ELSEVIER

Journal of Molecular Structure (Theochem) 537 (2001) 283–305

THEO CHEM

www.elsevier.nl/locate/theochem

Calibrating nucleic acids torsional energetics in force-field: insights from model compounds ☆

D. Bosch[a], N. Foloppe[b], N. Pastor[c], L. Pardo[a], M. Campillo[a,*]

[a]*Laboratori de Medicina Computacional, Unitat de Bioestadistica, Facultat de Medicina, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain*

[b]*Department of Bioscience, Center for Structural Biochemistry, Karolinska Institutet, S-141 57, Huddinge, Sweden*

[c]*Facultad de Ciencias, Universidad Autónoma del Edo. de Morelos, 62210 Cuernavaca, Morelos, Mexico*

Received 14 April 2000; accepted 28 June 2000

Abstract

The development of force fields that accurately describe both the structure and the dynamics of nucleic acids in condensed phase is an ongoing effort. The development of the latest versions of the CHARMM and AMBER nucleic acids relied on ab initio as well as on experimental target data for the parametrization. Here we compare the two latest versions of the AMBER and CHARMM force field, in their ability to reproduce the ab initio torsional energy surfaces for selected nucleic acid dihedral angles. A series of model compounds is instrumental in this analysis. This illustrates how dissecting the energetics of the force-field with model compounds allows to uncover deficiencies in the force-field, which may or may not be apparent in the simulated properties of the full nucleic acids. The positions of minima, the relative energies and barrier heights are discussed. This type of analysis is proposed as one useful diagnostic criterion, in combination with others, to assess how well balanced are the various contributions to the energetics of a nucleic acid force field.

Keywords: CHARMM; AMBER; Force field; DNA; RNA; Molecular mechanics

1. Introduction

Molecular dynamics (MD) simulations of nucleic acids are emerging as a powerful tool to study the structure, dynamics and energetics of these molecules in explicit solvent [1–17], but the reliability of the results obtained from MD simulations crucially depends on the quality of the underlying force field. This was recently illustrated with a number of simulations of duplex DNA, for which it was possible to show that the simulated properties are strongly influenced by the force field [18,19]. In particular, analysis of this force-field dependency of the results allowed to pinpoint to some specific shortcomings in the version 22 of CHARMM nucleic acid force field [20], and the AMBER Cornell et al. force field [21]. This led to the development of the subsequent CHARMM27 [22] and AMBER98 [23] force fields. The BMS nucleic acid force field was also developed recently [24]. The calibration of the AMBER98, BMS and CHARMM27 relied extensively on the correct representation of the structural properties of the simulated nucleic acids in condensed phase. For CHARMM27, both crystal and solution environments were used for

☆ Dedicated to Professor Serafín Fraga on the occasion of his 70th birthday.

* Corresponding author. Tel.: +34-93-581-2348; fax: +34-93-581-2344.

E-mail address: mercedes.campillo@uab.es (M. Campillo).

0166-1280/01/$ - see front matter
PII: S0166-1280(00)00685-0

condensed phase calibration, although only solution simulations were reported for the AMBER98 and BMS force field. The dependence of the DNA structural properties on the solvent was another criterion used in the BMS and CHARMM27 force field parametrization. The necessity to use the results of condensed phase simulations as a major test of the force field was illustrated by the development of CHARMM22, for which this test was not included [20]. Subsequent DNA simulations using CHARMM22 showed that the B form of DNA was not stable with that force field [4,11,15]. This deficiency has been corrected in CHARMM27 [22].

Using the structural properties of nucleic acids, when simulated in condensed phase, as a major test of the force field quality is important, but it has its limitations. First, this approach assumes that the structural properties of nucleic acids in condensed phase are well characterized experimentally. Although this is true, to a large extent, for high-resolution crystal structures, too few very high-resolution crystal structures are currently available for DNA [25,26]. Besides, crystallization is frequently obtained with non-physiological solvents, and it is well documented that the crystal structure for a given deoxyribo-oligonucleotide may be significantly influenced by crystal packing forces [27–30]. It has been argued that this dependence of deoxyribo-oligonucleotide structure on the crystal environment may be an opportunity to better understand some of the determining factors of DNA overall structure and flexibility [30]. In practice, it is not trivial, however, to separate what is contributed by the intrinsic properties of a specific DNA sequence from the influence of packing forces. This is a source of uncertainty in the context of force field development and testing, because it is unclear whether or not the structural information derived in the crystal environment is transferable to a solution situation, where most of the critical force field testing can be carried out. The usefulness of force field testing using simulations with the crystal environment is indeed questionable, because the packing forces may force the simulated structures to remain, overall, near their initial experimental coordinates. In this situation the simulation would not act as a test, and little information, if any, is gained concerning the quality of the force field. Therefore, having a number of well-characterized DNA structures in solution, including their dynamics, would be of great value as reference data for force field development. Obtaining this information has so far been elusive, although NMR has become increasingly powerful in deriving deoxyribo-oligonucleotide structures in solution [31–33]. However, the lack of long range distance restraints [34,35], in combination with a possible dependence of the NMR derived structures on the force field used in the refinement protocol, has plagued the characterization of DNA structure and dynamics in solution. These uncertainties severely limit the reliability of DNA force field testing by MD simulations in solution. Another, although related, limiting factor is the relatively small number of DNA sequences which have been structurally characterized so far. For instance, the widespread notion that DNA in solution populates only a well-defined number of discrete structural families was recently called into question [36]. This new insight suggests that we should resist the temptation to parametrize a force field so that it forces DNA to be confined in the already well characterized A, B and Z families. Finally, a major limitation of using condensed phase simulations as the only test for force field adjustments is that the results of these simulations do not specifically point to the type of parameters which have to be adjusted. In this situation, the force field parameters are adjusted in a fully empirical manner.

