Accurate Representation of B-DNA Double Helical Structure with Implicit Solvent and Counterions

Lihua Wang,* Brian E. Hingerty,† A. R. Srinivasan,‡ Wilma K. Olson,‡ and Suse Broyde*
*Biology Department, New York University, New York, New York 10003, †Life Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830-6480, and ‡Department of Chemistry, Rutgers, the State University of New Jersey, New Brunswick, New Jersey 08903 USA

ABSTRACT High-resolution nuclear magnetic resonance (NMR) and crystallographic data have been taken to refine the force field used in the torsion angle space nucleic acids molecular mechanics program DUPLEX. The population balance deduced from NMR studies of two carcinogen-modified DNA conformers in equilibrium was used to fine tune a sigmoidal, distance-dependent dielectric function so that reasonable relative energies could be obtained. In addition, the base-pair and backbone geometry from high-resolution crystal structures of the Dickerson–Drew dodecamer was used to re-evaluate the deoxyribose pseudorotation profile and the Lennard–Jones nonbonded energy terms. With a modified dielectric function that assumes a very steep distance-dependent form, a deoxyribose pseudorotation profile with reduced energy barriers between C2′- and C3′-endo minima, and a shift of the Lennard–Jones potential energy minimum to a distance ~0.4 Å greater than the sum of the van der Waals' radii, the sequence-dependent conformational features of the Dickerson–Drew dodecamer in both the solid state and the aqueous liquid crystalline phase are well reproduced. The robust performance of the revised force field, in conjunction with its efficiency through implicit treatment of solvent and counterions, provides a valuable tool for elucidating conformations and structure–function relationships of DNA, including those of molecules modified by carcinogens and other ligands.

INTRODUCTION

Computer-aided modeling of DNA, RNA, and protein structures has become a partner with experiment in the elucidation of molecular structure–function relationships. Molecular mechanics and dynamics techniques are routinely used to generate feasible structures in solution, both in the refinement of high-resolution nuclear magnetic resonance (NMR) data and for de novo predictions of conformation and interactions. Advances in computer architecture and in force-field development support the increasing utility of these methods, with elaborate simulations of biomolecular dynamics (with full-scale solvation and explicit treatment of electrostatic interactions) now performed beyond the nanosecond range and free energies derived from these simulations in reasonable accord with experiment (Jayaram et al., 1999; Beveridge and McConnell, 2000; Kollman et al., 2000; Cheatham and Young, 2001; Yan et al., 2001). Nonetheless, the "multiple minimum" or "sampling" problem remains unsolved, precluding identification of both the lowest energy state and all energetically accessible conformations. Furthermore, simulations of molecular dynamics incorporating full-scale solvation are still too costly to sample the full conformational landscape.

Continuum electrostatic models address the issue of computational cost involved in explicit molecular calculations by greatly reducing the number of degrees of freedom of the solvent (reviewed in Cramer and Truhlar, 1999). There are a number of possible strategies for treating the electrostatic component of the solvation free energy, including numerical solution of the Poisson–Boltzmann equation (Sharp and Honig, 1990; Honig and Nicholls, 1995), implementation of the generalized Born approximation (Still et al., 1990; Cramer and Truhlar, 1992; Hawkins et al., 1995, 1996; Tsui and Case, 2001), and use of distance-dependent dielectric functions (reviewed in Mehler, 1996). These methods can be used in conjunction with other molecular data, such as the solvent-accessible surface area (Lee and Richards, 1971), to evaluate nonelectrostatic energy contributions (Hermann, 1972; Still et al., 1990; Cramer and Truhlar, 1992). Solution of the nonlinear Poisson–Boltzmann equation, although very accurate, is computationally expensive and not yet well suited for routine conformational searches (Baker et al., 2001). The generalized Born method, by contrast, is computationally efficient and benchmarked to reproduce the solvation free energies of small organic molecules (Best et al., 1997; Jayaram et al., 1998). Moreover, recent parameterizations and modifications of the generalized Born approximation appear to have corrected inaccuracies in the treatment of macromolecules (Srinivasan et al., 1999; Tsui and Case, 2000, 2001), although further tests will be of interest.

Distance-dependent dielectric functions, based on Lorentz–Debye–Sack theory (Sack, 1926, 1927; Debye, 1929; Lorentz, 1952), have a long history in the treatment of molecular interactions and have been widely used, because of their efficiency, to account implicitly for solvent effects in molecular mechanics calculations. According to these schemes, the dielectric constant between charged atoms

Submitted January 18, 2002 and accepted for publication March 18, 2002.
Address reprint requests to Suse Broyde, Biology Department, New York Univ., New York, NY 10003. Tel: 212-998-8231; Fax: 212-995-4015; E-mail: broyde@nyu.edu.

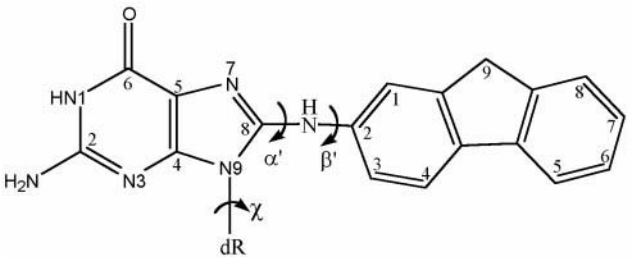
0006-3495/02/07/382/25 $2.00

increases with their distance of separation, tapering off at large interatomic displacements to the value of bulk solvent and thereby mimicking the expected diminution of electrostatic interactions as solvent molecules interpenetrate between the atom pairs. There is uncertainty, however, in the choice of a suitable functional form for the dielectric constant, particularly for the different environments of a macromolecule, such as the major and minor grooves of B-DNA. For example, a preferential build-up of counterions or co-ions around the double helix can change the dielectric medium and the distance-dependent interactions between charged atoms on the edges of the grooves (Young et al., 1998). The selection of a suitable dielectric function is thus key to the performance of a nucleic acid force field. The functional form introduced by Hingerty et al. (1985) is reasonably robust when compared against structural predictions based on more rigorous and expensive Poisson–Boltzmann calculations (Daggett and Kollman, 1990; Friedman and Honig, 1992, 1995).

A number of effective computational strategies for addressing the multiple minimum problem in the study of DNA and carcinogen-damaged DNA have been developed with the molecular mechanics program DUPLEX (Hingerty et al., 1989; Broyde and Hingerty, 1999). These include large-scale conformational searches for energetically favorable orientations of carcinogen with respect to DNA, progressive build-up of DNA structure from smaller mono- and oligonucleotide subunits, and restriction of nucleic acid conformational flexibility to the six exocyclic chain torsion angles and the puckering of the sugar rings. Bond lengths and bond angles are held fixed at standard equilibrium values and aromatic moieties are assigned planar conformations unless found otherwise by experiment. A sigmoidal, distance-dependent dielectric function (Hingerty et al., 1985) is used to treat solvation, and reduced partial charges are introduced on the pendant phosphate oxygens (Srinivasan and Olson, 1980) to mimic the effects of counterion condensation (Manning, 1978). Partial charges are taken from Ornstein and Rein (1979) with modifications noted in Hingerty et al. (1989). DUPLEX uses the linked atom algorithm (Scott and Scheraga, 1965) for coordinate generation from the bond lengths, bond angles, and dihedral angles, and the Powell algorithm (Powell, 1964) for energy minimization. The approach has proved useful both in prediction (Singh et al., 1991) and in refinement of DNA solution structures (Geacintov et al., 1997; Patel et al., 1998).

The NMR solution structures of DNA modified covalently at the C8 atom of guanine by the aromatic amine, 2-aminofluorene (AF) (Fig. 1 *A*), (Mao et al., 1998a,b) provide a useful starting point for reconsideration of the treatment of solvent and other aspects of the DUPLEX force field. Specifically, NMR studies of the AF-bound dodecamer duplex, d(CTCGG*CGCCATC)·d(GATGGCGCCGAG) (Sequence I, Fig. 1 *B*) with AF modification at G*, reveal a predominant, well-

A. Structure of AF-dG adduct



B. Sequence I

C1	T2	C3	G4	G5*	C6	G7	C8	C9	A10	T11	C12
G24	A23	G22	C21	C20	G19	C18	G17	G16	T15	A14	G13

C. Sequence II

C1	C2	A3	T4	C5	G6*	C7	T8	A9	C10	C11
G22	G21	T20	A19	G18	C17	G16	A15	T14	G13	G12

FIGURE 1 (*A*) Chemical structure of the AF-dG adduct (G*) showing the torsion angles (χ, α', β') that govern AF-DNA conformation. Angles are defined as follows: χ, O4′–C1′–N9–C4; α', N9–C8–N–C2; β': C8–N–C2–C1, where the C1, C2, and N are from 2-aminofluorene and the other atoms are from deoxyguanosine. (*B*, *C*) Sequences of high-resolution NMR structures, I (Mao et al.; 1998a) and II (Mao et al., 1998b), used to benchmark AF orientation in energy calculations involving the boxed 9-base pair sequences.

characterized conformer that places the AF in the B-DNA major groove while keeping all Watson–Crick base pairs (bp) intact. This conformer, present to the extent of ~90% of the population, is in equilibrium with another conformer that cannot be fully characterized with NMR distance restraints. Qualitative NMR data, however, indicate that the DNA in the second form adopts a structure with the AF inserted into the helix and the modified guanine displaced into the major groove. The latter conformer has been fully characterized in a different 11-bp duplex, d(CCATCG*CTACC)·d(GGTAGCGATGG) (Sequence II, Fig. 1 *C*) with AF modification again at G*. In this sequence context, the base-displaced, AF-intercalated conformer comprises ~70% of the conformational mixture. Moreover, the NMR data indicate that the base-displaced AF-intercalated states are essentially the same in the two sequence contexts. In both DNA adducts, there are two rapidly interconverting rotamers of comparable flexibility, with the C9-containing edge of the external, groove-bound AF directed toward either the 5′- or the 3′-side of the modified base and the C9-containing edge of the intercalated, base-displaced AF directed toward either the major groove or the minor groove of the double helix. The ~9:1 conformer ratio in Sequence I is indicative of a free energy difference of roughly 1.3 kcal/mol between the two states at 298 K, assuming equal conformational flexibilities. It should be noted that NMR structures within the bounds of the data represent ensemble aver-

ages rather than single points. This system serves as an excellent test case for the implicit treatment of solvent around DNA because the solvent exposure of carcinogen is so different in the two conformational states. We have thus evaluated and refined different distance-dependent dielectric functions for representative conformers of AF-bound DNA and further tested the capabilities of the expressions to reproduce various sequence-dependent structural features of unmodified B-DNA.

The right-handed, self-complementary B-DNA sequence, d(CGCGAATTCGCG)$_2$, known as the Dickerson–Drew dodecamer (Wing et al., 1980; Drew et al., 1981), has been crystallized under various conditions, and the many high-resolution crystal structures (Berger et al., 1998; Egli et al., 1998; Shui et al., 1998a,b; Tereshko et al., 1999a,b; Sines et al., 2000; Woods et al., 2000) now at hand in the Nucleic Acid Database (NDB; Berman et al., 1992) can be compared with DUPLEX-generated B-DNA structures based on different force fields. Among the most interesting aspects of the dodecamer structures are the subtle irregularities of the double-helix associated with the sequential positions of purine and pyrimidine bases along the complementary strands. This fine-scale information is not found in the idealized, regularly repeating B-DNA helical structure that can be derived from X-ray fiber diffraction measurements (Arnott et al., 1976). The sequence-dependent variation of base-pair geometry, backbone torsion angles, and groove dimensions computed on the basis of the DUPLEX force field is compared below against the corresponding conformational features observed in high-resolution Dickerson–Drew dodecamer structures. To improve the computational predictions, we reconsider the treatment of deoxyribose pseudorotation and the long-accepted strategy of minimizing the Lennard–Jones potential energy at a distance 0.2 Å greater than the van der Waals' separation (Brant et al., 1967; Olson and Flory, 1972; Srinivasan and Olson, 1980), in addition to modifying the sigmoidal, distance-dependent dielectric function.

Thus, we use the relative energies of AF-modified DNA duplexes in equilibrium between AF-external and AF-intercalated base-displaced conformers, and the sequence-dependent structural information of the Dickerson–Drew dodecamer as criteria for assessing the viability of DNA computational predictions based on the efficient and inexpensive treatment of solvation and counterions within the DUPLEX force field. A dielectric functional form that tapers more steeply than earlier functions to the dielectric constant of bulk water is found to reproduce fairly well the expected energy difference between the two conformers of AF-bound DNA. This function, together with a modified deoxyribose pseudorotation profile with reduced energy barriers between C2′- and C3′-endo minima, and a Lennard–Jones potential energy minimum at a distance 0.38 Å greater than the normal van der Waals' separation, accounts satisfactorily for the sequence-dependent structural features of Dickerson–Drew dodecamer crystals. The improved force field shows even better correspondence with the dodecamer structure of the same sequence deduced from NMR measurements in an aqueous dilute liquid crystalline phase (Tjandra et al., 2000). The robust performance of the revised force field, in conjunction with its efficiency through implicit treatment of solvent and counterions, provides a valuable tool for elucidating structure–function relationships of DNA, including those of chains modified by carcinogens and other ligands.

TABLE 1 Key torsion angles in high-resolution NMR solution structures and in remodeled structures of the AF-modified DNA

	(°)					
	NMR Structure*			Remodeled Structure		
Conformation	χ	α'	β'	χ	α'	β'
AF-external	−137	−160	38	−130	−166	35
AF-intercalated	65	−148	138	58	−143	134

*Mao et al., 1998a,b.

METHODS

Computer simulation of the AF-modified dodecamer, Sequence I (Fig. 1 *B*), was performed using the NMR solution structure of the major AF-external conformer as a starting model (Mao et al., 1998a) (see Table 1). Only one of the two rotamers, the form with the C9-containing edge of AF directed toward the 5′ side of the modified base ($\alpha' = -160°$, $\beta' = 38°$), was considered. The 12-mer was truncated to a 9-mer so that the modified G occupied the central position of the double helix (boxed segment in Fig. 1 *B*). Past experience with energy minimization of carcinogen-modified DNA (Shapiro et al., 1998) indicates that 9-mers are long enough to include all DNA-carcinogen interactions and concomitantly avoid, as much as possible, DNA end effects unrelated to adduct conformation. The secondary base-displaced, AF-intercalated conformer of the same sequence was modeled using the principal conformation detected in high-resolution NMR studies of the AF-modified Sequence II 11-mer (Mao et al., 1998b) as a starting state. The ($\alpha' = -148°$, $\beta' = 138°$) rotamer with the C9-containing edge of AF directed toward the minor groove was utilized. The 11-mer was then shortened by one residue at either end to create a 9-mer with a centrally modified G, and the sequence adjusted to match that of Sequence I (see Fig. 1, *B* and *C*). Subsequent minimization of the two 9-mers yielded structures very similar to the respective NMR parent structures with the key χ, α', β' torsion angles (Fig. 1 *A*) which govern the AF-DNA conformations departing no more than 7° from the starting states (Table 1).

