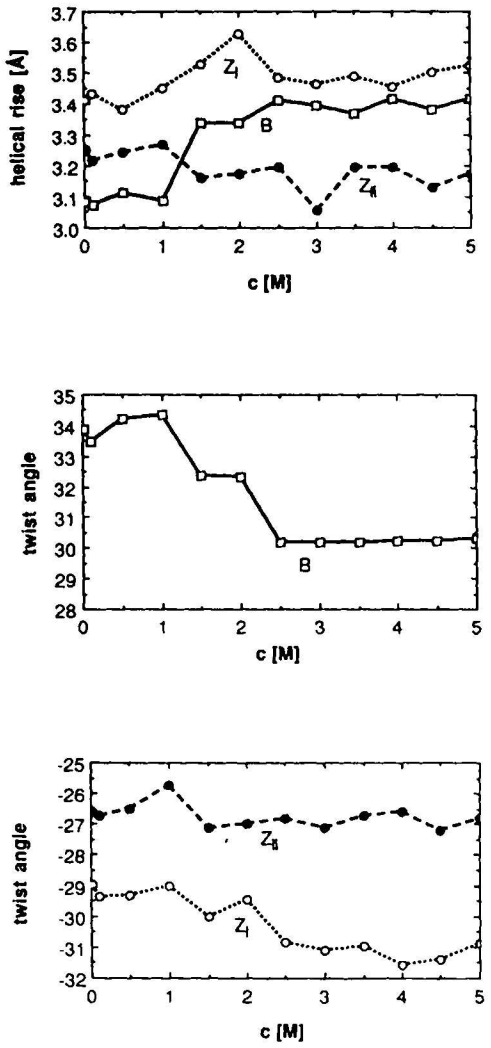

Figure 5. Salt dependence of the helical parameters for the optimized B, Z_I, and Z_{II} conformations. The average rise and twist were calculated from the central 4 base pairs of $d(C-G)_6 \cdot d(C-G)_6$ in order to minimize end effects.

tracts of $(dA \cdot dT)_n$ with $2 \leq n \leq 9$ in phase with the helical repeat have been shown to cause curvature of the helix axis (see Refs. 34 and 39 for reviews). It is widely assumed that such behavior is due to the propensity of this sequence to adopt right-handed conformations different from the classical B form.[40] An example is the heteronomous DNA (H-DNA) form with the sugar pucker C3'-*endo* for the dA strand and C2'-*endo* for the dT strand proposed by Arnott et al.[40] for conditions of low humidity. Vibrational spectroscopy[41] provides evidence for such a mixed sugar pucker.

At high humidities, x-ray fiber diffraction data suggest that the structure of poly(dA) · poly(dT) is only slightly heteronomous with each chain in a B-type conformation,[42,43] in agreement with nmr data.[44,45] Recently two crystal structures have been resolved having longer dA tracts[46,47] with both backbones adopting a C2'-*endo* sugar pucker. However, Raman scattering indicates that even under conditions of high humidity a certain percentage of riboses of the dA strand has a C3'-*endo* sugar pucker.[41,48-50] Immunological data also suggest that the dA strand has a conformation different from that of the B form.[51]

According to our calculations, the H form is found to be energetically more favorable than the B form and its relative stability increases with salt concentration (Table IV and Fig. 7). The lower energy of the H form (relative to B) found here may simply be due to the fact that our treatment does not include the free energy contribution of the water spine,[11] which should considerably stabilize the B form of dA · dT stretches relative to other conformations. However, this contribution, as well as that from vibrational entropy, is essentially salt independent. Thus, in Fig. 7, addition of the hydration and vibrational entropy terms would

Table II Typical Conformational Angles for B, Z_I, and Z_{II} Structures of $d(C-G)_6 \cdot d(C-G)_6$ Optimized at 0.5*M* NaCl and 300 K[a]