In view of the above arguments, it is necessary to supplement the force field parametrization process with another, less empirical component, which aims at dissecting what the individual contributions to the force field should be. This requires that the full polymer be broken down in a number of relevant building blocks, or model compounds, for which the intrinsic structural and energetic properties are investigated independently. Because nucleic acids contain many rotatable bonds per nucleotide, an important property to be studied at the model compound level is the intrinsic torsional energetics associated with each of these torsions. This was recently carried out by calculating ab initio the intrinsic torsional energy profiles corresponding to nucleic acids dihedrals β, γ, ϵ, and χ, using a series of model compounds [37]. The derived ab initio energetics was used to guide the development of the CHARMM27 force field, when results from condensed phase simulations suggested that some adjustments were needed. It must be

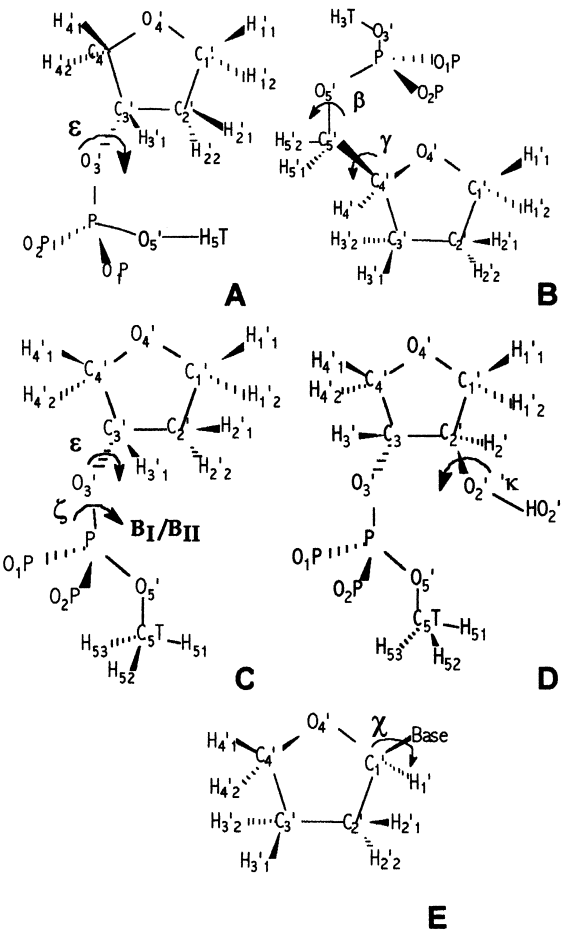

Fig. 1. Model compounds used to study the potential energy profile for the rotation about the C3′–O3′ (ϵ, compound A), O5′–C5′ (β, compound B), C5′–C4′ (γ, compound B), C1′–N1/N9 (χ, compound E), and C2′–O2′ (compound D) bonds. Compound C was used to investigate the energy difference between the B_I and B_{II} conformations.

stressed that the final CHARMM27 energetics was allowed to depart from its ab initio counterpart, to better fit the properties simulated in condensed phase to their experimental counterpart. These deviations between the molecular mechanics and the ab initio energetics may reflect limitations in the size of the model compounds, the level of theory applied in the ab initio calculations (no representation of solvation, for instance), the physics of the empirical energy function (no representation of polarizability, for instance), or simply the fact that the optimal parameters were not found yet. It should be realized, however, that these deviations may also reflect the limitations in the condensed phase target data (see above). In order to fit the properties of a given DNA sequence, in a particular environment, to condensed phased properties which were averaged over a large number or sequences and environments, one may inadvertently compromise the empirical energetics at the model compound level. It is therefore difficult to judge what should be the relative weights of the ab initio target data and the condensed phase target data, when calibrating a force field. Because the ab initio derived properties arguably represent the less empirical part of the target data, they should not be sacrificed without a clear justification.