We chose to model the minor (10%) AF-intercalated base-displaced structure of Sequence I rather than the secondary (30%) AF-external conformation of Sequence II because the latter form of the Sequence II complex is not as well defined by NMR distance restraints as the former arrangement of the Sequence I adduct. In addition, Sequence I is biologically more interesting in that it contains the *Nar*I mutagenic hotspot for aromatic amine carcinogens (reviewed in Hoffmann and Fuchs, 1997). The (χ, α', β') rotamer choices noted above correspond to the lower energy (and better characterized) states found in the refinement of the NMR structures (Mao et al., 1998a,b). In the first tests

TABLE 2 High-resolution dodecamer crystals used to benchmark calculations

NDB ID	Features	Modifier*	Resolution Å	Reference
BDL001	B-DNA double helix	–	1.9–8.0	(Drew et al., 1981)
BDL005	B-DNA double helix	–	1.9–8.0	(Holbrook et al., 1985)
BDL020	B-DNA double helix	–	1.9–5.0	(Westhof, 1987)
BD0005	B-DNA double helix	–	1.5–10.0	(Shui et al., 1998b)
BD0029	B-DNA double helix	–	1.8–10.0	(Woods et al., 2000)
BD0041	B-DNA double helix	–	1.2–10.0	(Sines et al., 2000)
BDL084	B-DNA double helix	–	1.4–10.0	(Shui et al., 1998a)
BD0007	B-DNA double helix, modified	TAF	1.1–20.0	(Tereshko et al., 1999a)
BD0012	B-DNA double helix, modified	TAF	1.2–20.0	(Tereshko et al., 1999b)
BD0013	B-DNA double helix, modified	TAF	1.5–20.0	(Tereshko et al., 1999b)
BD0030	B-DNA double helix, modified	TAF	0.9–20.0	(Egli et al., 1998)
BDLB84	B-DNA double helix, modified	TAF	1.6–8.0	(Berger et al., 1998)
BDLB85	B-DNA double helix, modified	TAF	1.6–8.0	(Berger et al., 1998)

*Modified B-DNA crystals have one or two 2′-deoxy-2′-fluoroarabino-thymidines (TAF) in each chain.

of various dielectric functions, the energies of AF-external and AF-intercalated, base-displaced 9-mer duplexes were computed without further energy minimization, to avoid additional structural variations associated with the multiple minimum problem. Specifically, the comparison of identical structures for each tested dielectric function provides direct information on how the steepness of the dielectric function affects the computed conformational energy difference. This strategy also closely preserves the structures of the NMR-characterized states.

Crystal and DUPLEX-generated Dickerson–Drew dodecamer structures were analyzed at the local base-pair level using the RNA (*R*un *N*ucleic *A*cid) software package (Babcock et al., 1993) and at the level of torsion angles and groove geometry with the routines of Curves 5.2 (Lavery and Sklenar, 1988, 1989). Thirteen crystal structures (Table 2) from the Nucleic Acid Database (Berman et al., 1992) with resolution ranging from 0.9 to 1.9 Å and without any mismatches, flipped-out bases, bound drug molecules, or bulky modifiers were chosen to benchmark the DUPLEX-generated structures. The selected crystals fall into three groups. The first group includes three independent refinements of a structure determined in the 1980s (Drew et al., 1981; Holbrook et al., 1985; Westhof, 1987) with a resolution of ~1.9 Å. The second group contains four structures from the Williams laboratory solved in 1998–2000 (Shui et al., 1998a,b; Sines et al., 2000; Woods et al., 2000), with resolution in the range of 1.2–1.8 Å. The third group is composed of six structures from the Egli laboratory solved in 1998–1999 (Berger et al., 1998; Egli et al., 1998; Tereshko et al., 1999a,b), with resolution in the range of 0.9–1.6 Å, and with one or two 2′-deoxy-2′-fluoroarabino-thymidines in place of thymine in each strand. The chemically modified residues in the latter molecules do not affect the overall duplex structure, but do enhance the thermodynamic stability of B-form DNA (Rosenberg et al., 1993; Ikeda et al., 1998). The unprecedented high resolution of this set of crystal structures makes them particularly attractive as benchmarks for the DUPLEX force field.

In attempting to reproduce the sequence-dependent structural features observed in dodecamer crystals, we also considered four different modifications of the deoxyribose pseudorotation profile: 1) the current list of variables—puckering amplitude τ_m, internal bond angles, and energies as a function of the pseudorotation phase angle P—used in DUPLEX (Sasisekharan, 1973; Hingerty and Broyde, 1982); 2) the current input from DUPLEX with energies reduced by half; 3) the computational treatment originally offered by Sato (Sasisekharan, 1973; Sato, 1983); and 4) the Sato treatment with energies reduced by half. Profiles 1 and 4 are shown in Fig. 2 (*gray* and *red lines*, respectively). It should be noted that the bases are replaced by a hydrogen in the Sato treatment, so that the conventional high barrier in the western region, i.e., $P \approx 270°$, (Saenger, 1984) of the pseudorotation cycle is not seen.

In the final stage of force-field improvements, we varied the location of the Lennard–Jones potential energy minimum over a range of distances 0.2–0.4 Å greater than the sum of the van der Waals' radii, first adding 0.2, 0.3, or 0.4 Å to the normal van der Waals' separation and then investigating finer (0.02-Å) increments of added distance between 0.3–0.4 Å.

RESULTS

Search for a dielectric functional form that reasonably reproduces the energetics of AF-modified DNA conformers

Our approach to the enhancement of the DUPLEX force field entailed an iterative procedure. We first tested the functional form of the distance-dependent dielectric constant with a pair of AF-modified DNA 9-mers derived from NMR solution studies (see Methods).

The original distance-dependent dielectric function of Hingerty et al. (1985) was recast to the analytical form

$$\varepsilon(r) = 78.3 - 77.3(r/2c)^e/(\sinh(r/2c))^e. \qquad (1)$$

Here $\varepsilon(r)$ is the dielectric constant when the two atoms of interest are separated by a distance r, whereas c and e are variable quantities that are used to adjust the expression to virtually any sigmoidal form. Values of 2.5 for c and 2.0 for e produce the function used previously (Hingerty et al., 1985). The representative functional forms in Fig. 3 illustrate how $\varepsilon(r)$ becomes steeper as c decreases and/or e increases.

The observed ~9:1 ratio between AF-external and base-displaced, AF-intercalated conformers of Sequence I corresponds to an expected energy difference of roughly 1.3 kcal/mol at 298 K, assuming comparable flexibility in the two conformational forms. The energies of representative NMR-derived states were computed with DUPLEX without minimization over the ranges of c and e shown in Table 3. The computed energy differences between the two conformers, in kcal/mol, are reported in the table for selected values of the two constants. We note that, as the steepness of the function increases (low c, high e), the energy difference between the two conformers becomes smaller. The original

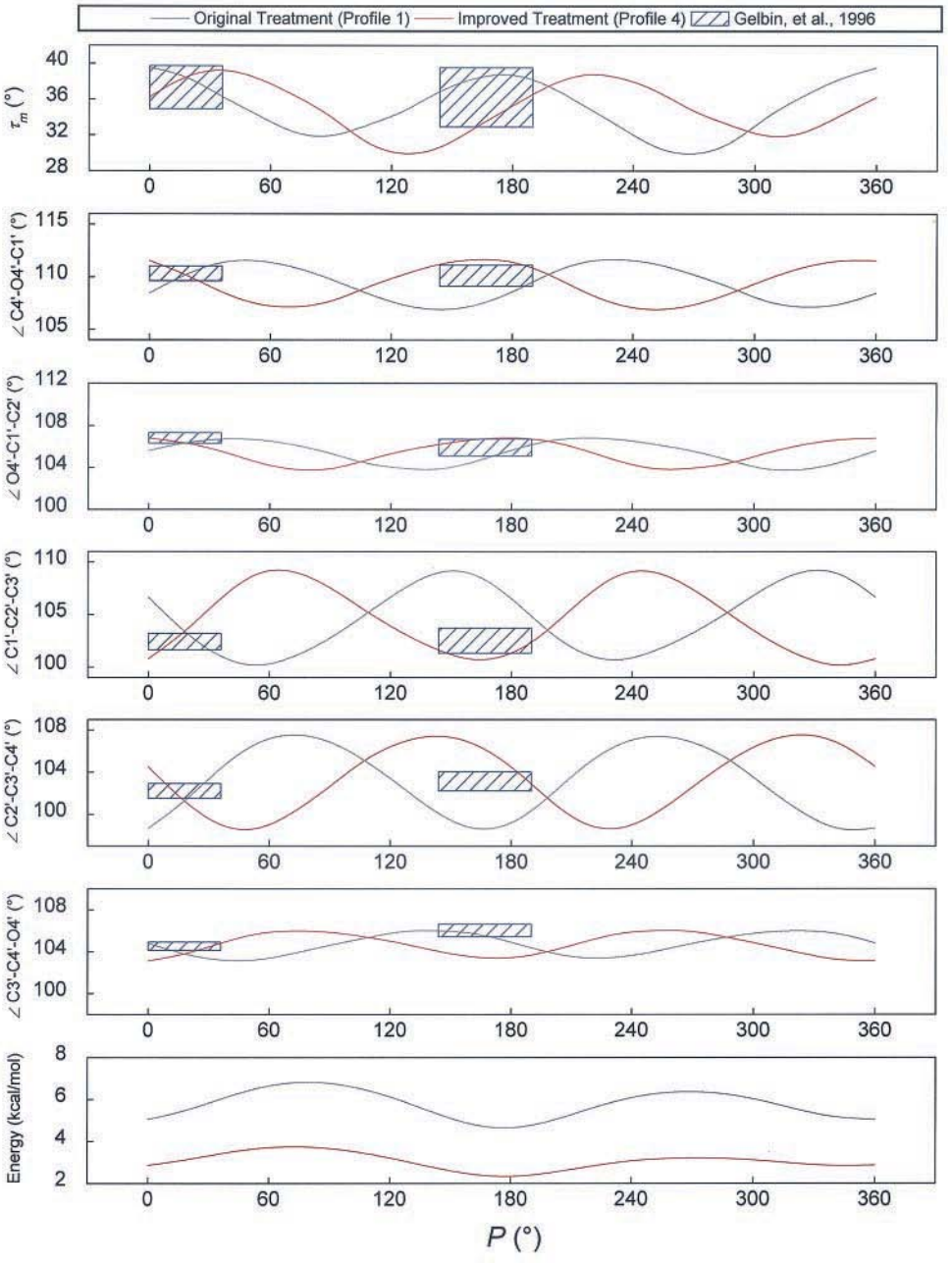


FIGURE 2 Deoxyribose pseudorotation profiles of puckering amplitude τ_m, internal bond angles, and conformational energy of the sugar ring as a function of the pseudorotation phase angle, *P*. Note that the base is replaced by a hydrogen in the profile (Sasisekharan, 1973), and a *gauche* energy term (Olson, 1982) is included to account for the known preference of C2′-endo ($P \approx 180°$) over C3′-endo ($P \approx 0°$) puckered forms. Color coding: *gray*, original DUPLEX pseudorotation treatment, profile 1 (Hingerty and Broyde, 1982); *red*, current improved treatment, profile 4. Shaded areas correspond to the observed averages and standard deviations of geometric parameters in high-resolution crystal structures of mononucleosides and mononucleotides (Gelbin et al., 1996).

Debye treatment (Hingerty et al., 1985) with $c = 2.5$ and $e = 2.0$ (Fig. 3, *curve C*) gives an energy difference of 5.3 kcal/mol, over-favoring the AF-external conformation compared to the AF-intercalated, base-displaced state. We chose $c = 0.5$, $e = 6.0$ for further study because that combination most closely approximates the expected energy difference of ~1.3 kcal/mol (Table 3), and further steepening in the functional form ($c = 0.5$, $e = 8.0$) does not improve the energetics. As shown below, this function performs quite well in reproducing the sequence-dependent, base-pair and backbone conformational parameters observed in crystals of the Dickerson–Drew dodecamer.

TABLE 3 Energy differences, ΔE (kcal/mol), between AF-external and AF-intercalated base-displaced DNA conformers for selected values of *c* and *e* in the distance-dependent dielectric function Eq. 1

	c						
e	0.5	2.5	4.5	6.5	8.5	10.5	12.5
0.25	3.8	8.2	7.5	6.9	6.9	7.3	8.0
0.5	3.3	7.8	8.1	7.5	7.0	6.8	6.9
1.0	8.4	6.6	8.1	8.1	7.6	7.2	6.9
2.0	4.3	5.3	7.5	8.2	8.1	7.9	7.5
4.0	3.6	4.2	6.2	7.5	8.1	8.2	8.1
6.0	2.7	3.8	5.5	6.9	7.7	8.1	8.2
8.0	2.7	3.5	5.3	6.7	7.3	7.9	8.1

Optimization of the deoxyribose energy and bond angle profiles to account better for the sequence-dependent structural parameters in crystal structures of the Dickerson–Drew dodecamer

In the next stage, we tested the revised force field with the new dielectric function by evaluating its performance in reproducing sequence-dependent base and backbone parameters observed in crystals of the Dickerson–Drew dodecamer. We found that the sequence-dependent pattern of most base-pair step parameters was reproduced reasonably well (data not shown). The DUPLEX-generated structure, however, was over-twisted, with an average dimer twist of 38.7 ± 1.6°, compared to an average value of 35.5 ± 4.4° in the crystal structures. In addition, the base-pair steps showed excessive negative roll into the minor groove, so that the minor groove width was much narrower on average (2.6 ± 1.5 Å) than that in the crystal structures (4.2 ± 1.2 Å), although not as narrow as with the original force field (Table 4). Further adjustments of the dielectric function did not improve these structural anomalies. Knowing that high twist is linked to a high phase angle of pseudorotation and negative roll (Olson and Zhurkin, 2000), we revisited the treatment of the pseudorotation phase angle in DUPLEX. The program previously used a modification of the deoxyribose pseudorotation potential (Hingerty and Broyde, 1982) devised by Sato (1983) and Sasisekharan (1973), in combination with a *gauche* energy term that is needed to reproduce the known preference of C2′-endo over C3′-endo puckered forms in deoxyribose (Olson, 1982). A test of the four deoxyribose pseudorotation profiles described in the Methods section showed that neither Sato's original deoxyribose treatment (Case 3 in the preceding section) nor the profile produced by halving the energy in the DUPLEX pseudorotation profile (Case 2 above) improved the overtwisting of the helix and the narrowness of the minor groove. A combination of the two (Case 4 above), however, satisfactorily produces a B-like minor groove and a reasonable sequence-dependent twisting pattern, without adverse effects on other measures of base-pair geometry, or the torsional parameters (data not shown). The change in barrier height follows from arguments developed to account for the

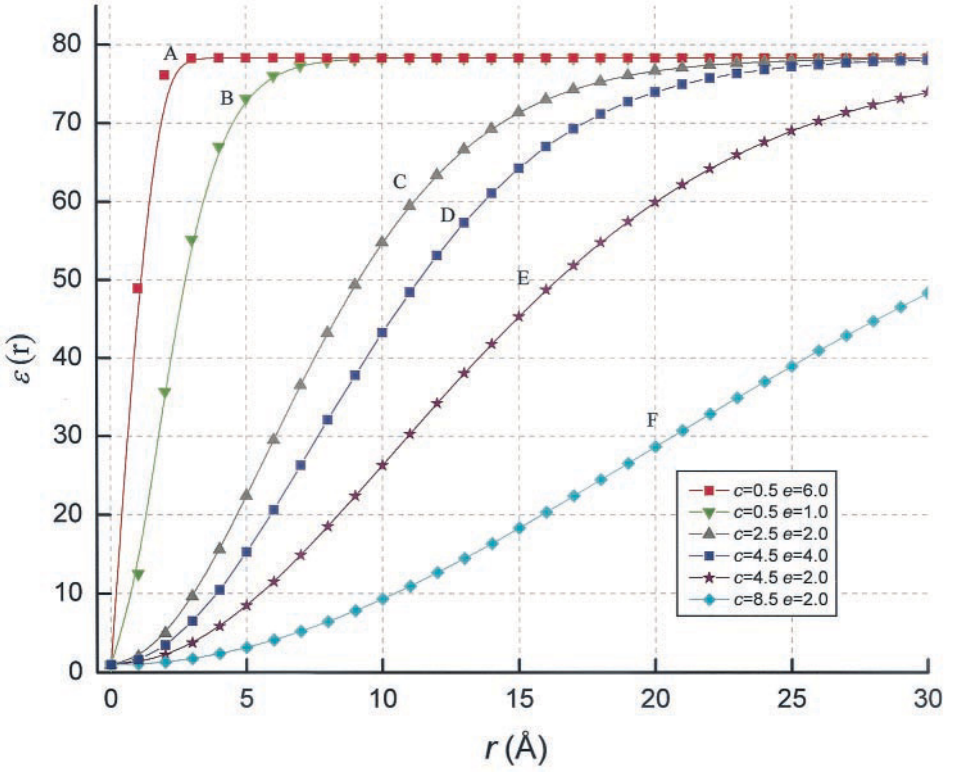


FIGURE 3 Representative functional forms of the distance-dependent dielectric constant, $\varepsilon(r)$, produced with selected values of the variables *c* and *e* in Eq. 1. The steepness of the function increases as *c* decreases and/or *e* increases. The original dielectric functional form (Hingerty et al., 1985) used by DUPLEX is shown in gray (*Curve C*); the functional form fitted to molecular dynamics simulations of Young et al. (1998) is shown in green (*Curve B*). The steep function used in current work is shown in red (*Curve A*). Curves D–F illustrate the change in $\varepsilon(r)$ as *c* and *e* change.