Bases	Sugar Pucker (Pseudo-Rotation)		α	β	γ	δ	ϵ	ζ	χ	ν_0	ν_1	ν_2	ν_3	ν_4
B-DNA														
Cytosine	C2'-*endo*	175.4	−73.3	−162.8	57.3	162.5	−178.2	−63.9	253.7	−19.2	42.0	−47.2	37.9	−11.7
Guanine	C3'-*exo*	190.6	−54.5	−69.6	−69.3	167.2	168.5	−97.2	247.2	−6.4	33.0	−45.0	42.5	−22.9
Z_I-DNA														
Cytosine	C2'-*endo*	173.8	−159.6	178.3	69.3	159.9	−74.0	61.7	208.8	−19.7	40.9	−45.8	35.4	−9.8
Guanine	C3'-*endo*	4.3	78.1	−166.9	179.4	90.3	−71.5	−64.2	49.5	10.0	−32.8	41.2	−35.8	16.5
Z_{II}-DNA														
Cytosine	C3'-*exo*	197.0	38.5	170.2	152.4	171.0	−71.6	65.5	216.7	−1.2	29.6	−45.1	45.3	−27.9
Guanine	C3'-*endo*	21.7	72.1	−178.9	−175.7	78.3	176.7	76.4	48.5	−2.1	−23.5	38.1	−40.8	27.2

[a] Definion of the torsion angles according to Ref. 57. All angles in degrees.