The degree to which ab initio target data, for the torsional energetics, were used to calibrate the current nucleic acids force fields varies widely. The development of the BMS force field [24] emphasized the importance of a correct representation of the condensed phase properties, and did not apparently rely much on ab initio data. The torsional energetics was parametrized explicitely only for α and ζ only in CHARMM22 [20], and for α, ζ and χ only in AMBER96 [21]. The details of the GROMOS nucleic acids force field are, to our knowledge, not yet published elsewhere than in the GROMOS96 [38] manual and user guide, making it awkward to comment on what was done. In CHARMM27 the energetics of all nucleic acid torsions was parametrized explicitly, using a series of model compound designed for this purpose [22]. These model compounds and the associated ab initio data were not, however, available for the parametrization of the other above-mentioned force fields. In the present work, we use the same model compounds to investigate the corresponding energetics in AMBER96, AMBER98 and CHARMM22. A comparison of the CHARMM27 energy profiles to the ab initio reference was already presented elsewhere [22], but it is included here to allow for comparison with the AMBER and CHARMM22 results. Indeed, one aim of present work is to compare side by side the energy profiles obtained from the CHARMM and AMBER force fields, in relationship to their ab initio counterpart. It is of great interest to compare CHARMM27 to AMBER98, which are currently two of the most widely used force-fields for nucleic acid molecular

dynamics (MD) simulations. Also, a comparison of the CHARMM22 and CHARMM27 energetics should be of interest to analyse, in hindsight, some of the shortcomings of CHARMM22, and understand if something of general interest for nucleic acid force-field development can be learned that exercise. A comparison of CHARMM22 and CHARMM27, at the model compound level, was not attempted during the presentation of CHARMM27 [22].

2. Methods

2.1. Model compounds and constraints

Fig. 1 shows the model compounds used to study: (i) the potential energy profile of torsions ϵ (compound A), β (compound B), γ (compound B), χ (compound E with adenine, guanine, cytosine and thymine), and C3′–C2′–O2′–HO2′ (compound D); (ii) the energy difference between the B_I (ϵ in *trans*; ζ in *gauche*$^-$) and B_{II} (ϵ in *gauche*$^-$; ζ in *trans*) conformations (compound C); and (iii) the relative energies of the A-DNA and B-DNA like conformations at the nucleoside level (compound E). We use the atom names and dihedral angle nomenclature corresponding to their nucleic acid counterpart [39]. A detailed presentation of the design of the model compounds A–E and the applied constraints can be found in Foloppe and MacKerell [37,40], but the most important points are briefly outlined here. (i) The potential energy profiles were calculated with the furanose ring of model compounds A, B, and E constrained (see below) to either the C2′*endo* (B-DNA) or C3′*endo* (A-DNA) conformations. Because compound D is related to RNA, its furanose was only constrained to the C3′*endo* pucker. Only the pseudorotation angle was constrained, but not the puckering amplitude. This was obtained by fixing the C3′–C4′–O4′–C1′ (C2′*endo*) or the C4′–O4′–C1′–C2′ (C3′*endo*) dihedral to 0.0°. Additional constraints on the dihedral angles α (291° for C3′*endo* and 298° for C2′*endo*), β (175 and 168°), γ (57 and 51°), ϵ (205 and 187°), ζ (287 and 262°), and χ (199 and 252°) were used, where applicable [37], during the energy optimization process, with the ab initio and molecular mechanics (AMBER and CHARMM) methodologies (see below). These correspond to the previously obtained (see [37, Table 1]) modal values of α, β, γ, ϵ, ζ and χ in crystal structures of A-DNA and B-DNA. The ϵ, γ, χ and C3′–C2′–O2′–HO2′ torsion energy profiles were sampled from 0 to 330° in a 30° increment, whereas β was sampled from 90 to 330° in a 30° increment. The exact location of each energy minima was obtained by further minimization of the dihedral of interest. (ii) The furanose ring was not constrained when modelling the energy difference between B_I and B_{II} substates with compound C. (iii) The energy minimizations at the nucleoside level were first performed with χ left unconstrained with either a north or a south furanose pucker. These are referred as the north and south energy minima, respectively. Subsequent minimizations were carried out with χ constrained to an A-DNA like conformation ($\chi = 201.1°$) and a B-DNA like conformation ($\chi = 258.1°$). $\Delta E_{A\text{-}n}$ is the difference in energy between the A-DNA like conformation and the north minimum; $\Delta E_{B\text{-}s}$ between B-DNA like conformation and the south minimum; $\Delta E_{B\text{-}A}$ between the B-DNA and A-DNA like conformations; and $\Delta E_{n\text{-}s}$ between the north and south minima.

2.2. Ab initio

The ab initio data on compounds A, B, C and E discussed in this work were obtained in a previous study [37,40], after geometry optimizations at the second order Møller–Plesset (MP2) level of theory, using the 6-31G* and the 6-31+G* basis sets, for neutral and anionic compounds, respectively. As part of the present work, additional ab initio calculations were carried out with the GAUSSIAN 98 suite of programs [41] and its default tolerances, to investigate the C3′–C2′–O2′–HO2′ torsional energetics in compound D. The corresponding geometries of compound D were first optimized at the restricted Hartree–Fock (HF) level of theory with the 6-31+G** basis set. These geometries were then used to derive single point MP2 energies with the 6-31+G** basis set (MP2/6-31+G**//HF/6-31+G**). Only these later results are presented and discussed. All ab initio calculations were carried out under the constraints detailed above.