TABLE 4 Average values of selected base-pair, torsional, and groove parameters for d(CGCGAATTCGCG)$_2$ dodecamer crystal structures and DUPLEX-generated B-DNA structures of the same sequence, based on original and improved force fields (FF)

Parameters	Crystal*†	Original FF*	Improved FF*
Complementary base-pair parameters‡			
Buckle (°)	1.8 ± 6.3	0.5 ± 7.0	1.1 ± 5.5
Propeller (°)	−12.5 ± 5.7	−4.1 ± 5.0	−8.8 ± 7.6
Opening (°)	6.5 ± 3.0	7.0 ± 3.0	4.7 ± 1.2
Base-pair step parameters‡			
Tilt (°)	−0.1 ± 2.5	−0.4 ± 2.4	−0.4 ± 1.7
Roll (°)	−0.1 ± 6.1	−9.7 ± 5.1	1.1 ± 2.5
Twist (°)	35.5 ± 4.4	38.7 ± 1.3	35.9 ± 1.4
Shift (Å)	0.0 ± 0.5	0.0 ± 0.3	0.0 ± 0.5
Slide (Å)	0.1 ± 0.5	−0.7 ± 0.3	−0.3 ± 0.4
Rise (Å)	3.3 ± 0.1	3.2 ± 0.2	3.2 ± 0.1
Local helical parameters‡			
Inclination (°)	0.6 ± 10.0	−13.8 ± 7.9	1.9 ± 4.1
Tip (°)	0.0 ± 4.1	0.6 ± 3.5	0.6 ± 2.7
Helical twist (°)	36.2 ± 4.4	40.2 ± 1.8	36.1 ± 1.4
x-displacement (Å)	0.1 ± 1.0	0.2 ± 0.7	−0.6 ± 0.7
y-displacement (Å)	0.0 ± 0.8	0.0 ± 0.4	−0.1 ± 0.7
Helical rise (Å)	3.3 ± 0.1	3.2 ± 0.2	3.2 ± 0.1
Backbone torsion angles§			
α (°)	−62.8 ± 13.3	−60.2 ± 5.7	−63.6 ± 6.6
β (°)	169.7 ± 13.4	179.4 ± 8.9	176.0 ± 7.5
γ (°)	62.1 ± 43.2	55.1 ± 4.8	51.4 ± 7.2
δ (°)	123.8 ± 18.7	144.4 ± 4.9	129.7 ± 13.8
ε (°)	−169.0 ± 21.1	−179.1 ± 4.6	−171.7 ± 3.7
ξ (°)	−107.5 ± 32.0	−115.2 ± 7.3	−103.8 ± 10.1
Sugar-base conformation§			
P (°)	130.0 ± 38.0	157.0 ± 8.2	135.7 ± 24.6
χ (°)	−113.4 ± 14.5	−120.5 ± 6.8	−121.6 ± 8.8
Groove dimensions§			
Minor-groove width (Å)	4.2 ± 1.2	1.2 ± 0.6	6.2 ± 0.6
Minor-groove depth (Å)	5.3 ± 0.4	5.3 ± 0.2	4.8 ± 0.1
Major-groove width (Å)	11.5 ± 0.8	12.6 ± 0.9	9.6 ± 0.6
Major-groove depth (Å)	4.7 ± 1.0	3.6 ± 0.8	6.2 ± 0.3

*Standard deviations follow the ± symbol.

†The conformational parameters in the multiple structures (*blue circles* in Fig. 4) were averaged at each base-pair/dimer step. Then the geometry of an average base-pair/dimer was obtained by averaging over all the base-pairs/steps in the dodecamer, including the ends. The standard deviations are the deviations over the averages at each base-pair/step.

‡Values computed with RNA (Babcock et al., 1993).

§Values computed with Curves 5.2 (Lavery and Sklenar, 1988, 1989).

undertwisting of DNA and the intermediate sugar puckering phase angles found in many computer simulations of B-DNA structures (Olson and Zhurkin, 2000). In the present work, the barrier is reduced to lower the twist, whereas the barrier must be raised in simulations based on other force fields, in which the computed structures are undertwisted.

Relocating the Lennard–Jones potential energy minimum to a distance 0.38 Å greater than the van der Waals' separation reasonably mimics the base-pair rise observed in crystal structures

At this stage, it was noted that the magnitude of base-pair rise and helical rise, nonsequence-dependent but important parameters for characterizing double helical structures (Sponer and Kypr, 1993; Hunter and Lu, 1997), were consistently smaller in the computer-generated dodecamer than the average values in the crystal structures. Consequently, the simulated dodecamer was visibly shorter than its crystal counterparts, with an average helical rise of 3.0 ± 0.0 Å compared to a value of 3.3 ± 0.1 Å in the crystal structures. In fact, the helical rise had diminished, following improvement of the dielectric function, from a value of 3.2 ± 0.2 Å obtained with the original force field (Table 4). The reduction in electrostatic interactions with the new dielectric function collapsed the DNA locally, allowing neighboring base pairs to approach more closely than in the original force field and in crystals. Changes in either the dielectric function or the deoxyribose pseudorotation profile did not alter either the local base pair rise or the helical rise significantly. Knowing that a large number of van der Waals' attractions can bring interacting moieties into closer contact than isolated atom pairs (Brant et al., 1967; Olson and Srinivasan, 1990), we investigated improving the rise values

by relocating the Lennard–Jones potential energy minimum. Specifically, we reevaluated whether the Lennard–Jones energy minimum should remain at the distance 0.2 Å greater than the van der Waals' separation used in the first empirical energy studies of polypeptide (Brant et al., 1967) and polynucleotide (Olson and Flory, 1972) random coils, or whether it should be increased to a value more appropriate to double helical DNA. Notably, the effective van der Waals' radii of nucleotide atoms in the Poltev force field (Zhurkin et al., 1981), which was parameterized on the basis of the interactions of aromatic molecules that mimic the stacking of nucleic acid bases, are larger by as much as 0.2 Å than those in force fields originally developed to treat alkanes and peptides (Olson and Srinivasan, 1990). As detailed in the Methods section, we performed a series of tests in which the minimum in the Lennard–Jones terms was relocated to distances in the range of 0.2–0.4 Å greater than the sum of the van der Waals' radii. Preliminary results showed that changes in the location of the energy minimum over this range did not have any adverse effects on the sequence-dependent structural features of the double helix, but did, as expected, systematically increase the local rise between neighboring base-pair planes. An adjusted Lennard–Jones potential energy minimum at a distance 0.38 Å greater than the van der Waals' separation produced average local base-pair and helical rise values of 3.2 ± 0.1 Å, in remarkably good agreement with crystallographic observations (3.3 ± 0.1 Å). Additional displacement of the Lennard–Jones potential energy minimum to a distance 0.4 Å greater than the sum of van der Waals' radii, however, bent the computed DNA structure severely, indicating that further adjustment along these lines was not feasible.

Figures 4–6 and A1 (Appendix) compare selected sequence-dependent conformational parameters of the bases and chain backbones in the DNA dodecamer structures predicted on the basis of the original and fully improved DUPLEX force fields with the corresponding values in high-resolution crystal (Table 2) and NMR (Tjandra et al., 2000) structures. The sequence-dependent features of the revised model more closely mimic the experimentally observed trends than those based on the original potential. Table 4 shows similar improvements in the mean values and standard deviations of complementary base-pair orientation angles, base-pair step parameters, local helical angles and distances, chain torsion angles, deoxyribose pseudorotation phase angle, and major/minor groove widths/depths in the computer-generated dodecamers compared to the crystal data. The stereo images in Fig. 7 show the approximate global similarity of the structure based on the improved force field, the original Dickerson–Drew dodecamer (NDB_ID: BDL001; Drew et al., 1981), and the dodecamer structure in solution (PDB_ID: 1DUF, model 1, Tjandra et al., 2000). The original crystal structure is virtually indistinguishable at this level from the higher-resolution Dickerson–Drew dodecamer structures listed in Table 2 (with heavy-atom root-mean-square differences less than 0.5 Å from BDL001).

We returned to the AF-modified DNA conformer mixture as an additional check of the revised force field. Reminimization of the representative structures yielded an energy difference of 1.8 kcal/mol in favor of the AF-intercalated, base-displaced form, compared to the experimentally expected ~1.3 kcal/mol preference for the AF-external form.

We then recreated Table 3, using the preceding two newly energy-minimized states, to complete the optimization cycle. The results in Table 5 show that energies within ±3 kcal/mol from the theoretical target of ΔE = 1.3 kcal/mol favoring the AF-external conformer are reproduced by functions within the shaded area (between *Curves A* and *D* of Fig. 3). This observation prompted us to compare the new steep dielectric function (Fig. 3, *Curve A*) with the dielectric expression proposed for DNA by Young et al. (1998), (Fig. 3, *Curve B*), and the function originally used in DUPLEX (Fig. 3, *Curve C*), represented, respectively, in Table 5 with ΔE of −1.5 kcal/mol (c = 0.5, e = 1.0) and 2.3 kcal/mol (c = 2.5, e = 2.0), in reproducing sequence-dependent features of the Dickerson–Drew dodecamer. In these comparisons, the improved pseudorotation profile and relocated Lennard–Jones potential energy minimum were used. Figs. A2 and A3 (Appendix) show that the new steep dielectric function performs as well as or slightly better than the other two functions, in accounting for both local base-pair geometry and groove dimensions of the crystal structures.

Finally, the conformational features of the dodecamer structure generated with the fully revised force field are in excellent agreement with the recent NMR solution structure of the Dickerson–Drew dodecamer in an aqueous dilute liquid crystalline phase (Tjandra et al., 2000), using the lowest energy and best representative conformer in the ensemble of five reported structures (PDB_ID: 1DUF; Berman et al., 2000). The results in Figs. 4–5 and A1 (Appendix) show that the conformational predictions of the revised force field match the details of the solution structure even better than those of the crystals.

DISCUSSION

Dielectric function

Electrostatic interactions are among the most important forces in determining structural and functional features of proteins and nucleic acids. The calculation of the electrostatic contribution to the potential energy is dependent on many factors, including the assigned magnitude of partial charges, the behavior of solvent and counterions, the accessibility of solvent to solute, the surface hydrophobicity, and the polarizability of solute. In all-atom models of biopolymers surrounded by explicit waters, the dielectric constant is assigned a value of unity and the electrostatic interactions of

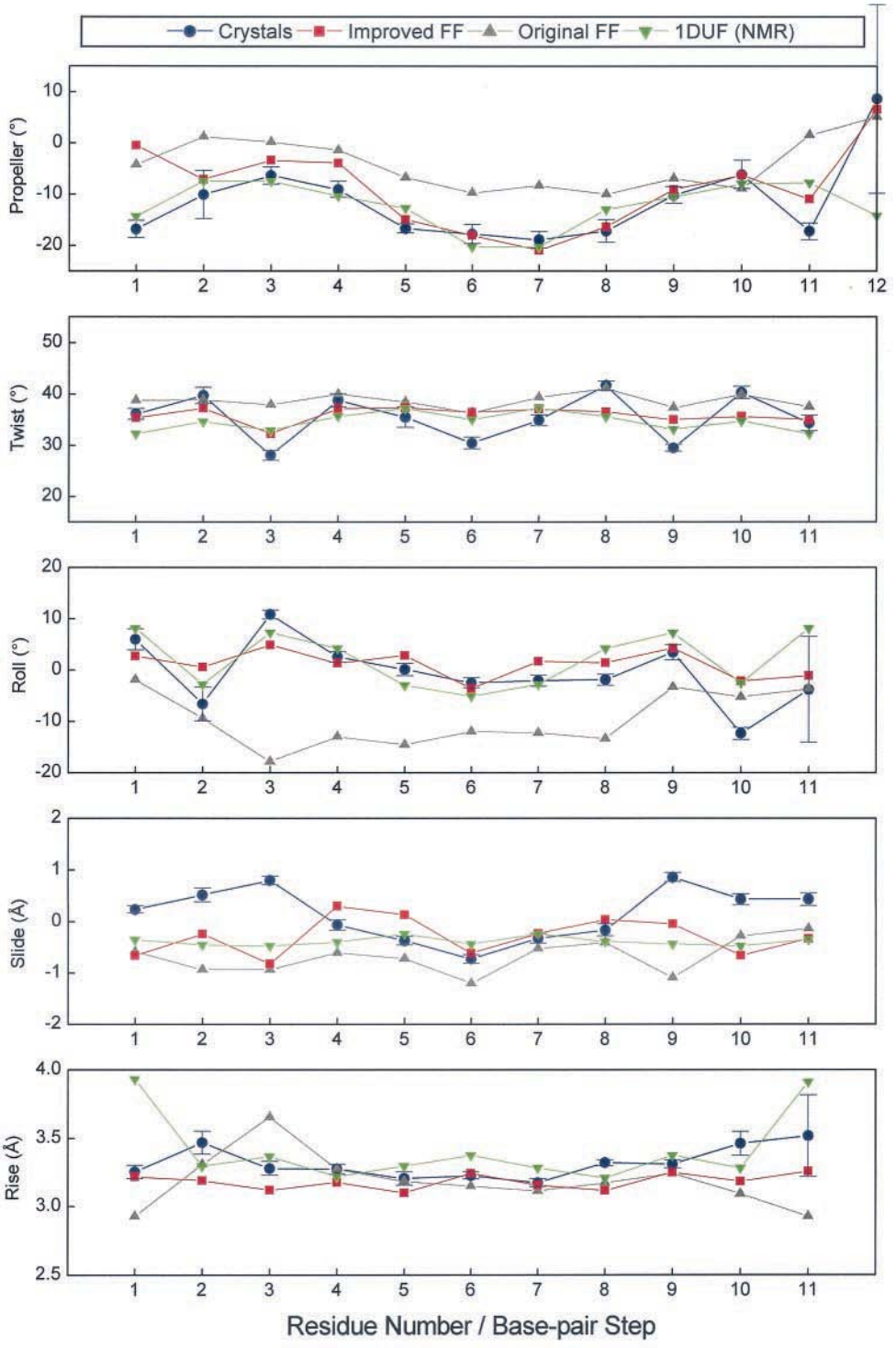


FIGURE 4 Variation of selected sequence-dependent base-pair parameters of high resolution Dickerson–Drew dodecamer crystals (Table 2), the dodecamer structure in an aqueous dilute liquid crystalline phase (PDB_ID: 1DUF, model 1) (Tjandra et al., 2000, and the DUPLEX-generated B-DNA structures of the same sequence, using original and improved force fields. Values calculated using RNA (Babcock et al., 1993). Color coding: *blue circles*, crystals; *green inverted triangles*, 1DUF; *red squares*, improved force field (FF); *gray triangles*, original force field. Crystal data are averages at corresponding sites in different structures, with the brackets denoting standard deviations.