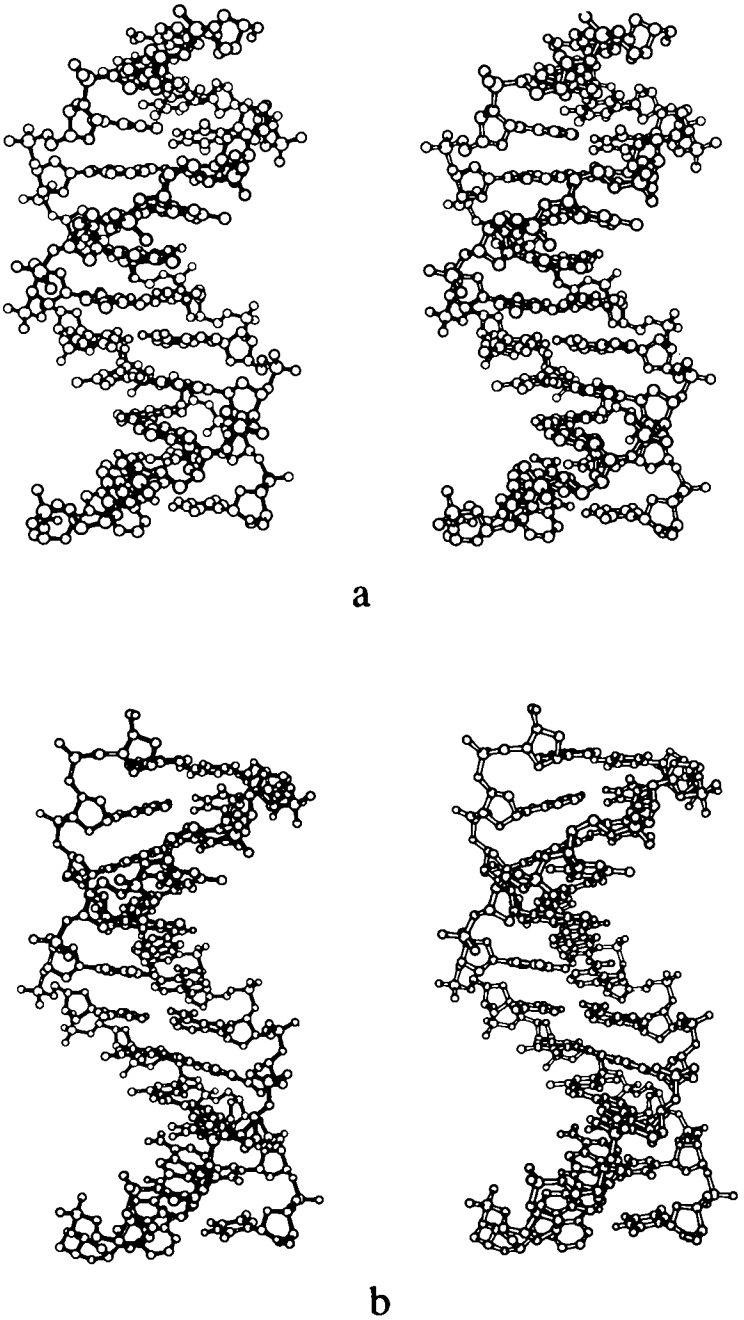

Figure 6. Stereo pair of (a) the optimized B–DNA $dA_{12} \cdot dT_{12}$ structure and (b) the optimized heteronomous DNA $dA_{12} \cdot dT_{12}$ structure.

simply shift the curve upward by a constant amount, implying that salt-induced B–H transitions might occur.

The propeller twist of the optimized B form [Fig. 3(a)] is large (22°). The wedge roll angle[34] equals zero, the wedge tilt angle is −4°, and the base pairs open toward the dAs by 4°. This structure is in fairly good agreement with the dA blocks in the recently resolved crystal structures.[46,47] The bifurcated hydrogen bonds resolved in the latter are also observed in our optimized B form. The greatest difference with respect to the crystallographic data is the large twist angle of 42° (see Table II) compared to 36° found experimentally.[52]

Table III Typical Conformational Angles for B and H Structures of ($dA_{12} \cdot dT_{12}$) Optimized for 0.01*M* NaCl and 300 K[a]

Bases	Sugar Pucker (Pseudo-Rotation)		α	β	γ	δ	ϵ	ζ	χ	ν_0	ν_1	ν_2	ν_3	ν_4
B–DNA														
Adenine	C2′-*endo*	171.9	−65.4	−177.0	55.3	159.0	179.8	−121.4	263.5	−21.7	42.6	−46.4	35.3	−8.6
Thymine	C2′-*endo*	178.1	−64.9	178.8	58.3	163.8	−177.6	−132.4	263.6	−17.0	40.3	−46.9	38.7	−13.6
H–DNA														
Adenine	C2′-*exo*	359.1	−68.0	173.0	56.4	78.8	−171.9	−58.6	199.2	16.3	−40.6	48.2	−39.8	14.8
Thymine	C2′-*endo*	168.1	−66.5	−173.1	55.1	158.4	179.5	−123.0	251.8	−25.5	45.4	−47.4	34.2	−5.5

Helical Parameters of the $dA_{12} \cdot dT_{12}$ Dodecamer[b]

Conformer	Helical Twist	Propeller Twist	Cylinder Tilt	Cylinder Roll	Wedge Roll	Wedge Tilt
B–DNA	42	22	0	5	−4	0
H–DNA	36	22	−8	2	−2	−5

[a]Same definitions as in Table I.
[b]For definitions of the helical parameters used, see Ref. 39.