2.3. AMBER

The energy optimized ab initio structures (see above)

were employed as starting points for energy minimization to the default tolerances (no truncation of nonbonded interactions, dielectric constant of 1.0) with the Sander module of AMBER5 [42]. All constraints on dihedrals (see above) were applied using flat harmonic restraints ($\pm 0.01°$, 5000 kcal mol^{-1} rad^{-1}). All the energy profiles were calculated with both the all atom force field developed by Cornell et al. [21] (named parm94.dat or AMBER96), and the latest modification of this parameter set [23] (parm98.dat or AMBER98). The two parameter sets differ by the periodicity, the peak height, and the phase of the OS–CT–CT–OS, OH–CT–CT–OS, OH–CT–CT–OH, OS–CT–N*–CK and OS–CT–N*–CM torsional potentials. Besides, AMBER98 adds a specific CT–OS–CT–N* torsional potential (Table 5 in Cheatham et al. [23]). Only the calculations carried out with AMBER98 are presented, because their yield results similar to those obtained with AMBER96 (see Section 3). The atom centered point charges were derived by restrained electrostatic potential (RESP) fits [21], with model compounds A, B, C, and E in the C2′*endo* conformation, and D in the C3′*endo* conformation. The hydrogens bonded to the same heavy atom were restrained to have equal charge. An additional restraint of 0.1216 was applied to the C1′ atom and its two hydrogens, in model compounds A–D which lack a base. This value represents the total charge of the base plus C1′ and H1′ in the Cornell et al. force field [21]. This procedure provides atomic charges in the sugar-backbone moieties that do not deviate significantly from the charges reported in the Cornell et al. force field [21]. Appendices A–C show the atomic RESP charges used in the present work.

2.4. CHARMM

The program CHARMM [43] was used in combination with the associated versions 23 [20] and 27 [22] of its nucleic acid force field, referred to as CHARMM22 and CHARMM27, respectively. The details of the calculations carried out with CHARMM27 were already described [22], and the same protocol was used in the present work with CHARMM22. The calculations were carried out with no truncation of nonbonded interactions, and a dielectric constant of 1.0. Energy minimizations were performed with 50 steps of steepest descent, followed by 500 steps of adopted-basis Newton–Raphson, and 50 steps of Newton–Raphson, to a final energy gradient $\leq 10^{-6}$ kcal/mol/Å. Energy surfaces were calculated by harmonically constraining the selected dihedral with a force constant of 10 000 kcal/mol/degree^2.

For CHARMM27 atom types and partial charges, see Foloppe et al. [22]. The CHARMM22 atom types and partial charges for model compounds A-E were assigned following the same guidelines as in Foloppe et al. [22], that is the atom types and partial charges are as similar as possible in the model compounds as in their nucleic acid counterpart. Although these atom types and partial charges may not be optimal for the model compound itself (considered independently of nucleic acids), this strategy provides a more direct and practical relationship between the energetics derived at the model compound level, and its transferability to the full polymer. The CHARMM22 partial charges for model compounds A–E were transferred from the relevant building blocks, as explained in MacKerell et al. [20]. The partial charges on the phospho-ester moieties (compounds A-B) and the phospho-diester moieties (compounds C–D) were as in Fig. 5 of MacKerell et al. [20]. The charges on the furanose moiety (compounds A–E) and on the bases (compound E) were as in Fig. 2 and 10 of MacKerell et al. [20], respectively. When these various building blocks are assembled to form a larger model compound, a chemical bond has to be formed between them, which typically replaced two previous bonds involving hydrogen atoms. Doing that, the partial charges that were initially assigned to these hydrogens are transferred to the non-hydrogen atoms to which the hydrogens were bound. The charges for the nucleosides were exactly those already present in CHARMM22 and CHARMM27.

3. Results and discussion

We present the intrinsic energetics of the ϵ, β, γ, and χ torsions, the torsional energetics about the C2′–O2′ bond in RNA, the contribution of the torsional energetics to the relative energies of the B_I and B_{II} substates in DNA, and the relative energies of the A-DNA and B-DNA like conformations. These

torsions are investigated with model compounds A-E (Fig. 1), for which the ab initio energetics was already presented and discussed elsewhere [37,40], except for the torsional energetics about the C2′–O2′ bond in compound D. We also present the relative intrinsic conformational energies, at the nucleoside level, of the A-DNA and B-DNA like conformations. In the following, in each section we present the ab initio reference data first, the AMBER data second, and the CHARMM22 and CHARMM27 data third. For convenience, the previously obtained relevant ab initio results are briefly summarized, to facilitate comparison with their molecular mechanics counterparts.

All the AMBER calculations were performed with both the AMBER96 and AMBER98 force fields (see Section 2). The differences in torsional parameters between both force fields influence the potential energy surfaces for dihedrals γ, χ, and C3′–C2′–O2′–HO2′. However, the overall shape of the energy profiles calculated for compounds A-E, with AMBER96 and AMBER98, are similar to a large extent. The largest deviations in the location of the minima and the energy difference between them are 1.6° and 1.0 kcal/mol for γ, 10.8° and 0.5 kcal/mol for χ, and 0.5° and 0.5 kcal/mol for C3′–C2′–O2′–HO2′. Therefore, only the torsional energy profiles calculated with AMBER98 are reported for compounds A-E. However, the relative energies of the A-DNA and B-DNA like conformations, at the nucleoside level, obtained with AMBER96 and AMBER98 are both reported. The differences between CHARMM22 and CHARMM27 being very significant, the results obtained with both versions are compared for all compounds.