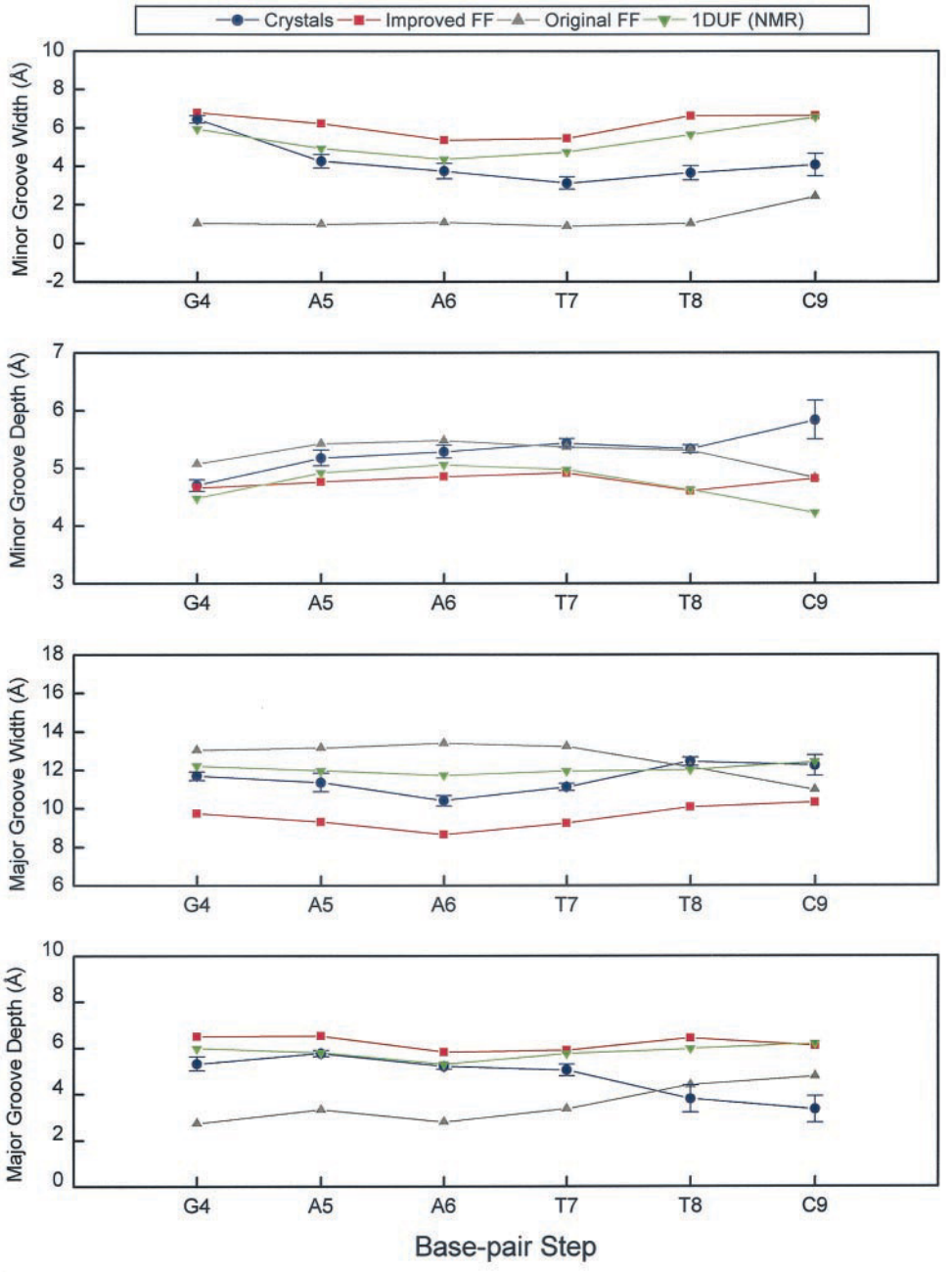


FIGURE 5 Variation of groove widths and depths of high-resolution Dickerson–Drew dodecamer crystals (Table 2), the dodecamer structure in an aqueous dilute liquid crystalline phase (PDB_ID: 1DUF, model 1) (Tjandra et al., 2000), and the DUPLEX-generated B-DNA structures of the same sequence, using the original and improved force fields. Values calculated using Curves 5.2 (Lavery and Sklenar, 1988, 1989). Color coding and symbols are the same as in Fig. 4.

all possible atom pairs are considered. This exact and computationally very expensive approach requires elaborate treatments, such as molecular dynamics or Monte Carlo simulations, to equilibrate the solvent continuously. Early solvent-implicit treatments of nucleic acids performed with limited computer resources, by contrast, assumed a dielectric constant of 2–4 to mimic the organic medium of the sugars and bases between successive negatively charged phosphates (Ramachandran and Srinivasan, 1970; Olson and Flory, 1972). Such treatments, although highly effective for homogeneous macromolecules such as randomly coiling chains composed of identical chemical repeating units (Flory, 1953), are very rough approximations for ordered biological macromolecules such as nucleic acid helices or globular proteins, which, respectively, contain grooves or clefts that can partially sequester counterions and solvent

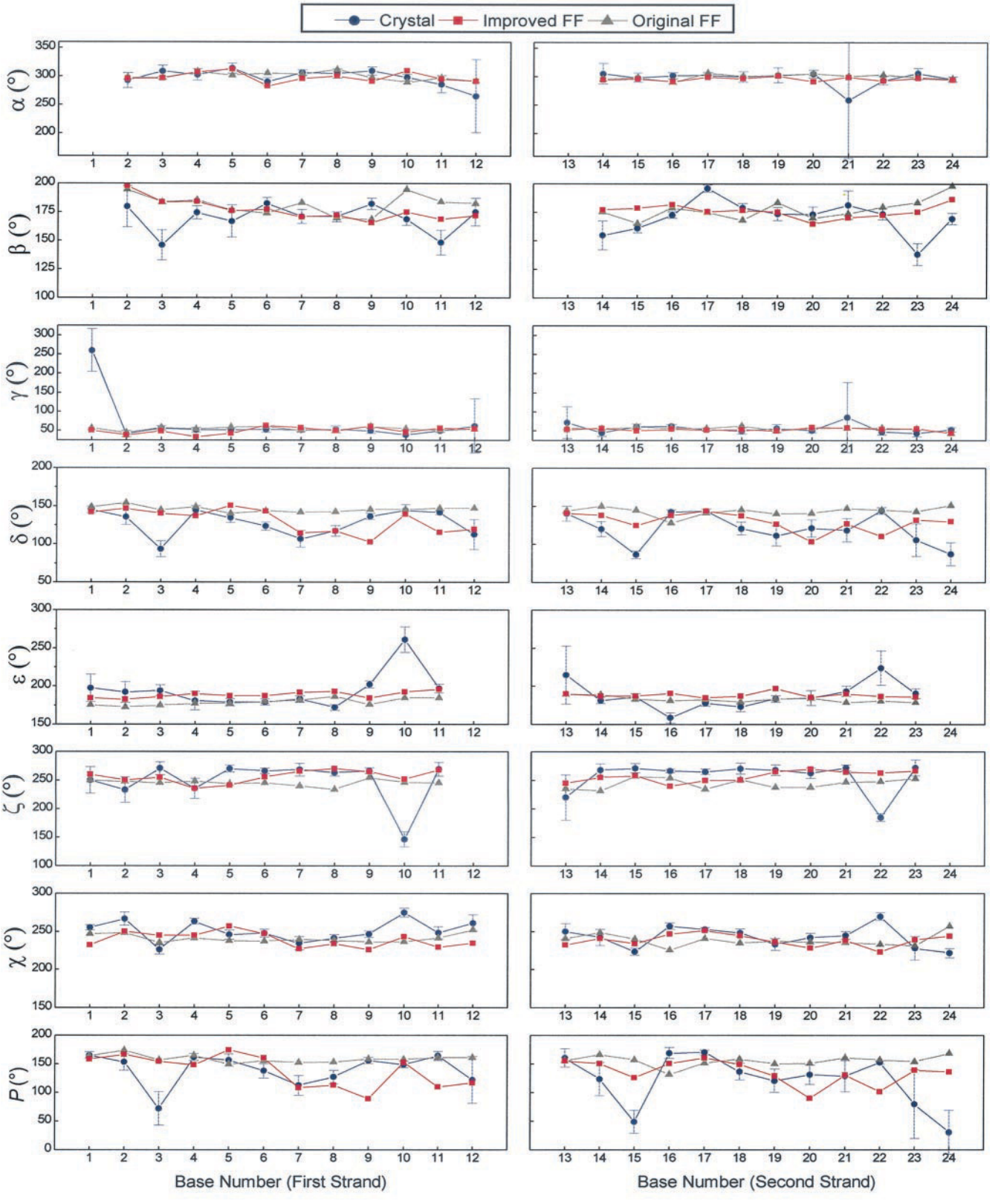


FIGURE 6 Variation of torsion angles and deoxyribose pseudorotation phase angle *P* of high-resolution Dickerson–Drew dodecamer crystals (Table 2), and the DUPLEX-generated B-DNA structures of the same sequence, using the original and improved force fields. Values calculated using Curves 5.2 (Lavery and Sklenar, 1988, 1989). Color coding and symbols are the same as in Fig. 4. Residue numbers reported in Curves 5.2 were shifted +1 for α and β, in accordance with the IUPAC convention (Saenger, 1984).

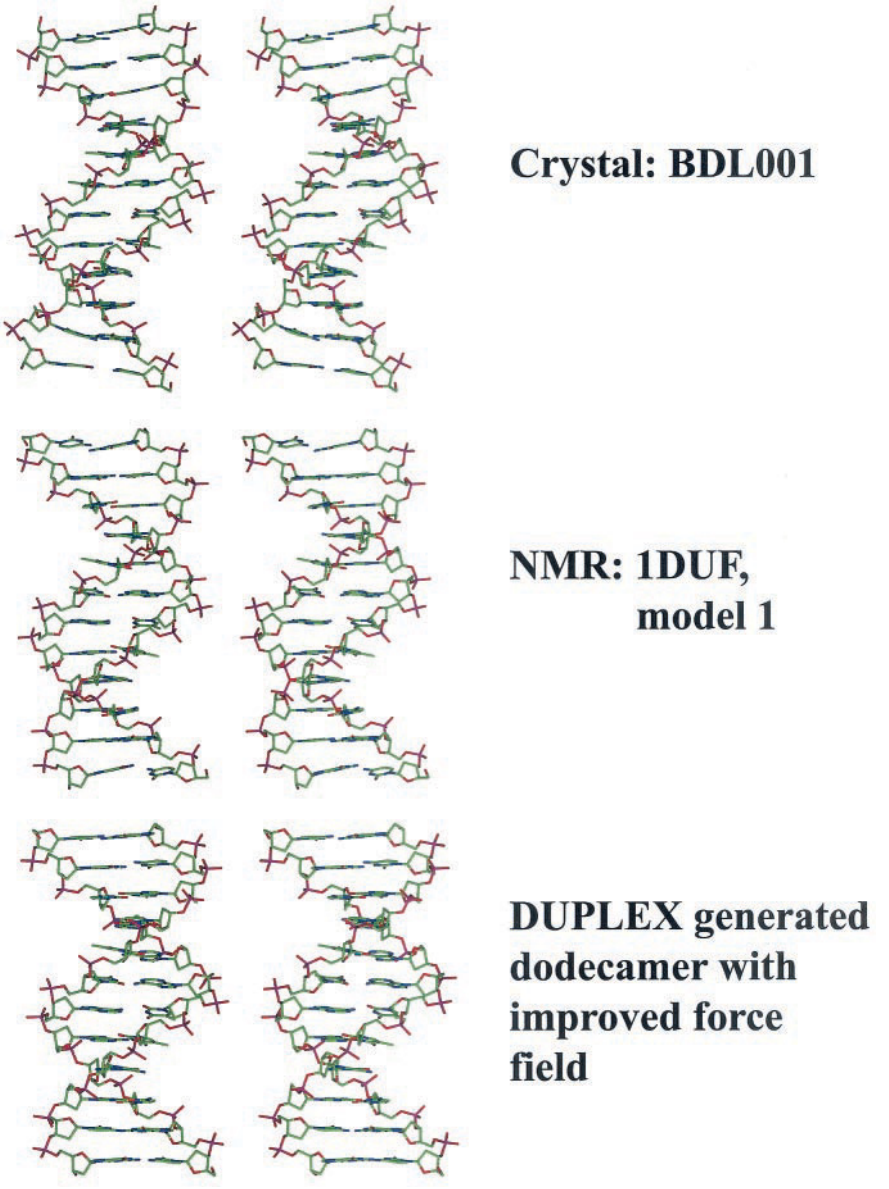


FIGURE 7 Stereo views of the original Dickerson–Drew dodecamer crystal (NDB_ID: BDL001; *top*) (Drew et al., 1981), the high-resolution NMR solution structure of the same dodecamer in an aqueous dilute liquid crystalline phase (PDB_ID: 1DUF, model 1; *middle*) (Tjandra et al., 2000), and the DUPLEX generated B-DNA of the identical sequence obtained with the fully revised force field (*bottom*).

molecules (Jayaram et al., 1989, 1990). The distance-dependent treatment of the dielectric constant, where $\varepsilon(r) = r$ (Brooks et al., 1983), is an improvement over the preceding model, but still does not account accurately for the screening of solvent between charged atoms (Jayaram et al., 1989).

As a further improvement of implicit modeling, Hingerty et al. (1985) offered a modification of Debye's original sigmoidal, distance-dependent dielectric function (Debye, 1929). This analytical expression extended the dependence of the dielectric constant on the distance r between charges to less than 3 Å, yielding a value of $\varepsilon = 1$ at $r = 0$, and $\varepsilon = 78.3$, the dielectric constant of water, as r approached 20 Å (Fig. 3, *Curve C*). Ramstein and Lavery (1988) reformulated this function by introducing adjustable parameters to specify the plateau value of the dielectric term at large distances and the location of its half value. Mazur and Jernigan (1991) subsequently presented an adaptable sigmoidal function that can simulate the functions of both Hingerty et al. (1985) and Ramstein and Lavery (1988), as well as other functional forms that mimic the salt-dependent B- to A-form transition of neighboring base pairs in a poly dG·poly dC duplex. Here we propose a further modification of the earlier function (Hingerty et al., 1985), with two empirical variables so that the function can adopt virtually any sigmoidal form. The form selected here on the basis of the equilibria in solution of AF-modified DNA duplexes derived from high-resolution NMR measurements, reaches solvent saturation rather fast, at an interatomic distance of ~3 Å. The electrostatic field is thus strongly damped and

TABLE 5 Energy differences, ΔE (kcal/mol), between AF-external and AF-intercalated base-displaced DNA conformers for selected values of *c* and *e* in the distance-dependent dielectric function Eq. 1 with improved deoxyribose pseudorotation profile and relocated Lennard–Jones potential energy minimum

	c						
e	0.5	2.5	4.5	6.5	8.5	10.5	12.5
0.25	−0.2	9.4	10.1	9.1	8.2	7.6	7.2
0.5	−1.0	7.4	10.1	10.0	9.3	8.6	8.1
1.0	−1.5	4.8	8.9	10.2	10.2	9.8	9.2
2.0	−1.7	2.3	6.6	9.0	10.0	10.3	10.1
4.0	−1.8	0.5	4.0	6.8	8.6	9.6	10.1
6.0	−1.8	−0.2	2.6	5.2	7.3	8.6	9.5
8.0	−1.8	−0.6	1.7	4.1	6.2	7.7	8.8

Shaded regions represent functions that yield energy differences within ±3 kcal/mol of the expected value, ~1.3 kcal/mol.

close to the dielectric screening of bulk water at a distance corresponding to just one solvation layer away from the charged DNA skeleton. This functional form is much steeper than our earlier dielectric expression, but is closer to functions suggested by others (Grahame, 1950; Takashima and Schwan, 1965; Pennock and Schwan, 1969; Harvey and Hoekstra, 1972; Young et al., 1998; Hassan et al., 2000), including the functional form for B-DNA proposed by Mazur and Jernigan (1991).

We find that steep dielectric functions systematically disfavor the AF-external conformer relative to the AF-intercalated, base-displaced one (Table 5, negative entries). This preference appears to stem from an overdampening of electrostatic interactions between phosphate groups across the minor groove in the intercalated conformer, the width of which is widened relative to the external conformer; P···P distances in the intercalated and external conformers are respectively: P5 ··· P17, 13.2 vs. 12.9 Å; P6 ··· P16, 15.4 vs. 13.3 Å; P7 ··· P15, 16.4 vs. 13.0 Å, where the numbers on the phosphorus atoms refer to the residues in Fig. 1 *B*. Specific distinctions of this type between different conformers of carcinogen-modified DNAs highlight the fact that the dielectric environment differs in various regions of DNA, both modified and unmodified, as previously noted by Young et al. (1998). Accordingly, benchmark structures can offer only approximate energetic criteria.