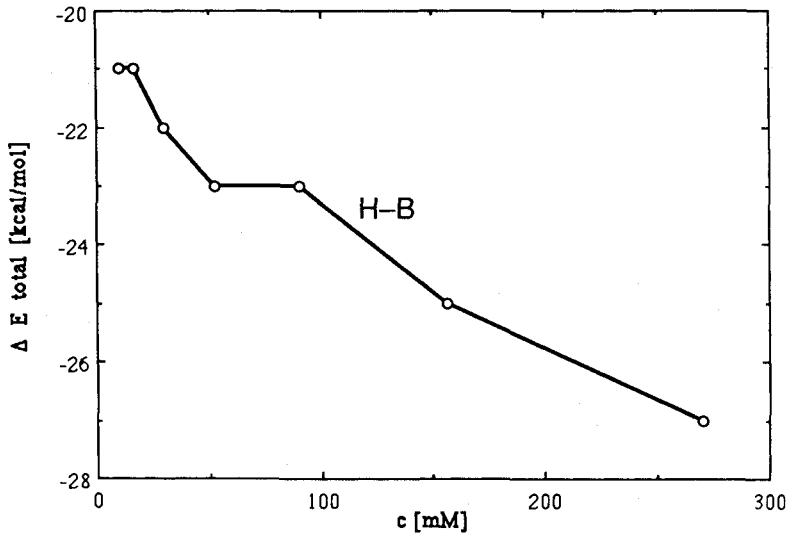

Figure 7. Energy difference of the B- and H-DNA structures. With increasing salt concentration, the H form becomes progressively more stable than the standard B form of DNA.

The tendency to generate larger twist angles seems to be a general problem of the AMBER force field, at least for $dA_n \cdot dT_n$ sequences. Optimization of the $dA_{10} \cdot dT_{10}$ structures after introducing nmr hydrogen distance constraints leads to a structure similar to the optimized B form, that is, with a helical twist angle of 39°.[45]

The optimized H-DNA structure of $dA_{12} \cdot dT_{12}$ [see Fig. 3(b)] has a helical rise and twist angle per base pair of 3.3 Å and 36°, respectively. The sugar pucker of the dA strand is C2′-*exo* with a pseudo-rotation angle of −1° and thus very close to C3′-*endo*. The propeller twist is 22° and the wedge roll angle is −5°. Thus, two adjacent base pairs close toward the minor groove; their base-pair planes make an angle of 5°. The wedge tilt angle is −2°, that is, the planes of two adjacent base pairs open by 2° toward the dA strand. This optimized H-DNA deviates from the crystal structure mainly in the sugar pucker of the dA strand and the negative wedge roll angle, but is in agreement with recent Raman data.[48]

Table IV Energies of the B- and H-DNA (B'-DNA) Structures of $dA_{12} \cdot dT_{12}$ in Kilocalories per Mole at Different Bulk Salt Concentrations

c (mM)	E_T[a]	$E_{\text{bond angle}}$[b]	E_{dihedral}[c]	E_{VDW}[d]	E_{estat}[e]
B-DNA					
10	−1202.	212.4	390.6	−405.0	167.2
17	−1205.	212.3	390.4	−405.0	164.2
30	−1208.	213.5	389.0	−404.1	159.5
52	−1212.	213.8	388.6	−404.0	155.6
90	−1216.	214.5	387.7	−403.6	151.3
156	−1219.	215.8	386.2	402.7	146.1
270	−1223.	216.8	384.3	−401.0	140.1
H-DNA					
10	−1223.	225.4	360.6	−403.6	168.5
17	−1226.	225.4	360.5	−403.5	164.7
30	−1230.	225.6	360.2	−403.4	160.2
52	−1235.	225.8	360.0	−403.1	155.0
90	−1239.	226.9	358.7	−402.1	149.1
156	−1244.	227.3	358.3	−401.8	143.3
270	−1249.	228.2	357.3	−401.8	137.0

[a] Total energy (internal plus PMF terms).
[b] Bond angle energy contribution.
[c] Dihedral angle energy contribution.
[d] van der Waals energy contribution.
[e] Electrostatic energy contribution including PMF terms.

CONCLUDING REMARKS

The AMBER force field, up to now restricted to calculations of internal macromolecular energies, has been extended through inclusion of PMF terms to treat solvent-mediated Coulomb interactions. The resultant AMBER-PMF force field enables one to treat structural stabilities and conformational equilibria of charged biomolecules under realistic conditions with respect to supporting electrolyte effects. In addition, it permits the refinement of structures in varying ionic environments. These features are absent from treatments using distance-dependent dielectric constants and charges effectively reduced according to the condensation hypothesis or compensating (bound) counterions, i.e., the recipes introduced solely for the sake of computational convenience in the standard AMBER and similar force fields. The latter do not derive from a rigorous formulation of the problem of electrostatic interactions in a molecule such as DNA immersed in an aqueous electrolyte. In contrast, as discussed elsewhere,[16,17] the PMF approach is based on a set of consistent approximations. Recent work[18] shows that it yields excellent results not only in the high but also in the low salt regime (0.05–0.5M NaCl), in contrast to traditional approaches for estimating polyionic contributions to DNA structural stability and transitions such as the Manning counterion condensation hypothesis and the Poisson–Boltzmann equation.