The torsional energy profiles of compounds A-E are characterized by the value of the investigated dihedral in its energy minima (location of the minima), and the energy difference between each secondary minimum and the global minimum (ΔE) in a given energy profile. These properties are listed in Tables 1–11. The discussion emphasizes the similarities and deviations between the ab initio energy profiles and their molecular mechanics counterpart. This is assessed by analyzing the difference in location ($\Delta°$) between the molecular mechanics energy minima relative to their ab initio counterpart, and the difference between the relative energies ($\Delta\Delta E$) separating two minima, in a force field on one hand, and in the ab initio on the other hand.

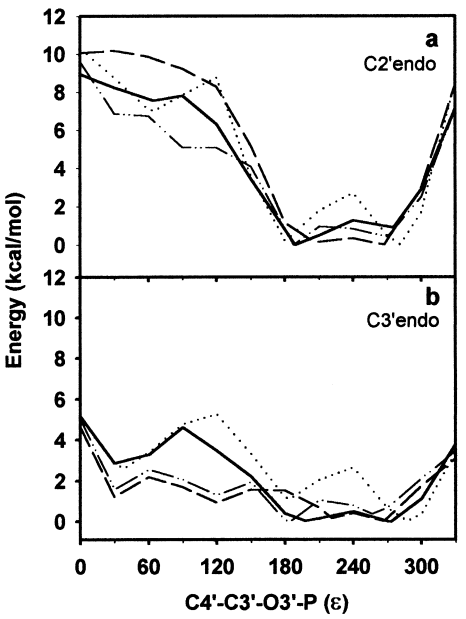

Fig. 2. Potential energy profiles of ϵ in model compound A obtained ab initio (solid line), with AMBER (dotted line), CHARMM22 (dashed line) and CHARMM27 (line-dot-dot-line), for (a) C2′*endo* and (b) C3′*endo* conformations of the furanose ring. Each torsional profile was offset relative to its global energy minimum.

3.1. Torsion angle C4′–C3′–O3′–P (ϵ)

Fig. 2 shows the previously obtained [37] ab initio potential energy profile for torsion ϵ, obtained with model compound A (Fig. 1) with the furanose ring in both the C2′*endo* and C3′*endo* conformations. The ϵ potential energy profiles display three energy minima in the *gauche*$^+$, *trans* and *gauche*$^-$ regions for both conformations of the furanose (Table 1). The ab initio minimum of lowest energy for the C2′*endo* conformation is in the *trans* region at 188.9° (relative energies, ΔE, between minima are reported in Table 1), the minimum in the *gauche*$^-$ region at 275.4° being only 0.9 kcal/mol higher in energy. In contrast, the energy minimum in the *gauche*$^+$ region at 64.3° is higher in energy by 7.6 kcal/mol. When the furanose is C3′*endo* both ab initio minima in the *trans* (197.6°)

Table 1
Energy minima of the torsional profile of ϵ (Fig. 2) and energy differences among minima (ΔE) calculated ab initio (in bold), with AMBER, CHARMM22, and CHARMM27, for the C2′*endo* and C3′*endo* conformations of the furanose ring. Numbers in parenthesis show the difference in position (Δ°) and energy ($\Delta\Delta E$) relative to ab initio

Pucker	Method	*gauche*$^+$		*trans*				*gauche*$^-$	
		ϵ	ΔE	ϵ	ΔE	ϵ	ΔE	ϵ	ΔE
C2′*endo*	Ab initio[a]	**64.3°**	**7.6**	–	–	**188.9°**	**0.0**	**275.4°**	**0.9**
	AMBER	65.2°	7.0	–	–	180.8°	0.2	281.7°	0.0
		(*0.9°*)	(*0.6*)			(*8.1°*)	(*0.2*)	(*6.3°*)	(*0.9*)
	CHARMM22	–	–	–	–	205.6°	0.1	267.5°	0.0
						(*16.7°*)	(*0.1*)	(*7.9°*)	(*0.9*)
	CHARMM27	–	–	–	–	187.8°	0.0	260.3°	0.3
						(*1.1°*)	(*0.0*)	(*15.1°*)	(*0.6*)
C3′*endo*	Ab initio[a]	**40.4°**	**2.9**	–	–	**197.6°**	**0.03**	**273.3°**	**0.0**
	AMBER	38.7°	2.6	–	–	184.6°	1.1	287.7°	0.0
		(*1.7°*)	(*0.3*)			(*13.0°*)	(*1.1*)	(*14.4°*)	(*0.0*)
	CHARMM22	30.0°	1.2	120.0	0.9	221.1°	0.1	267.0°	0.0
		(*10.4°*)	(*1.7*)	–	–	(*23.5°*)	(*0.1*)	(*6.3°*)	(*0.0*)
	CHARMM27	30.0°	1.6	120.0	1.3	184.4°	0.0	256.6°	0.4
		(*10.4°*)	(*1.3*)	–	–	(*13.2°*)	(*0.0*)	(*16.7°*)	(*0.4*)

[a] Adapted from Ref. [37].

and *gauche*$^-$ (273.3°) regions have similar energies. The C3′*endo*/*gauche*$^+$ minimum (40.4°) is higher in energy than the other minima by only 2.9 kcal/mol, suggesting that this *gauche*$^+$ minimum might be more easily accessible if the furanose ring adopts the C3′*endo* conformation.