Our new steep dielectric function is remarkably similar to the dielectric behavior deduced by Young et al. (1998) from the computed positions of explicit waters and Na^+ counterions around a B-form DNA oligonucleotide duplex during a 14-ns molecular dynamics trajectory, using an adaptation of Kirkwood dielectric theory (Kirkwood, 1939). The average dielectric constant ($\varepsilon = 59$) at a distance of 2.8 Å from the DNA surface—reflecting the different dielectric behavior of interactions involving major groove ($\varepsilon = 53$), minor groove ($\varepsilon = 51$), and backbone ($\varepsilon = 66$) atoms in the simulations—is thus somewhat smaller than the value of 78.2 in our revised function at the same distance. Currently available experimental information on the dielectric constant of solvent in the DNA grooves—a value of $\varepsilon = 55$ in the major groove according to fluorescence studies of dansylated-DNA-drug complexes (Barawkar and Ganesh, 1995) and a value of $\varepsilon = 20$ in the minor groove based on the fluorescence properties of bound bisbenzimide molecules (Jin and Breslauer, 1988)—suggests, however, that the average dielectric constant at 2.8 Å may be even lower than the values deduced by Young et al. The empirical Hingerty–Lavery sigmoidal function to which the dielectric behavior of the solvent was fitted by Young et al. (1998) is compared with our new steep dielectric function in Fig. 3 (*Curves B* and *A*, respectively). The set of dielectric functions in water, developed by Mehler and associates for peptides and proteins (Mehler, 1996; Mehler and Guarnieri, 1999; Hassan et al., 2000), reaches a bulk value of 78.3 at ~5 Å, suggesting further similarity in the distance-dependent dielectric activity of proteins and nucleic acids.

The fluctuations in the NMR solution structures of AF-modified duplexes (Mao et al., 1998a,b) coupled with the multiple minimum problem make energetic comparison of the two representative conformers used in this study only an approximate benchmark for force-field optimization. The dielectric function presented here, however, can be easily adjusted to fit different contexts as other reliable benchmarks arise. It appears that functional forms like the present expression or that of Young et al. (1998) will be applicable in other cases.

As pointed out by a reviewer, our proposed steep dielectric function resembles a two-step dielectric model implemented in the Delphi version (Nicholls et al., 1990) of the finite difference adaptation of Poisson–Boltzmann theory (Honig and Nicholls, 1995). To explore whether such a two-step dielectric might be useful instead of the more costly steep function, we recomputed the energy difference between the two conformers used to generate Table 5, using $\varepsilon = 4$ if $r < 2.71$ Å and $\varepsilon = 78.3$ if $r > 2.71$ Å. This produced an energy difference of 1.7 kcal/mol, very close to the expected value of 1.3 kcal/mol, suggesting that such a two-step dielectric might be worthwhile. Of course, considerable further testing of this approach would be needed. Systematic comparison with finite difference adaptations of Poisson–Boltzmann theory would be interesting future work.

Deoxyribose pseudorotation profile

The DUPLEX package previously used a modification of the Sato deoxyribose pseudorotation profile (Sasisekharan, 1973) where the puckering amplitude τ_m, and bond angles were shifted with respect to the phase angle *P* of pseudorotation (Fig. 2, *gray line*) (Hingerty and Broyde, 1982). This modification was made primarily to match experimental and theoretical data available (Olson, 1982) when the force field was first constructed. The current work, how-

ever, shows that this DUPLEX-generated Dickerson–Drew dodecamer was overtwisted with an extremely narrow minor groove, if this pseudorotation treatment is used. These deficiencies have now been remedied with the deoxyribose pseudorotation profile shown in Fig. 2 (*red line*), which corresponds to the series of structures originally proposed by Sato (1983) but with their energies halved. Local energy minima occur at $P = 0°$ (C3′-endo) and $P = 180°$ (C2′-endo) with a slight favoring of the C2′-endo region by ~0.6 kcal/mol imposed by the *gauche* term (Olson, 1982). The absence of the well-known high-energy barrier through the western region ($P \approx 270°$) of the pseudorotation cycle (Saenger, 1984) reflects the omission of the base in the energetic treatment by Sato (see Methods). The steric effects, which give rise to this barrier, are accounted for explicitly by base–sugar interactions in the DUPLEX force field. Notably, the internal bond angles of the deoxyribose ring in the improved profile agree quite well with values (Fig. 2, *shaded boxes*) found in ultra-high-resolution crystal structures of mononucleosides and mononucleotides (Gelbin et al., 1996). The correct treatment of the deoxyribose is key not only to simulating B-DNA structures with proper twist (Fig. 4), as suggested by Olson and Zhurkin (2000), but also to reproducing proper groove widths (Fig. 5) when implicit solvent and counterions are used.

Lennard–Jones potential energy minimum location

It has been known for many years, in computing Lennard–Jones interactions, that the individual van der Waals' radii need to be adjusted to counter lattice compression effects (Brant et al., 1967; Olson and Flory, 1972; Srinivasan and Olson, 1980; Olson and Srinivasan, 1990). Without such adjustment, the interaction potential between neighboring base pairs is minimized at stacking distances up to 0.6 Å less than the normal B-DNA base-pair separation of 3.4 Å (Olson, 1978; Yoon et al., 1984). Minimizing the Lennard–Jones potential energy at a distance 0.2 Å greater than the normal van der Waals' separation, following the early studies of Brant et al. (1967) and Olson and Flory (1972), has become a standard practice. The low rise between neighboring base pairs in the Dickerson–Drew dodecamer structure generated with the new steep dielectric function and modified deoxyribose pseudorotation profile was largely corrected by relocating the Lennard–Jones potential energy minimum to a distance 0.38 Å greater than the sum of the van der Waals' radii.

The improved van der Waals' term closely resembles the expressions in other commonly used nucleic acid force fields, including one by Poltev (Zhurkin et al., 1981), AMBER (Cornell et al., 1995), and CHARMM (Foloppe and MacKerell, 2000; Mackerell and Banavali, 2000), with the exception of the depth of the energy well of the phosphorus–phosphorus Lennard–Jones interactions. Table A1 (Appendix) compares values of r_{min}, the distance between an atom pair at the potential energy minimum, and V_{min}, the well-depth in our improved force field, with the corresponding parameters of the AMBER, CHARMM, and Poltev force fields. The value of V_{min} for phosphorus–phosphorus (P · · · P) interactions in our improved force field is −1.06 kcal/mol, but only −0.58 to −0.20 in the other force fields. Modification of the depth of the P· · ·P well to values comparable to those used in other force fields (−0.30), however, has only a minimal effect on the conformational parameters of the dodecamer. Figures A4 and A5 (Appendix) show these results.

Solution versus crystal structures

A key issue in testing a force field involves selection of the appropriate benchmark. Although crystal structures represent the largest body of structural data with the least ambiguity, the question of just how these observations relate to solution structures is an actively debated question (Harvey et al., 1995; Dickerson et al., 1996; Shatsky-Schwartz et al., 1997; Beveridge and McConnell, 2000). Notably, the many crystal structures of the Dickerson–Drew dodecamer are very similar to one another, despite ongoing controversies concerning issues, such as the number of layers of water in the hydration spine (Shui et al., 1998b) and the types of ions that play crucial coordinating roles (Tereshko et al., 1999b; McConnell and Beveridge, 2000). The subtle interplay of different counterions, crystallization conditions, packing forces, and their effects on structural nuances remains to be unraveled. Gross end effects, such as the rupture of terminal base pairs, which then pack with neighboring helices (Liu et al., 1998), are thought to account for some of the irregularities in conformational parameters of the dodecamer crystals in Figs. 4 and 6. Notable sequence-dependence in certain base-pair step parameters, especially roll and twist (Fig. 4), has been shown to stem from packing forces (Subirana and Faria, 1997).

The present force-field treatment, with implicit solvent and counterions, is directed toward simulating reasonable solution structures. The collective crystal data are simply used as a benchmark of the sequence-dependent, conformational structure of DNA rather than as measures of precise atomic arrangements in the crystal lattice. This approach, with the goal of reproducing sequence-dependent trends within the range of observed data, transcends specific issues concerning counterions and the hydration spine. In this connection, it should be noted that, although the improved force field accounts reasonably for the sequence-dependence of widths and depths of the major and minor grooves, the magnitudes of the parameters deviate somewhat from crystal values (Fig. 5, *red squares* and *blue circles*). Moreover, computations using the less steep, distance-dependent dielectric function of Young et al. (1998) (Fig. 3, *Curve B*) in combination with

the improved pseudorotation profile and relocated Lennard–Jones potential energy minimum perform similarly (Fig. A3, *magenta squares*; Appendix). Interestingly, a 15-ns molecular dynamics study of the Dickerson–Drew dodecamer with explicit solvent and counterions shows a comparable average minor groove width (McConnell and Beveridge, 2000), suggesting that the groove width may differ in solution compared to the crystalline state, possibly due to differing dielectric environments. In addition, both the DUPLEX-generated dodecamer structure and the dynamical simulations of McConnell and Beveridge (2000) reproduce the narrowness in the central AT-tracts of the DNA sequence, even though the average major groove width in the present study is smaller than the crystal average, and that of the molecular dynamics simulation is larger.

The recent NMR solution structure of the Dickerson–Drew dodecamer in an aqueous liquid crystalline phase (Tjandra et al., 2000) provides an opportunity for comparing the DUPLEX-generated dodecamer structure with a highly resolved DNA solution structure. As shown in Figs. 4–5, and A1 (Appendix), the computed structure is overall very close to the lowest energy and best representative member of the ensemble of five solution structures. In particular, the minor groove width is somewhat wider on average in the solution structure than in the crystals, and this is reproduced with the revised force field. Interestingly, the efficiently generated dodecamer produced by DUPLEX reproduces sequence-dependent conformational features in the NMR structure at least as closely as other force fields with state-of-the-art molecular dynamics simulations using explicit solvent and counterions (Figs. 4–6 and A1; Appendix) (Cheatham and Young, 2001).

CONCLUSIONS

The improved force field in the present work satisfactorily reproduces the sequence-dependent structural features of the Dickerson–Drew dodecamer, namely the average geometry of complementary and neighboring base pairs, as well as the less clear sequence-dependent variation of torsion angles and pseudorotation parameters. The revised force field also reproduces the relative sequence-dependent trends of dimer tilt, roll, and twist angles found in the dynamical solution structure based on a 15-ns molecular dynamics study with AMBER (McConnell and Beveridge, 2000), although, as reported to date, the AMBER force field yields undertwisted DNA duplexes with neighboring base pairs unwound by ~3–4° compared to those in crystal structures (Cheatham and Young, 2001). We also note that Mazur (1998) has reported a successful reproduction of the original Dickerson–Drew dodecamer crystal structure (Wing et al., 1980; Drew et al., 1981) using a combination of partial explicit hydration in the grooves and implicit treatment of counterions according to Manning's theory (Manning, 1978), as in our approach, but with a linear distance-dependent dielectric function, $\varepsilon(r) = r$. The 5-ns trajectories based on this strategy converge to conformations close to that of the crystal structure in terms of root-mean-square atomic deviations, but the degree to which the sequence-dependent conformational parameters are reproduced remains to be reported.

Given the computational cost involved in molecular dynamics studies and the relatively robust performance of the revised DUPLEX force field, it is reassuring that the implicit treatment of solvent and counterions can still play an important role in molecular mechanics calculations. These computational shortcuts are particularly valuable for the extensive conformational searches with large numbers of energy minimization trials that must be carried out in studying carcinogen-DNA adducts; such detailed searches are beyond the present-day capabilities of molecular dynamics or Monte Carlo calculations with explicit solvent and counterions. In addition, NMR solution structures of carcinogen-damaged DNAs, which can be efficiently delineated in conjunction with molecular mechanics searches using distance restraints, benefit from a reliable force field with implicit treatment of solvent. Molecular dynamics simulations, using all-atom force fields such as AMBER (Case et al., 1999) with explicit solvent and counterions are, of course, essential for deciphering the structural codes that govern protein–DNA interactions, and for estimating conformational and binding free energies (Kollman et al., 2000). The force field treatments described here apply in the context of the DUPLEX molecular mechanics program, but similar approaches can be conveniently tested and modified, both in other force fields and with other benchmarks. We anticipate that these simple and efficient treatments will be useful in elucidating structure–function relationships for other applications, such as DNA–ligand interactions.

APPENDIX

The Appendix contains five figures (Figs. A1–A5) presenting the variation of sequence-dependent base-pair parameters, groove widths, and depths in computed and experimental structures of the Dickerson–Drew dodecamer, along with one table (Table A1) presenting van der Waals' parameters, r_{min} and V_{min}, of nucleic acid atom pairs in different force fields.

TABLE A1 Van der Waals' parameters, r_{min} and V_{min}, of nucleic acid atom pairs in different force fields*

Atoms†	r_{min} (Å) DUPLEX‡	r_{min} (Å) Poltev§	r_{min} (Å) AMBER¶	r_{min} (Å) CHARMM‖	V_{min} (kcal/mol) DUPLEX‡	V_{min} (kcal/mol) Poltev§	V_{min} (kcal/mol) AMBER¶	V_{min} (kcal/mol) CHARMM‖
H-H	2.78	2.70	2.77	2.20	−0.12	−0.05	−0.02	−0.05
C1-H	3.28	3.35	3.30	3.14	−0.11	−0.04	−0.04	−0.06
C1-C1	3.78	4.00	3.82	4.08	−0.13	−0.03	−0.11	−0.08
C2-H	3.28	3.30	3.30	3.00	−0.11	−0.05	−0.04	−0.07
C2-C1	3.78	3.95	3.82	3.94	−0.13	−0.04	−0.10	−0.09
C2-C2	3.78	3.90	3.82	3.80	−0.13	−0.06	−0.09	−0.10
C3-H	3.28	3.35	3.30	3.38	−0.11	−0.04	−0.04	−0.03
C3-C1	3.78	4.00	3.82	4.32	−0.13	−0.03	−0.11	−0.04
C3-C2	3.78	3.95	3.82	4.18	−0.13	−0.04	−0.10	−0.04
C3-C3	3.78	4.00	3.82	4.55	−0.13	−0.03	−0.11	−0.02
N1-H	3.13	3.23	3.21	2.95	−0.18	−0.06	−0.05	−0.10
N1-C1	3.63	3.88	3.73	3.89	−0.19	−0.05	−0.14	−0.12
N1-C2	3.63	3.83	3.73	3.75	−0.19	−0.06	−0.12	−0.14
N1-C3	3.63	3.88	3.73	4.12	−0.19	−0.05	−0.14	−0.06
N1-N1	3.48	3.76	3.65	3.70	−0.29	−0.07	−0.17	−0.20
N2-H	3.13	3.25	3.21	2.95	−0.18	−0.04	−0.05	−0.10
N2-C1	3.63	3.90	3.73	3.89	−0.19	−0.04	−0.14	−0.12
N2-C2	3.63	3.85	3.73	3.75	−0.19	−0.05	−0.12	−0.14
N2-C3	3.63	3.90	3.73	4.12	−0.19	−0.04	−0.14	−0.06
N2-N1	3.48	3.78	3.65	3.70	−0.29	−0.06	−0.17	−0.20
N2-N2	3.48	3.80	3.65	3.70	−0.29	−0.05	−0.17	−0.20
N3-H	3.13	3.20	3.21	2.95	−0.18	−0.07	−0.05	−0.10
N3-C1	3.63	3.85	3.73	3.89	−0.19	−0.06	−0.14	−0.12
N3-C2	3.63	3.80	3.73	3.75	−0.19	−0.08	−0.12	−0.14
N3-C3	3.63	3.85	3.73	4.12	−0.19	−0.06	−0.14	−0.06
N3-N1	3.48	3.73	3.65	3.70	−0.29	−0.09	−0.17	−0.20
N3-N2	3.48	3.75	3.65	3.70	−0.29	−0.07	−0.17	−0.20
N3-N3	3.48	3.70	3.65	3.70	−0.29	−0.11	−0.17	−0.20
O1-H	3.10	3.00	3.11	2.87	−0.12	−0.06	−0.06	−0.08
O1-C1	3.60	3.65	3.63	3.81	−0.14	−0.05	−0.15	−0.11
O1-C2	3.60	3.60	3.63	3.67	−0.14	−0.06	−0.13	−0.12
O1-C3	3.60	3.65	3.63	4.05	−0.14	−0.05	−0.15	−0.06
O1-N1	3.45	3.53	3.55	3.62	−0.21	−0.07	−0.19	−0.17
O1-N2	3.45	3.55	3.55	3.62	−0.21	−0.06	−0.19	−0.17
O1-N3	3.45	3.50	3.55	3.62	−0.21	−0.09	−0.19	−0.17
O1-O1	3.42	3.30	3.44	3.54	−0.15	−0.08	−0.21	−0.15
O2-H	3.10	3.00	3.07	2.87	−0.12	−0.06	−0.05	−0.08
O2-C1	3.60	3.65	3.59	3.81	−0.14	−0.05	−0.14	−0.11
O2-C2	3.60	3.60	3.59	3.67	−0.14	−0.06	−0.12	−0.12
O2-C3	3.60	3.65	3.59	4.05	−0.14	−0.05	−0.14	−0.06
O2-N1	3.45	3.53	3.51	3.62	−0.21	−0.07	−0.17	−0.17
O2-N2	3.45	3.55	3.51	3.62	−0.21	−0.06	−0.17	−0.17
O2-N3	3.45	3.68	3.51	3.62	−0.21	−0.07	−0.17	−0.17
O2-O1	3.42	3.30	3.40	3.54	−0.15	−0.08	−0.19	−0.15
O2-O2	3.42	3.30	3.37	3.54	−0.15	−0.08	−0.17	−0.15