As seen from the example of the $B–Z_I$ transition of $d(C–G)_n$ helices (Refs. 16–18 and this work), the PMF contribution alone can describe salt-induced structural transitions provided that the intramolecular energy terms do not come into play due to cancellations. However, we do not expect this to be true in general. For example, in the case of salt-induced B–A transitions we observe a sequence dependence that necessitates the generalized approach presented in this paper (R. Klement, D. M. Soumpasis, and T. M. Jovin, manuscript in preparation). More generally, the influence of sequence, structural hydration, specific ion binding, and chemical modifications will have to be considered in more advanced future versions of the theory.

The generalization of the AMBER force field discussed in this paper permits simultaneous consideration of both intramolecular and ion-mediated charge interactions for arbitrary conformations and varying ionic environments. Thus, a more realistic conformational analysis of charged biomolecules, for which an adequate treatment of solvent-averaged electrostatic interactions is of paramount importance, becomes possible. In addition, one can use the PMF approach to calculate the ionic distri-

butions around such macromolecules in solution (R. Klement, D. M. Soumpasis, and T. M. Jovin, in preparation).

APPENDIX

RPM Ionic PMFs in the EXP-MSA Approximation

Consider a homogeneous system of ionic species $\alpha, \beta, \gamma, \ldots$ in water interacting through the pairwise additive solvent averaged potential $\psi_{\alpha\beta}(r)$

$$\psi_{\alpha\beta}(r) = \psi^0_{\alpha\beta}(r) + \frac{Z_\alpha Z_\beta e^2}{\epsilon r}$$

$$\psi^0_{\alpha\beta} = \begin{cases} \infty & r < \sigma \\ 0 & r > \sigma \end{cases} \qquad \text{(A1)}$$

where Z_α, Z_β are the valencies of ions α, β, separated at distance r, ϵ the bulk dielectric constant of water, and σ an effective average distance of closest approach. $\psi^0_{\alpha\beta}$ approximates the short-range repulsion of two hydrated ions by an effective hard-sphere repulsion, and the second term in Eq. (A1) approximately describes their solvent-screened Coulomb interaction by introducing the measured bulk dielectric constant of water. Specifically, $\psi_{\alpha\beta}$ of Eq. (A1) corresponds to the well-known McMillan–Mayer level model usually called the *restricted primitive model* (RPM).[53] Starting with the RPM description, one of the central goals of the statistical mechanical theory is to evaluate the PMFs $W_{\alpha\beta}(r)$, related to the pair correlation function $g_{\alpha\beta}(r)$ via

$$\frac{W_{\alpha\beta}(r)}{k_B T} = -\ln g_{\alpha\beta}(r) \qquad \text{(A2)}$$

where T is the temperature and k_B is Boltzmann's constant (for definitions and general properties, see Hill[53]).

The PMF above describes the effective interaction of two hydrated ions at a distance r when all others are statistically averaged, or equivalently (up to a distance-independent term), the Helmholtz free energy of the system consisting of the two fixed ions with the rest of the ions freely mobile.