3.1.1. AMBER

The ϵ torsional energy profiles in AMBER (Fig. 2) are also characterized by three energy minima at *gauche*$^+$ ($\Delta° = 0.9°$), *trans* ($\Delta° = 8.1°$), and *gauche*$^-$ ($\Delta° = 6.3°$) for the C2′*endo* conformation; and at *gauche*$^+$ ($\Delta° = 1.7°$), *trans* ($\Delta° = 13.0°$), and *gauche*$^-$ ($\Delta° = 14.4°$) for the C3′*endo* conformation (Table 1). Thus, the values of ϵ corresponding to the minima in the AMBER energy profiles are in reasonable agreement with the ab initio calculations. However, there are some discrepancies regarding the relative energies of the AMBER minima, as compared to the ab initio. In contrast to the ab initio, the AMBER C2′*endo*/*gauche*$^-$ minimum is lower in energy than the C2′*endo*/*trans* minimum ($\Delta E =$ 0.2 kcal/mol). This departure from the ab initio, however, does not compromise the relative energies of the B_I and B_{II} conformations in the larger compound C (see below). The C2′*endo*/*gauche*$^+$ minimum is properly placed 6.8 kcal/mol higher in energy than the C2′*endo*/*trans* conformation.

With a C3′*endo* pucker, AMBER *gauche*$^-$ minimum is also lower in energy than the *trans* minimum ($\Delta E = 1.1$ kcal/mol, Table 1), whereas these two minima are of almost identical energy in the ab initio calculations (see above). The distribution of ϵ in crystal structures of A-DNA duplexes is in the *trans* range [44]. The effect of AMBER favoring the C3′*endo*/*gauche*$^-$ over the C3′*endo*/*trans* conformation, in the context of its representation of A-DNA, is not clear yet. The AMBER C3′*endo*/*gauche*$^+$ minimum is 1.5 kcal/mol higher in energy than the *trans* conformation.

Another important property is the energy barrier between the $\epsilon =$ *trans* and $\epsilon =$ *gauche*$^-$ minima, because it is likely to be involved in the equilibrium between the B_I and B_{II} substates. This is clearly higher in AMBER than in the ab initio results, for both the C2′*endo* and C3′*endo* puckers (Fig. 2). In the ab initio structure corresponding to the energy maximum at $\epsilon = 240°$, the C3′···O1P and H3′···O1P distances are 2.88 and 2.28 Å (C2′*endo*) and 2.95 and 2.45 Å (C3′*endo*), respectively. These distances may indicate the existence of some hydrogen bond character in the interaction between the C3′–H3′ donor group and the

Table 2
Selected AMBER and CHARMM descriptors for BI/BII and energy differences between B_{II} and B_I conformations using compound C as model, calculated ab initio (in bold), with AMBER, CHARMM22, and CHARMM27. $\Delta E(B_{II} - B_I)$ is the energy difference between the B_{II}-like minimum and the B_I-like minimum

Method	Epsilon(ϵ)		Zeta (ζ)		Pseudo-rotation angle		$\Delta E(B_{II} - B_I)$
	B_I	B_{II}	B_I	B_{II}	B_I	B_{II}	
Ab initio	**194.3°**	**267.4°**	**274.2°**	**161.4°**	**152.3°**	**159.1°**	**1.5**
AMBER96	181.6°	269.2°	265.3°	168.1°	137.3°	120.9°	1.5
	(*12.7°*)	(*2.2°*)	(*8.9°*)	(*6.7°*)	(*15.0°*)	(*38.2°*)	(*0.0*)
AMBER98	181.3°	268.3°	265.6°	167.3°	142.3°	126.1°	1.5
	(*13.0°*)	(*0.9°*)	(*8.6°*)	(*5.9°*)	(*10.0°*)	(*33.0°*)	(*0.0*)
CHARMM22[a]	–	268.4°	–	185.2°	–	129.9°	0.5
		(*1.0°*)		(*23.8°*)		(*29.2°*)	(*1.0*)
CHARMM27	188.2°	261.0°	259.0°	183.7°	150.6°	143.2°	1.1
	(*6.1°*)	(*6.4°*)	(*15.2°*)	(*22.3°*)	(*1.7°*)	(*15.9°*)	(*0.4*)

[a] The ϵ and ζ values for the B_I conformation in CHARMM22 are not reported because this conformation is not an energy minimum with CHARMM22 and compound C.