(Continued)

TABLE A1 (Continued)

Atoms[†]	r_{min} (Å) DUPLEX[‡]	r_{min} (Å) Poltev[§]	r_{min} (Å) AMBER[¶]	r_{min} (Å) CHARMM[‖]	V_{min} (kcal/mol) DUPLEX[‡]	V_{min} (kcal/mol) Poltev[§]	V_{min} (kcal/mol) AMBER[¶]	V_{min} (kcal/mol) CHARMM[‖]
O3-H	3.10	2.98	3.05	2.80	−0.12	−0.09	−0.06	−0.07
O3-C1	3.60	3.63	3.57	3.74	−0.14	−0.07	−0.15	−0.10
O3-C2	3.60	3.58	3.57	3.60	−0.14	−0.09	−0.13	−0.11
O3-C3	3.60	3.63	3.57	3.98	−0.14	−0.07	−0.15	−0.05
O3-N1	3.45	3.51	3.49	3.55	−0.21	−0.10	−0.19	−0.15
O3-N2	3.45	.353	3.49	3.55	−0.21	−0.08	−0.19	−0.15
O3-N3	3.45	3.40	3.49	3.55	−0.21	−0.15	−0.19	−0.15
O3-O1	3.42	3.20	3.38	3.47	−0.15	−0.13	−0.21	−0.14
O3-O2	3.42	3.20	3.34	3.47	−0.15	−0.13	−0.19	−0.14
O3-O3	3.42	3.11	3.32	3.40	−0.15	−0.22	−0.21	−0.12
P-H	3.38	3.35	3.49	3.25	−0.31	−0.13	−0.06	−0.16
P-C1	3.88	4.00	4.01	4.19	−0.37	−0.11	−0.15	−0.21
P-C2	3.88	3.95	4.01	4.05	−0.37	−0.15	−0.13	−0.24
P-C3	3.88	4.00	4.01	4.43	−0.37	−0.11	−0.15	−0.11
P-N1	3.73	3.88	3.92	4.00	−0.54	−0.17	−0.18	−0.34
P-N2	3.73	3.90	3.92	4.00	−0.54	−0.14	−0.18	−0.34
P-N3	3.73	3.85	3.92	4.00	−0.54	−0.21	−0.18	−0.34
P-O1	3.70	3.65	3.82	3.92	−0.38	−0.18	−0.21	−0.30
P-O2	3.70	3.65	3.78	3.92	−0.38	−0.18	−0.18	−0.30
P-O3	3.70	3.54	3.76	3.85	−0.38	−0.30	−0.20	−0.26
P-P	3.98	4.00	4.20	4.30	−1.06	−0.43	−0.20	−0.58

*r_{min} is the distance between an atom pair at the potential energy minimum; V_{min} is the minimum pairwise interaction energy. Specifically: $V_{vdw}(r_{ij}) = -V_{min}[(r_{min}/r_{ij})^{12} - 2(r_{min}/r_{ij})^{6}]$ where $V_{vdw}(r_{ij})$ is the Lennard–Jones potential at distance r_{ij} (Olson and Srinivasan, 1990).

[†]Atoms are defined as follows: H, hydrogen; C1, sp^3 carbon; C2, sp^2 carbon; C3, CH_3 united carbon; N1, sp^2 glycosyl nitrogen; N2, sp^2 nitrogen of base (NH2); N3, sp^2 nitrogen in aromatic rings; O1, hydroxyl oxygen; O2, ether and ester oxygen; O3, carbonyl oxygen; P, phosphorus.

[‡]Hingerty et al., (1985).

[§]Zhurkin et al., (1981).

[¶]Cornell et al., (1995).

[‖]Foloppe and MacKerell (2000); Mackerell and Banavali (2000).

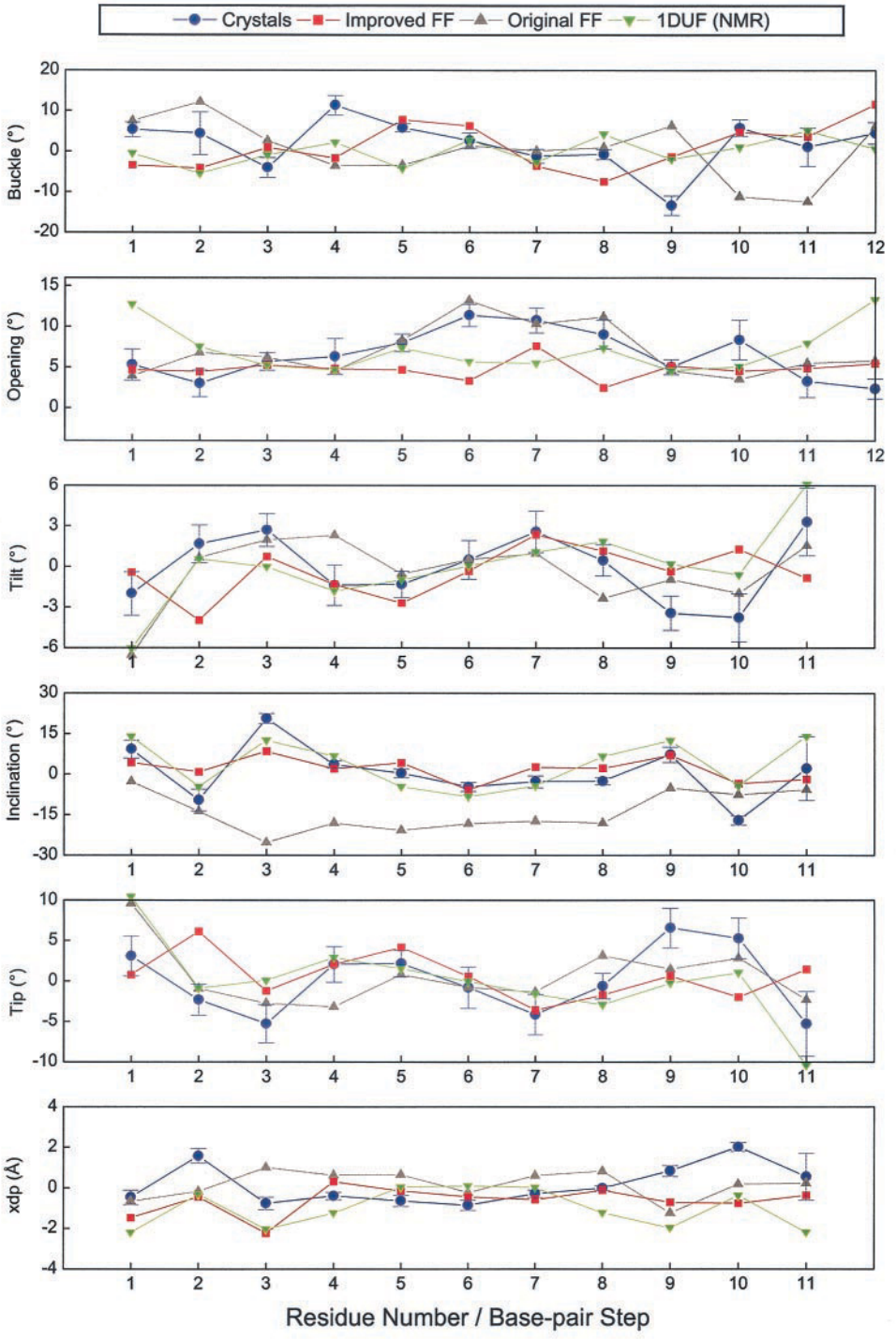


FIGURE A1 Variation of additional sequence-dependent base-pair parameters of high resolution Dickerson–Drew dodecamer crystals (Table 2), the dodecamer structure in an aqueous dilute liquid crystalline phase (PDB_ID: 1DUF, model 1) (Tjandra et al., 2000), and the DUPLEX-generated B-DNA structures of the same sequence, using the original and improved force fields. Values calculated using RNA (Babcock et al., 1993). Color coding and symbols are the same as in Fig. 4.

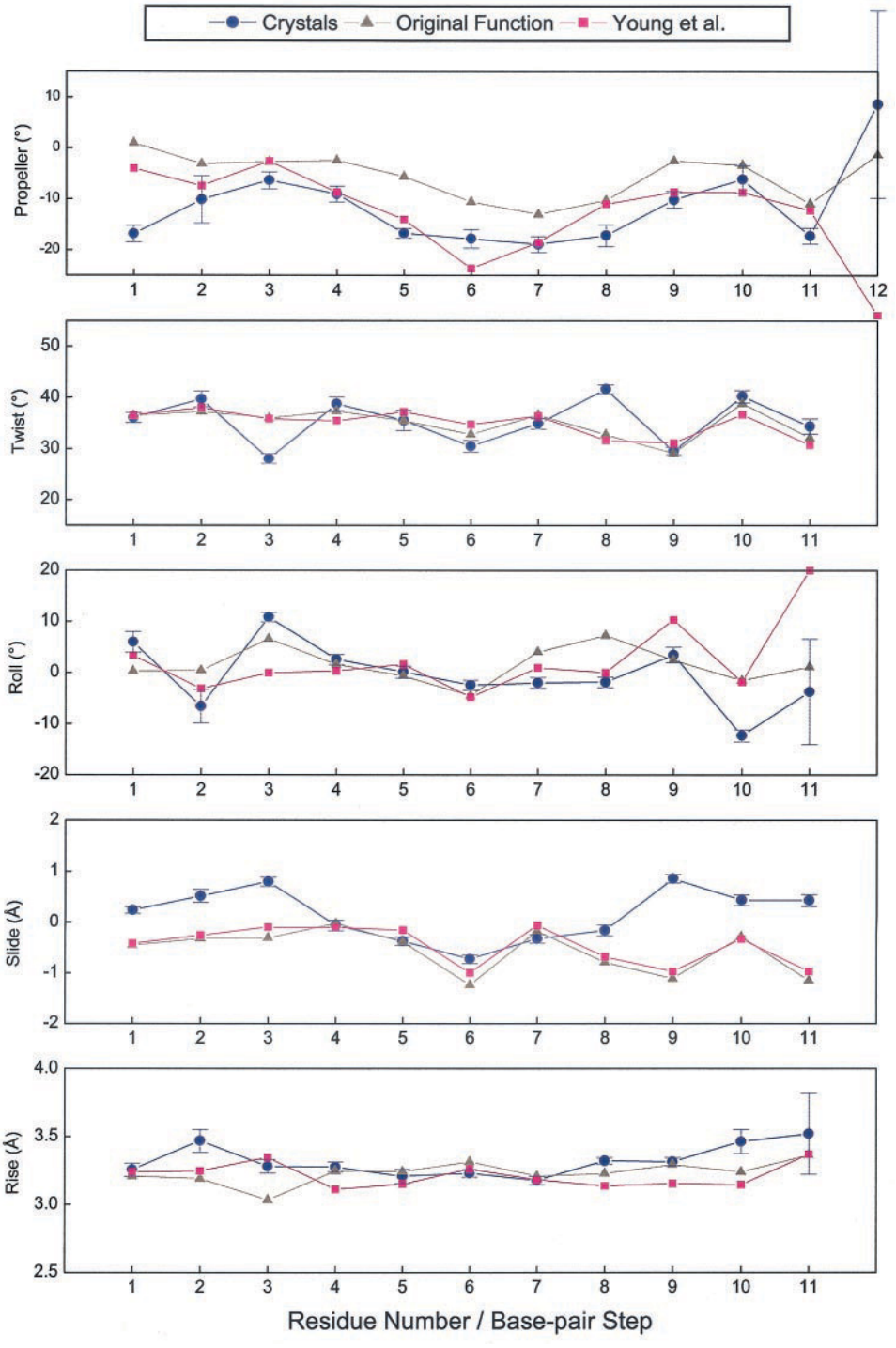


FIGURE A2 Variation of selected sequence-dependent base-pair parameters of high-resolution Dickerson–Drew dodecamer crystals (Table 2), and the DUPLEX-generated B-DNA structures of the same sequence, using the original dielectric function (Hingerty et al., 1985) and the function of Young et al. (1998), both with the improved pseudorotation profile and relocated Lennard–Jones potential energy minimum. Values calculated using RNA (Babcock et al., 1993). Color coding: *blue circles*, crystals; *gray triangles*, original function; *magenta squares*, Young et al. (1998) function.

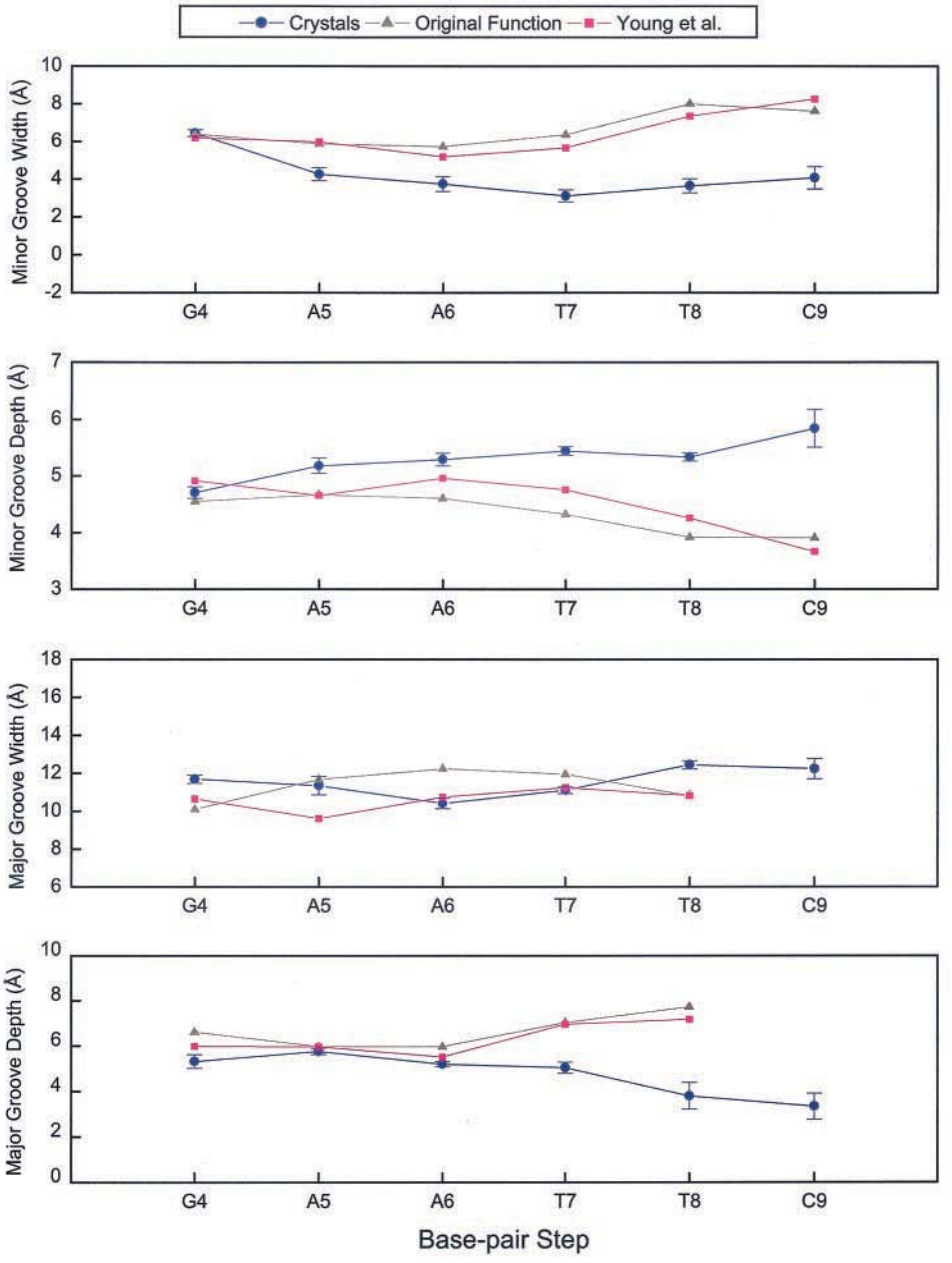


FIGURE A3 Variation of groove widths and depths of high-resolution Dickerson–Drew dodecamer crystals (Table 2), and the DUPLEX-generated B-DNA structures of the same sequence, using the original dielectric function (Hingerty et al., 1985) and the function of Young et al. (1998), both with the improved pseudorotation profile and relocated Lennard–Jones potential energy minimum. Values calculated using Curves 5.2 (Lavery and Sklenar, 1988, 1989). Color coding and symbols are the same as in Fig. A2.