The EXP-MSA treatment[24,25] of a RPM electrolyte approximates the PMFs in the form

$$\frac{W_{\alpha\beta}}{k_B T} = -\ln g^0(r) - Z_\alpha Z_\beta h_D(r) \qquad r \geq \sigma \quad \text{(A3)}$$

where $-\ln g^0$ is the *exact* PMF of the hard-sphere system (all ions uncharged) described by $\psi^0_{\alpha,\beta}$ of Eq. (A1) alone, and all the charge dependence of the PMF is contained in the function $h_D(r)$ evaluated in the framework of the well-known MSA theory of Waisman and Lebowitz[29,30] who, however, obtained only the *Laplace transform* $\hat{G}(p)$ of the related function $G(\tilde{r}) = \tilde{r} h_D(\tilde{r})$, $\tilde{r} = r/\sigma$, namely,

$$\mathcal{L}(G; \tilde{p}) = \hat{G}(p) = -\frac{e^2}{k_B T \epsilon (1+\chi)^2} \times \frac{p}{\left[(p+\chi)^2 + \chi^2\right] - 2\chi^2 \exp(-p)} \qquad \text{(A4)}$$

where $\chi = \frac{1}{2}[(1+2\chi_0)^{\frac{1}{2}} - 1]$, $\chi_0 = \kappa\sigma$, κ is the Debye–Hückel screening parameter, $\kappa = [(4\pi e^2/k_B T\epsilon)\Sigma_\alpha \rho_\alpha Z_\alpha^2)]^{\frac{1}{2}}$, and ρ_α the ionic number densities.

Equation (A4) cannot be analytically inverted to yield a single expression over the whole range of distances, but use of zonal expansion techniques[27,55] in conjunction with results from the theory of Laplace transforms finally leads to the following explicit formula for the range $1 \leq \tilde{r} \leq 4$, for h_D introduced in the main text [Eq. (3)]:

$$h_D(\tilde{r}) = \frac{\lambda}{(1+\chi)^2} \frac{1}{\tilde{r}} (f_1(\tilde{r}) + f_2(\tilde{r}) + f_3(\tilde{r}) \cdots) \qquad \text{(A5)}$$

where

$$\lambda = \frac{e^2}{\epsilon k_B T \sigma} \qquad \text{(A5a)}$$

$$f_1(\tilde{r}) = \exp(-X_1)[\sin X_1 - \cos X_1] \quad \text{for } \tilde{r} > 1 \qquad \text{(A5b)}$$

$$f_2(\tilde{r}) = \exp(-X_2)[(1 - X_2)\sin X_2 - X_2 \cos X_2] \quad \text{for } \tilde{r} > 2 \qquad \text{(A5c)}$$

$$f_3(\tilde{r}) = \tfrac{1}{2}\exp(-X_3)\left[(X_3^2 - 3X_3)\cos X_3 - (X_3^2 + X_3 - 3)\sin X_3\right] \quad \text{for } \tilde{r} > 3 \qquad \text{(A5d)}$$

and

$$X_i = \chi(\tilde{r} - i), \qquad i = 1,2,3 \qquad \text{(A5e)}$$

Additional terms of similar structure and increasing complexity come into play for $\tilde{r} > 4$, but in this range correlations in almost all cases of interest decay very fast. Therefore, $h_D(\tilde{r})$ can be approximated by the asymptotic form

$$h_D(\tilde{r}) = \frac{\lambda}{(1+\chi)^2} \frac{\exp[-\chi(\tilde{r}-1)]}{\tilde{r}} \quad \text{for } \tilde{r} > 4 \quad \text{(A6)}$$

The essentially exact hard-sphere potential of mean force $W_0/k_BT = -\ln g^0$ has been obtained by using Monte Carlo techniques[26] for a wide range of reduced distances $\tilde{r}$ and reduced densities $\tilde{\rho} = \rho\sigma^3$. It is important to note that for a $1:P$ electrolyte, the density ρ to be used for this PMF is $1+P$ times the salt density (or concentration) since dissociation of one salt molecule gives rise to $1+P$ hard spheres.

The Monte Carlo data for the hard-sphere correlation function g_0 can be accurately parameterized as suggested by Verlet and Weis.[56] A FORTRAN program effecting this parameterization has been written by D. Henderson and is reproduced in Ref. 54.