O1P acceptor. This type of C–H···O hydrogen bonds is supported by both experimental and theoretical data [45–48]. AMBER is not able to reproduce this type of interaction in the same manner as the MP2 treatment of correlation effects with the 6-31+G* basis set.

3.1.2. CHARMM

Fig. 2 shows the CHARMM22 and CHARMM27 energy profiles for the ϵ torsion. In contrast to the three ab initio minima (see above), both the CHARMM22 and CHARMM27 C2′*endo* energy profiles display only two energy minima (Table 1). The C2′*endo*/*gauche*$^+$ minimum is present neither in CHARMM22 nor CHARMM27. The two minima in *trans* ($\Delta° = 16.7°$ and $\Delta° = 1.1°$ for CHARMM22 and CHARMM27, respectively) and *gauche*$^-$ ($\Delta° = 7.9°$ and $\Delta° = 15.1°$) are in reasonable agreement with the ab initio calculations. Thus, the location of the C2′*endo*/*trans* minimum is significantly closer to the ab initio in CHARMM27 than in CHARMM22, although the trend is opposite for the C2′*endo*/*gauche*$^-$ minimum. This represents an improvement in CHARMM27 over CHARMM22 because the $\epsilon =$ *trans* conformation is much more populated in nucleic acids than the $\epsilon =$ *gauche*$^-$ conformation. CHARMM27 properly estimates the *trans* minimum as lower in energy than the *gauche*$^-$ minimum ($\Delta E =$ 0.3 kcal/mol and $\Delta\Delta E = 0.6$ kcal/mol, Table 1), following the ab initio trend. This energy difference is better treated in CHARMM27 than in AMBER or CHARMM22 (see Table 1). Moreover, the CHARMM27 barrier for the ϵ transition from *trans* to *gauche*$^-$ is of the same order as the ab initio value. A proper treatment of this barrier may be important to reproduce the rate of interconversion between the B_I and B_{II} substates.

The C3′*endo* energy profile contains four energy minima at ~30°, ~120° (not observed in ab initio),

Table 3
Energy minima of the torsional profile of β (Fig. 3) calculated ab initio (in bold), with AMBER, CHARMM22, and CHARMM27, for the C2′*endo* and C3′*endo* conformations of the furanose ring. Numbers in parenthesis show the difference in position ($\Delta°$) and energy ($\Delta\Delta E$) relative to ab initio

Pucker	Method	β (*anti*)
C2′*endo*	Ab initio[a]	**238.2°**
	AMBER	186.7°
		(*51.6°*)
	CHARMM22	232.5°
		(5.7°)
	CHARMM27	235.0°
		(*3.2°*)
C3′*endo*	Ab initio[a]	**240.6°**
	AMBER	183.8°
		(*56.8°*)
	CHARMM22	232.2°
		(*8.4°*)
	CHARMM27	188.2°
		(*52.4°*)

[a] Adapted from Ref. [37].

Table 4
Energy minima of the γ torsional profile (Fig. 4) and energy differences between minima (ΔE) calculated ab initio (in bold), with AMBER, CHARMM22, and CHARMM27, for the C2′*endo* and C3′*endo* conformations of the furanose ring. Numbers in parenthesis show the difference relative in position (Δ°) and energy ($\Delta\Delta E$) relative to ab initio

Pucker	Method	*gauche*$^+$		*trans*		*gauche*$^-$	
		γ	ΔE	γ	ΔE	γ	ΔE
C2′*endo*	Ab initio[a]	**44.7°**	**0.0**	**182.6°**	**5.0**	**292.6°**	**0.3**
	AMBER	41.1°	1.5	201.5°	2.4	295.9°	0.0
		(*3.6°*)	(*1.5*)	(*18.9°*)	(*2.6*)	(*3.3°*)	(*0.3*)
	CHARMM22	44.1°	6.4	195.6°	5.8	292.3°	0.0
		(*0.6°*)	(*6.4*)	(*13.0°*)	(*0.8*)	(*0.3°*)	(*0.3*)
	CHARMM27	25.7°	1.3	–	–	292.1°	0.0
		(*19.0°*)	(*1.3*)			(*0.5°*)	(*0.3*)
C3′*endo*	Ab initio[a]	**45.9°**	**0.7**	**189.9°**	**3.6**	**297.4°**	**0.0**
	AMBER	42.1°	1.3	202.1°	2.2	299.0°	0.0
		(*3.8°*)	(*0.6*)	(*12.2°*)	(*1.4*)	(*1.6°*)	(*0.0*)
	CHARMM22	46.2°	5.8	195.9°	5.5	294.6°	0.0
		(*0.3°*)	(*5.1*)	(*6.0*)	(*1.9*)	(*2.8°*)	(*0.0*)
	CHARMM27	21.0°	0.8	–	–	297.7°	0.0
		(*24.9°*)	(*0.1*)			(*0.3°*)	(*0.0*)

[a] Adapted from Ref. [37].