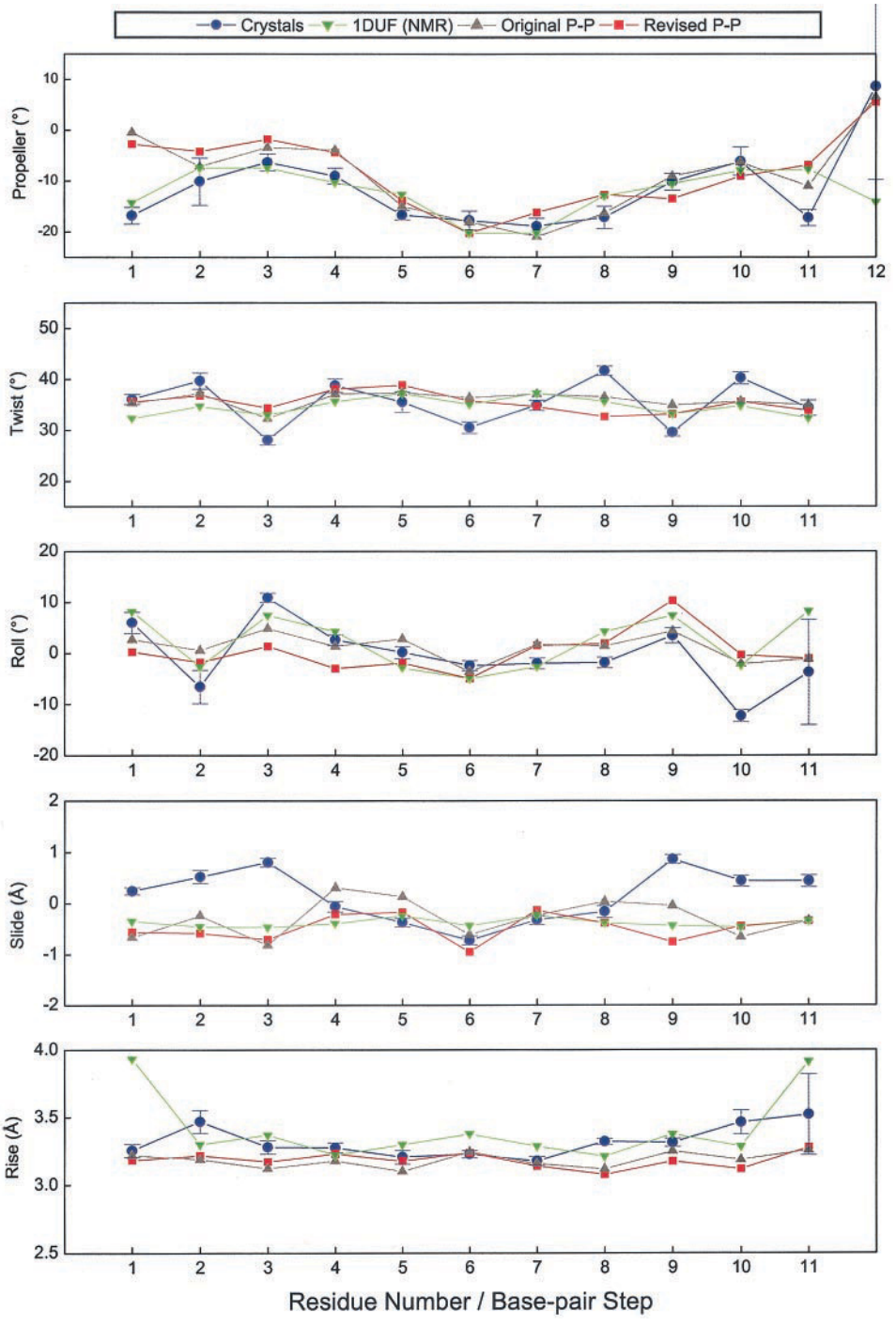


FIGURE A4 Variation of selected sequence-dependent base-pair parameters of high-resolutoin Dickerson–Drew dodecamer crystals (Table 2), the dodecamer structure in an aqueous dilute liquid crystalline phase (PDB_ID; 1DUF, model 1) (Tjandra et al., 2000), and the DUPLEX-generated B-DNA structure of the same sequence, obtained with the improved force field using original and revised phosphorus–phosphorus (P · · · P) well depths. Values calculated uisng RNA (Babcock et al., 1993). Color coding: *blue circles*, crystals; *green inverted triangles*, 1DUF; *gray triangles*, original P · · · P interactions; *red squares*, revised P · · · P interactions.

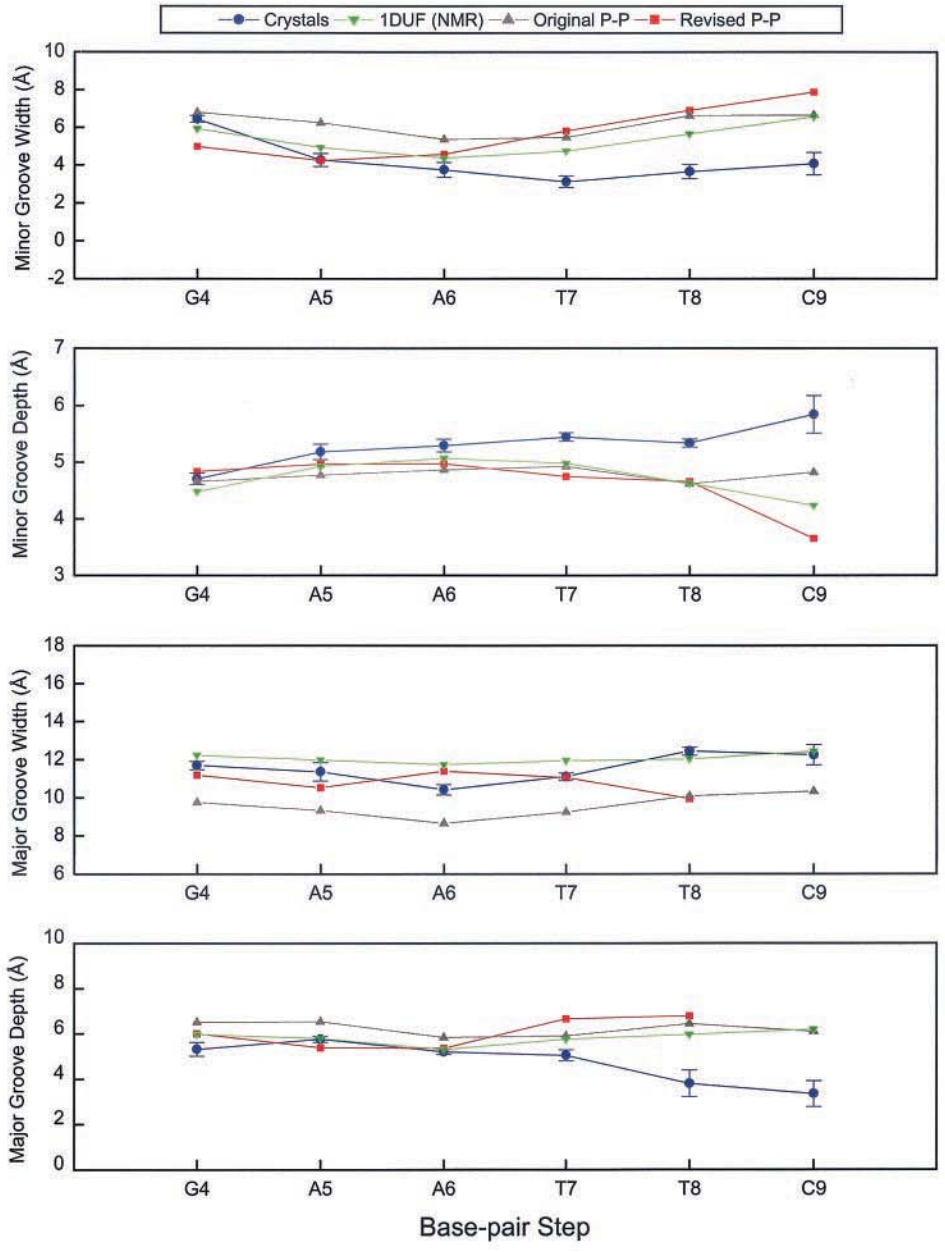


FIGURE A5 Variation of groove widths and depths of high-resolution Dickerson–Drew dodecamer crystals (Table 2), the dodecamer structure in an aqueous dilute liquid crystalline phase (PDB_ID: 1DUF, model 1) (Tjandra et al., 2000), and the DUPLEX-generated B-DNA structures of the same sequence, obtained with the improved force field using original and revised phosphorus–phosphorus (P · · · P) well depths. Values calculated using Curves 5.2 (Lavery and Sklenar, 1988, 1989). Color coding and symbols are the same as in Fig. A4.

We thank Prof. Robert Shapiro, Chemistry Department, New York University, for discussions that inspired this project and Prof. Helen M. Berman, Department of Chemistry, Rutgers University, for discussion of structural data. B.E.H. would like to thank Dr. Gerard Bunick, Life Sciences Division, Oak Ridge National Laboratory and Joel Harp, University of Tennessee/Oak Ridge Graduate Program in Genome Science and Technology for helpful discussions.

This research was supported by National Institutes of Health grants CA75449 and CA28038 and Department of Energy (DOE) grant DE-FG02–90ER60931 to Suse Broyde and National Institutes of Health grant GM20861 to Wilma K. Olson. Computations were carried out at the DOE National Energy Research Supercomputer Center, the National Science Foundation San Diego Supercomputer Center, and the Center for Computational Chemistry at Rutgers University.

REFERENCES

Arnott, S., P. J. Campbell-Smith, and R. Chandrasekaran. 1976. Atomic coordinates and molecular conformations for DNA–DNA, RNA–RNA, and DNA–RNA helices. *In* CRC Handbook of Biochemistry and Molecular Biology. G. Fasman, editor. CRC Press, Cleveland, OH. 411–422.

Babcock, M. S., E. P. Pednault, and W. K. Olson. 1993. Nucleic acid structure analysis: a users guide to a collection of new analysis programs. *J. Biomol. Struct. Dyn.* 11:597–628.

Baker, N. A., D. Sept, S. Joseph, M. J. Holst, and J. A. McCammon. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* 98:10037–10041.

Barawkar, D. A., and K. N. Ganesh. 1995. Fluorescent d(CGCGAATT CGCG): characterization of major groove polarity and study of minor groove interactions through a major groove semantophore conjugate. *Nucleic Acids Res.* 23:159–164.

Berger, I., V. Tereshko, H. Ikeda, V. E. Marquez, and M. Egli. 1998. Crystal structures of B-DNA with incorporated 2′-deoxy-2′-fluoro-arabino-furanosyl thymines: implications of conformational preorganization for duplex stability. *Nucleic Acids Res.* 26:2473–2480.

Berman, H. M., W. K. Olson, D. L. Beveridge, J. Westbrook, A. Gelbin, T. Demeny, S. H. Hsieh, A. R. Srinivasan, and B. Schneider. 1992. The nucleic acid database. A comprehensive relational database of three-dimensional structures of nucleic acids. *Biophys. J.* 63:751–759.

Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. 2000. The Protein Data Bank. *Nucleic Acids Res.* 28:235–242.

Best, S. A., K. M. Merz, Jr., and C. H. Reynolds. 1997. GB/SA-based continuum solvation model for octanol. *J. Phys. Chem. B.* 101: 10479–10487.

Beveridge, D. L., and K. J. McConnell. 2000. Nucleic acids: theory and computer simulation, Y2K. *Curr. Opin. Struct. Biol.* 10:182–196.

Brant, D. A., W. G. Miller, and P. J. Flory. 1967. Conformational energy estimates for statistically coiling polypeptide chains. *J. Mol. Biol.* 23: 47–65.

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

Broyde, S., and B. E. Hingerty. 1999. Effective computational strategies for determining structures of carcinogen-damaged DNA. *J. Comp. Phys.* 151:313–332.

Case, D. A., D. A. Pearlman, J. W. Caldwell, T. E. Cheatham, III, W. S. Ross, C. L. Simmerling, T. A. Darden, K. M. Merz, R. V. Stanton, A. L. Cheng, J. J. Vincent, M. Crowley, V. Tsui, R. J. Radmer, Y. Duan, J. Pitera, I. Massova, G. L. Seibel, U. C. Singh, P. K. Weiner, and P. A. Kollman. 1999. AMBER 6. University of California, San Francisco.

Cheatham, T. E., III, and M. A. Young. 2001. Molecular dynamics simulation of nucleic acids: successes, limitations, and promise. *Biopolymers.* 56:232–256.

Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. 1995. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 117: 5179–5197.

Cramer, C. J., and D. G. Truhlar. 1992. An SCF solvation model for the hydrophobic effect and absolute free-energies of aqueous solvation. *Science.* 256:213–217.

Cramer, C. J., and D. G. Truhlar. 1999. Implicit solvation models: equilibria, structure, spectra, and dynamics. *Chem. Rev.* 99:2161–2200.

Daggett, V., and P. A. Kollman. 1990. Molecular dynamics simulations of active site mutants of triosephosphate isomerase. *Protein Eng.* 3:677–690.

Debye, P. 1929. Polar Molecules. Dover, New York.

Dickerson, R. E., D. S. Goodsell, and M. L. Kopka. 1996. MPD and DNA bending in crystals and in solution. *J. Mol. Biol.* 256:108–125.

Drew, H. R., R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura, and R. E. Dickerson. 1981. Structure of a B-DNA dodecamer: conformation and dynamics. *Proc. Natl. Acad. Sci. U.S.A.* 78:2179–2183.

Egli, M., V. Tereshko, M. Teplova, G. Minasov, A. Joachimiak, R. Sanishvili, C. M. Weeks, R. Miller, M. A. Maier, H. An, P. Dan Cook, and M. Manoharan. 1998. X-ray crystallographic analysis of the hydration of A- and B-form DNA at atomic resolution. *Biopolymers.* 48:234–252.

Flory, P. J. 1953. Principles of Polymer Chemistry. Cornell University Press, Ithaca, NY.

Foloppe, N., and A. D. MacKerell, Jr. 2000. All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comp. Chem.* 21: 86–104.