REFERENCES

1. Burkert, N. L. & Allinger, N. L. (1982) *Molecular Mechanics*, American Chemical Society Monograph **177**, Washington, DC.
2. Lifson, S. (1981) in *Structural Molecular Biology, Methods and Applications*, Davis, D. B., Saenger, W. & Danyluk, S. S., Eds., Plenum, New York, pp. 359–385.
3. Levitt, M. (1982) *Cold Spring Harbor Symp. Quant. Biol.*, **47**, 251–262.
4. Momany, F. A., McGuire, R. F., Burgess, A. W. & Scheraga, H. A. (1975) *J. Phys. Chem.* **79**, 2361–2381.
5. Brooks, B. R., Bruccoleri, R. G., Olafson, B. D., Stater, D. J., Jwarinathan, S. V. & Karplus, M. (1983) *J. Comp. Chem.* 4, 187–217.
6. Weiner, P. K. & Kollman, P. A. (1981) *J. Comput. Chem.* **2**, 287–303.
7. Weiner, P. K., Kollman, P. A., Cuse, U. C., Singh, C., Ohio, G., Alogerma, S., Profeta, J. R. & Weiner, P. (1984) *J. Am. Chem. Soc.* **106**, 765–784.
8. v. Kitzing, E. (1986) *Molekülsimulation mit Hilfe von Kraftfeldrechnungen am Beispiel der Aggregation von Nukleinsäuren verschiedener Konformation zu einem Komplex mit Übersetzungsfunktion*, edition herodot, Aachen, Rader Verlag.
9. Go, N. & Scheraga, H. A. (1969) *J. Chem. Phys.* **53**, 4751–4767.
10. Karplus, M. & Kushik, J. N. (1981) *Macromolecules* **14**, 325–332.
11. Drew, H. & Dickerson, R. E. (1981) *J. Mol. Biol.* **151**, 535–556.
12. Kennard, O., Cruse, W. B. T., Nachman, J., Prange, T., Shakked, Z. & Rabinovich, D. (1986) *J. Biomol. Struct. Dynam.* **3**, 623–647.
13. Wang, A. H., Hokashima, T., van der Marel, G., van Boom, J. H. & Rich, A. (1984) *Cell* **37**, 321–331.
14. Westhof, E., Prangé, Th., Chevier, B. & Moras, D. (1985) *Biochemie* **67**, 811–817.
15. Saenger, W., Hunter, W. N. & Kennard, O. (1986) *Nature* **324**, 385–388.
16. Soumpasis, D. M. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5116–5120.
17. Soumpasis, D. M., Wiechen, P. & Jovin, T. M. (1987) *J. Biomol. Struct. Dynam.* **4**, 535–552.
18. Soumpasis, D. M. (1988) *J. Biomol. Struct. Dynam.* **6**, 563–574.
19. Soumpasis, D. M., Robert-Nicoud, M. & Jovin, T. M. (1987) *FEBS Lett.* **213**, 341–344.
20. Pohl, F. M. (1983) *Cold Spring Harbor Symp. Quant. Biol.* **47**, 113–118.
21. Kollman, P. A., Weiner, P. Quigley, G. & Wang, A. (1982) *Biopolymers* **21**, 1945–1969.
22. Tidor, B., Irikura, K. K., Brooks, B. R. & Karplus, M. (1983) *J. Biomol. Struct. Dynam.* **1**, 231–252.
23. Irikura, K. K., Tidor, B., Brooks, B. R. & Karplus, M. (1985) *Science* **229**, 571–572.
24. Anderson, H. C. & Chandler, D. (1972) *J. Chem. Phys.* **57**, 1918–1929.
25. Anderson, H. C., Chandler, D. & Weeks, J. D. (1972) *J. Chem. Phys.* **57**, 2626–2631.
26. Barker, J. A. & Henderson, D. (1971) *Mol. Phys.* **21**, 187–191.
27. Wertheim, M. S. (1963) *Phys. Rev. Lett.* **19**, 321–323.
28. Throop, G. J. & Bearman, R. J. (1965) *J. Chem. Phys.* **42**, 2408–2411.
29. Waisman, E. & Lebowitz, J. L. (1972) *J. Chem. Phys.* **56**, 3086–3092.
30. Waisman, E. & Lebowitz, J. L. (1972) *J. Chem. Phys.* **56**, 3093–3099.
31. Arnott, S. & Hukins, D. W. L. (1972) *Biochem. Biophys. Res. Commun.* **47**, 1504–1509.
32. Wang, A. H., Quigley, G. J., Kolpak, F. J., van der Marel, G., van Boom, J. H. & Rich, A. (1981) *Science* **211**, 171–176.
33. Bremermann, H. (1970) *Math. Biosci.* **9**, 1–15.
34. v. Kitzing, E. & Diekmann, S. (1987) *Eur. Biophys. J.* **15**, 13–26.
35. Pohl, F. M. & Jovin, T. M. (1972) *J. Mol. Biol.* **67**, 375–380.
36. Garcia, A. & Soumpasis, D. M. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 3160–3164.
37. Behe, M. J. (1986) *Biopolymers* **25**, 519–523.
38. Wu, H.-Y. & Behe, M. J. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 7284–7287.
39. Diekmann, S. (1987) in *Nucleic Acids and Molecular Biology*, Vol. 1, Eckstein, F. & Lilley, D., Eds.,

Springer Verlag, New York, pp. 138–156.
40. Arnott, S., Chandrasekaran, R., Hall, I. H. & Puig janer, L. C. (1983) *Nucleic Acids Res.* **11**, 4141–4155.
41. Taillandier, E., Ridoux, J. P., Liquier, J., Leupin, W., Denny, W. A., Wang, Y., Thomas, G. A. & Peticolas, W. L. (1987) *Biochemistry* **26**, 3361–3368.
42. Alexeev, D. G., Lipanov, A. A. & Skuratovskii, I. Y. (1987) *Nature* **325**, 821–823.
43. Park, H.-S., Arnott, S., Chandrasekaran, R. & Millane, R. P. (1987) *J. Mol. Biol.* **197**, 513–523.
44. Sarma, M. H., Gupta, G. & Sarma, R. H. (1985) *J. Biomol. Struct. Dynam.* **2**, 1057–1084.
45. Behling, R. W., Rao, S. N., Kollman, P. & Kearns, D. R. (1987) *Biochemistry* **26**, 4674–4681.
46. Nelson, H. C. M., Finch, J. T., Luisi, B. F. & Klug, A. (1987) *Nature* **330**, 221–226.
47. Coll, M., Frederick, C. A., Wang, A. H.-J. & Rich, A. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 8385–8389.
48. Patapoff, T. W., Thomas, G. A., Wang, Y. & Peticolas, W. L. (1988) *Biopolymers* **27**, 493–507.
49. Wartell, R. M. & Harrell, J. T. (1986) *Biochemistry* **25**, 2664–2671.
50. Jolles, B., Laigle, A., Chinsky, L. & Turpin, P. Y. (1985) *Nucleic Acids Res.* **13**, 2075–2085.
51. Diekmann, S. & Zarling, D. (1987) *Nucleic Acids Res.* **15**, 6063–6074.
52. Peck, L. J. & Wang, J. C. (1981) *Nature* **292**, 375–378.
53. Hill, T. L. (1962) *An Introduction to Statistical Thermodynamics*, Reading, MA, Addison & Wesley.
54. McQuarrie, D. A. (1977) *Statistical Mechanics*, New York, Harper & Row.
55. Henderson, D. & Smith, W. R. (1978) *J. Stat. Phys.* **19**, 191–200.
56. Verlet, L. & Weis, J. J. (1972) *Phys. Rev.* **A5**, 939–952.
57. Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer Verlag, New York, p. 17.

Received October 14, 1988
Accepted April 12, 1989