Friedman, R. A., and B. Honig. 1992. The electrostatic contribution to DNA base-stacking interactions. *Biopolymers.* 32:145–159.

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

Geacintov, N. E., M. Cosman, B. E. Hingerty, S. Amin, S. Broyde, and D. J. Patel. 1997. NMR solution structures of stereoisometric covalent polycyclic aromatic carcinogen-DNA adduct: principles, patterns, and diversity. *Chem. Res. Toxicol.* 10:111–146.

Gelbin, A., B. Schneider, L. Clowney, S. Hsieh, W. K. Olson, and H. M. Berman. 1996. Geometric parameters in nucleic acids: sugar and phosphate constituents. *J. Am. Chem. Soc.* 118:519–529.

Grahame, D. C. 1950. Effects of dielectric saturation upon the diffuse double layer and the free energy of hydration of ions. *J. Chem. Phys.* 18:903–909.

Harvey, S. C., and P. Hoekstra. 1972. Dielectric relaxation spectra of water adsorbed on lysozyme. *J. Phys. Chem.* 76:2987–2994.

Harvey, S. C., M. Dlakic, J. Griffith, R. Harrington, K. Park, D. Sprous, and W. Zacharias. 1995. What is the basis of sequence-directed curvature in DNAs containing A tracts? *J. Biomol. Struct. Dyn.* 13:301–307.

Hassan, S. A., F. Guarnieri, and E. L. Mehler. 2000. A general treatment of solvent effects based on screened Coulomb potentials. *J. Phys. Chem. B.* 104:6478–6489.

Hawkins, G. D., C. J. Cramer, and D. G. Truhlar. 1995. Pairwise solute descreening of solute charges from a dielectric medium. *Chem. Phys. Lett.* 246:122–129.

Hawkins, G. D., C. J. Cramer, and D. G. Truhlar. 1996. Parameterized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium. *J. Phys. Chem.* 100:19824–19839.

Hermann, R. B. 1972. Theory of hydrophobic bonding. 2. Correlation of hydrocarbon solubility in water with solvent cavity surface-area. *J. Phys. Chem.* 76:2754–2759.

Hingerty, B. E., and S. Broyde. 1982. Base displacement in AAF and AF modified dCpdG: *syn* and *anti* guanine. *Int. J. Quantum Chem. Quantum Biol. Symp.* 9:125–136.

Hingerty, B. E., T. L. Ferrell, and J. E. Turner. 1985. Dielectric effects in biopolymers: the theory of ionic saturation revisited. *Biopolymers.* 24: 427–439.

Hingerty, B. E., S. Figueroa, T. L. Hayden, and S. Broyde. 1989. Prediction of DNA structure from sequence: a build-up technique. *Biopolymers.* 28:1195–1222.

Hoffmann, G. R., and R. P. Fuchs. 1997. Mechanisms of frameshift mutations: insight from aromatic amines. *Chem. Res. Toxicol.* 10: 347–359.

Holbrook, S. R., R. E. Dickerson, and S. H. Kim. 1985. Anisotropic thermal-parameter refinement of the DNA dodecamer CGCGAATTCGCG by the segmented rigid-body method. *Acta Crystallogr. B.* 41:255–262.

Honig, B., and A. Nicholls. 1995. Classical electrostatics in biology and chemistry. *Science.* 268:1144–1149.

Hunter, C. A., and X. J. Lu. 1997. DNA base-stacking interactions: a comparison of theoretical calculations with oligonucleotide X-ray crystal structures. *J. Mol. Biol.* 265:603–619.

Ikeda, H., R. Fernandez, A. Wilk, J. J. Barchi, and V. E. Marqez. 1998. The effect of two antipodal fluorine-induced sugar puckers on the conformation and stability of the Dickerson–Drew dodecamer duplex [d(CGCGAATTCGCG)]$_2$. *Nucleic Acids Res.* 26:2237–2244.

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

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

Jayaram, B., D. Sprous, and D. L. Beveridge. 1998. Solvation free energy of biomacromolecules: parameters for a modified generalized Born model consistent with the AMBER force field. *J. Phys. Chem. B.* 102: 9571–9576.

Jayaram, B., K. J. McConnell, S. B. Dixit, and D. L. Beveridge. 1999. Free energy analysis of protein-DNA binding: the *Eco*RI endonuclease-DNA complex. *J. Comp. Phys.* 151:333–357.

Jin, R., and K. J. Breslauer. 1988. Characterization of the minor groove environment in a drug-DNA complex: bisbenzimide bound to the poly-[d(AT)].poly[d(AT)]duplex. *Proc. Natl. Acad. Sci. U.S.A.* 85: 8939–8942.

Kirkwood, J. G. 1939. The dielectric polarization of polar liquids. *J. Chem. Phys.* 7:911–919.

Kollman, P. A., I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case, and T. E. Cheatham, III. 2000. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* 33:889–897.

Lavery, R., and H. Sklenar. 1988. The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.* 6:63–91.

Lavery, R., and H. Sklenar. 1989. Defining the structure of irregular nucleic acids: conventions and principles. *J. Biomol. Struct. Dyn.* 6:655–667.

Lee, B., and F. M. Richards. 1971. The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* 55:379–400.

Liu, J., L. Malinina, T. Huynh-Dinh, and J. A. Subirana. 1998. The structure of the most studied DNA fragment changes under the influence of ions: a new packing of d(CGCGAATTCGCG). *FEBS Lett.* 438: 211–214.

Lorentz, H. A. 1952. Theory of Electrons. Dover, New York.

Mackerell, A. D., Jr., and N. Banavali. 2000. All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution. *J. Comp. Chem.* 21:105–120.

Manning, G. S. 1978. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Quart. Revs. Biophys.* 11:179–246.

Mao, B., B. E. Hingerty, S. Broyde, and D. J. Patel. 1998a. Solution structure of the aminofluorene [AF]-external conformer of the *anti*-[AF]-C8-dG adduct opposite dC in a DNA duplex. *Biochemistry.* 37: 95–106.

Mao, B., B. E. Hingerty, S. Broyde, and D. J. Patel. 1998b. Solution structure of the aminofluorene [AF]-intercalated conformer of the *syn*-[AF]-C8-dG adduct opposite dC in a DNA duplex. *Biochemistry.* 37: 81–94.

Mazur, A. K. 1998. Accurate DNA dynamics without accurate long-range electrostatics. *J. Am. Chem. Soc.* 120:10928–10937.

Mazur, J., and R. L. Jernigan. 1991. Distance-dependent dielectric constants and their application to double-helical DNA. *Biopolymers.* 31: 1615–1629.

McConnell, K. J., and D. L. Beveridge. 2000. DNA structure: what's in charge? *J. Mol. Biol.* 304:803–820.

Mehler, E. L. 1996. The Lorentz–Debye–Sack theory and dielectric screening of electrostatic effects in proteins and nucleic acids. *In* Molecular Electrostatic Potentials: Concepts and Applications. Theoretical and Computational Chemistry. J. S. Murray and K. Sen, editors. Elsevier Science B.V., New York. 371–405.

Mehler, E. L., and F. Guarnieri. 1999. A self-consistent, microenvironment modulated screened Coulomb potential approximation to calculate pH-dependent electrostatic effects in proteins. *Biophys. J.* 77:3–22.

Nicholls, A., K. A. Sharp, and B. Honig. 1990. Delphi. Department of Biochemistry and Molecular Biophysics, Columbia Univ., New York.

Olson, W. K., and P. J. Flory. 1972. Spatial configurations of polynucleotide chains. II. Conformational energies and the average dimensions of polyribonucleotides. *Biopolymers.* 11:25–56.

Olson, W. K. 1978. Spatial configuration of ordered polynucleotide chains. V. Conformatonal energy estimates of helical structure. *Biopolymers.* 17:1015–1040.

Olson, W. K. 1982. How flexible is the furanose ring? 2. An updated potential energy estimate. *J. Am. Chem. Soc.* 104:278–286.

Olson, W. K., and A. R. Srinivasan. 1990. Classical energy calculations of DNA, RNA and their constituents. In Landolt–Börnstein, Numerical Data and Functional Relationships in Science and Technology. New Series, Group VII, Biophysics 1d. W. Saenger, editor. Springer-Verlag, Berlin.415–433.

Olson, W. K., and V. B. Zhurkin. 2000. Modeling DNA deformations. *Curr. Opin. Struct. Biol.* 10:286–297.

Ornstein, R., and R. Rein. 1979. Energetics of intercalation specificity. I. Backbone unwinding. *Biopolymers*. 18:1277–1291.

Patel, D. J., B. Mao, Z. Gu, B. E. Hingerty, A. Gorin, A. K. Basu, and S. Broyde. 1998. Nuclear magnetic resonance solution structures of covalent aromatic amine-DNA adducts and their mutagenic relevance. *Chem. Res. Toxicol.* 11:391–407.

Pennock, B. E., and H. P. Schwan. 1969. Further observations on the electrical properties of hemoglobin-bound water. *J. Phys. Chem.* 73: 2600–2610.

Powell, M. J. D. 1964. Efficient method for finding minimum of function of several variables without calculating derivatives. *Comput. J.* 7:155–159.

Ramachandran, G. N., and R. Srinivasan. 1970. Effective dielectric constant values to be used in biopolymer energy calculations. *Indian J. Biochem.* 7:95–97.

Ramstein, J., and R. Lavery. 1988. Energetic coupling between DNA bending and base pair opening. *Proc. Natl. Acad. Sci. U.S.A.* 85: 7231–7235.

Rosenberg, I., J. F. Soler, Z. Tocik, W. Y. Ren, L. A. Ciszewski, P. Kois, K. W. Pankiewicz, M. Spassova, and K. A. Watanabe. 1993. Synthesis of oligodeoxynucleotides containing the C-nucleoside and 2′-deoxy-2′-fluoro-ara-nucleoside moieties by the H-phosphonate method. *Nucleosides Nucleotides.* 12:381–401.

Sack, V. H. 1926. The dielectric constant of electrolytes. *Phys. Z.* 27: 206–208.

Sack, V. H. 1927. The dielectric constants of solutions of electrolytes at small concentrations. *Phys. Z.* 28:199–210.

Saenger, W. 1984. Principles of Nucleic Acid Structure. Springer-Verlag, New York.

Sasisekharan, V. 1973. Conformation of polynucleotides. *Jerusalem Symp. Quantum Chem. Biol.* 5:247–260.

Sato, T. 1983. Another method for specifying furanose ring puckering. *Nucleic Acids Res.* 11:4933–4938.

Scott, R. A., and H. A. Scheraga. 1965. Method for calculating internal rotation barriers. *J. Chem. Phys.* 42:2209–2215.

Shapiro, R., S. Ellis, B. E. Hingerty, and S. Broyde. 1998. Effect of ring size on conformations of aromatic amine-DNA adducts: the aniline-C8 guanine adduct resides in the B-DNA major groove. *Chem. Res. Toxicol.* 11:335–341.

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

Shatsky-Schwartz, M., N. D. Arbuckle, M. Eisenstein, D. Rabinovich, A. Bareket-Samish, T. E. Haran, B. F. Luisi, and Z. Shakked. 1997. X-ray and solution studies of DNA oligomers and implications for the structural basis of A-tract-dependent curvature. *J. Mol. Biol.* 267:595–623.

Shui, X., L. McFail-Isom, G. G. Hu, and L. D. Williams. 1998a. The B-DNA dodecamer at high resolution reveals a spine of water on sodium. *Biochemistry.* 37:8341–8355.

Shui, X., C. C. Sines, L. McFail-Isom, D. VanDerveer, and L. D. Williams. 1998b. Structure of the potassium form of CGCGAATTCGCG: DNA deformation by electrostatic collapse around inorganic cations. *Biochemistry.* 37:16877–16887.

Sines, C. C., L. McFail-Isom, S. B. Howerton, D. VanDerveer, and L. D. Williams. 2000. Cations mediate B-DNA conformational heterogeneity. *J. Am. Chem. Soc.* 122:11048–11056.

Singh, S. B., B. E. Hingerty, U. C. Singh, J. P. Greenberg, N. E. Geacintov, and S. Broyde. 1991. Structures of the (+)- and (−)-*trans*-7,8-dihydroxy-*anti*-9,10-epoxy- 7,8,9,10-tetrahydrobenzo(a)pyrene adducts to guanine-N2 in a duplex dodecamer. *Cancer Res.* 51:3482–3492.

Sponer, J., and J. Kypr. 1993. Theoretical analysis of the base stacking in DNA: choice of the force field and a comparison with the oligonucleotide crystal structures. *J. Biomol. Struct. Dyn.* 11:277–292.

Srinivasan, A. R., and W. K. Olson. 1980. Polynucleotide conformation in real solution. A preliminary theoretical estimate. *Fed. Proceed.* 39:2199.

Srinivasan, J., M. W. Trevathan, P. Beroza, and D. A. Case. 1999. Application of a pairwise generalized Born model to proteins and nucleic acids: inclusion of salt effects. *Theor. Chem. Acc.* 101:426–434.

Still, W. C., A. Tempczyk, R. C. Hawley, and T. Hendrickson. 1990. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* 112:6127–6129.

Subirana, J. A., and T. Faria. 1997. Influence of sequence on the conformation of the B-DNA helix. *Biophys. J.* 73:333–338.

Takashima, S., and H. P. Schwan. 1965. Dielectric dispersion of crystalline powders of amino acids, peptides, and proteins. *J. Phys. Chem.* 69: 4176–4182.

Tereshko, V., G. Minasov, and M. Egli. 1999a. The Dickerson–Drew B-DNA dodecamer revisited at atomic resolution. *J. Am. Chem. Soc.* 121:470–471.

Tereshko, V., G. Minasov, and M. Egli. 1999b. A "hydrat-ion" spine in a B-DNA minor groove. *J. Am. Chem. Soc.* 121:3590–3595.

Tjandra, N., S. Tate, A. Ono, M. Kainosho, and A. Bax. 2000. The NMR structure of a DNA dodecamer in an aqueous dilute liquid crystalline phase. *J. Am. Chem. Soc.* 122:6190–6200.

Tsui, V., and D. A. Case. 2000. Molecular dynamics simulations of nucleic acids with a generalized Born solvation model. *J. Am. Chem. Soc.* 122:2489–2498.

Tsui, V., and D. A. Case. 2001. Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers.* 56: 275–291.

Westhof, E. 1987. Re-refinement of the B-dodecamer d(CGCGAATTCGCG) with a comparative analysis of the solvent in it and in the Z-hexamer d(5BrCG5BrCG5BrCG). *J. Biomol. Struct. Dyn.* 5:581–600.

Wing, R., H. Drew, T. Takano, C. Broka, S. Tanaka, K. Itakura, and R. E. Dickerson. 1980. Crystal structure analysis of a complete turn of B-DNA. *Nature.* 287:755–758.

Woods, K. K., L. Mcfail-Isom, C. C. Sines, S. B. Howerton, R. K. Stephens, and L. D. Williams. 2000. Monovalent cations sequester within the A-tract minor groove of $[d(CGCGAATTCGCG)]_2$. *J. Am. Chem. Soc.* 122:1546–1547.

Yan, S., R. Shapiro, N. E. Geacintov, and S. Broyde. 2001. Stereochemical, structural, and thermodynamic origins of stability differences between stereoisomeric benzo[a]pyrene diol epoxide deoxyadenosine adducts in a DNA mutational hot spot sequence. *J. Am. Chem. Soc.* 123: 7054–7066.

Yoon, C. N., Y. K. Kang, and M. S. Jhon. 1984. Conformational study of trinucleoside tetraphosphate d(pCpGpCp): transition of right-handed form to left-handed form. *Biopolymers.* 23:511–536.

Young, M. A., B. Jayaram, and D. L. Beveridge. 1998. Local dielectric environment of B-DNA in solution: results from a 14 ns molecular dynamics trajectory. *J. Phys. Chem. B.* 102:7666–7669.

Zhurkin, V. B., V. I. Poltev, and V. L. Florent'ev. 1981. Atom–atom potential functions for conformational calculations of nucleic acids. *Mol. Biol. USSR (Eng. Ed.).* 14:882–895.