Table 5
Energy minima of the χ torsional profile for compound E with adenine (Fig. 7a and b) and energy differences between minima (ΔE) calculated ab initio (in bold), with AMBER, CHARMM22, and CHARMM27, for the C2′*endo* and C3′*endo* conformations of the furanose ring. Numbers in parenthesis show the difference in position (Δ°) and energy ($\Delta\Delta E$) relative to ab initio

Pucker	Method	*syn*		*anti*	
		χ	ΔE	χ	ΔE
C2′*endo*	Ab initio[a]	**65.5°**	**0.5**	**189.9°**	**0.0**
	AMBER	54.8°	3.9	205.2°	0.0
		(*10.7°*)	(*3.4*)	(*15.3°*)	(*0.0*)
	CHARMM22	70.6°	5.1	186.4°	0.0
		(*5.1°*)	(*4.6*)	(*3.5°*)	(*0.0*)
	CHARMM27	63.1°	3.1	194.3°	0.0
		(*2.4°*)	(*2.6*)	(*4.4°*)	(*0.0*)
C3′*endo*	Ab initio[a]	**75.4°**	**1.2**	**184.1°**	**0.0**
	AMBER	41.7°	4.9	203.7°	0.0
		(*33.7°*)	(*3.7*)	(*19.6°*)	(*0.0*)
	CHARMM22	–	–	186.4°	0.0
				(*2.3°*)	(*0.0*)
	CHARMM27	104.9°	3.6	191.1°	0.0
		(*29.5°*)	(*2.4*)	(*7.0°*)	(*0.0*)

[a] Adapted from Ref. [37].

Table 6
Energy minima of the χ torsional profile for compound E with guanine (Fig. 7c and d) and energy differences between minima (ΔE) calculated ab initio (in bold), with AMBER, CHARMM22, and CHARMM27, for the C2′*endo* and C3′*endo* conformations of the furanose ring. Numbers in parenthesis show the difference in position (Δ°) and energy ($\Delta\Delta E$) relative to ab initio

Pucker	Method	*syn*		*anti*	
		χ	ΔE	χ	ΔE
C2′*endo*	Ab initio[a]	**62.1°**	**0.6**	**198.5°**	**0.0**
	AMBER	50.7°	2.1	211.0°	0.0
		(*11.4°*)	(*1.5*)	(*12.5°*)	(*0.0*)
	CHARMM22	90.0°	8.0	186.3°	0.0
		(*27.9°*)	(*7.4*)	(*12.2°*)	(*0.0*)
	CHARMM27	60.5°	2.5	213.3°	0.0
		(*1.6°*)	(*1.9*)	(*14.8°*)	(*0.0*)
C3′*endo*	Ab initio[a]	**70.4°**	**1.9**	**187.7°**	**0.0**
	AMBER	35.4°	2.6	207.2°	0.0
		(*35.0°*)	(*0.7*)	(*19.5°*)	(*0.0*)
	CHARMM22	–	–	186.3°	0.0
				(*1.4°*)	(*0.0*)
	CHARMM27	43.4°	3.9	198.5°	0.0
		(*27.0°*)	(*2.0*)	(*10.8°*)	(*0.0*)

[a] Adapted from Ref. [37].

Table 7
Energy minima of the torsional profile of χ for cytosine (Fig. 8a and b) and energy differences among minima (ΔE) calculated ab initio (in bold), with AMBER, CHARMM22, and CHARMM27, for the C2′*endo* and C3′*endo* conformations of the furanose ring. Numbers in parenthesis show the difference in position ($\Delta°$) and energy ($\Delta\Delta E$) relative to ab initio

Pucker	Method	*syn*		*anti*	
		χ	ΔE	χ	ΔE
C2′*endo*	Ab initio[a]	**61.6°**	**4.0**	**194.1°**	**0.0**
	AMBER	56.3°	8.0	210.1°	0.0
		(*5.3°*)	(*4.0*)	(*16.0°*)	(*0.0*)
	CHARMM22	65.1°	6.1	221.6°	0.0
		(*3.5°*)	(*2.1*)	(*27.5°*)	(*0.0*)
	CHARMM27	63.9°	6.0	193.9°	0.0
		(*2.3°*)	(*2.0*)	(*0.2°*)	(*0.0*)
C3′*endo*	Ab initio[a]	**66.6°**	**4.2**	**191.5°**	**0.0**
	AMBER	62.0°	9.4	203.7°	0.0
		(*4.6°*)	(*5.2*)	(*12.2°*)	(*0.0*)
	CHARMM22	98.9°	5.5	214.5°	0.0
		(*32.3°*)	(*1.3*)	(*23.0°*)	(*0.0*)
	CHARMM27	99.8°	6.3	192.1°	0.0
		(*33.2°*)	(*2.1*)	(*0.6°*)	(*0.0*)

[a] Adapted from Ref. [37].

Table 8
Energy minima of the χ torsional profile of compound E with thymine (Fig. 8c and d) and energy differences between minima (ΔE) calculated ab initio (in bold), with AMBER, CHARMM22, and CHARMM27, for the C2′*endo* and C3′*endo* conformations of the furanose ring. Numbers in parenthesis show the difference in position ($\Delta°$) and energy ($\Delta\Delta E$) relative to ab initio

Pucker	Method	*syn*		*anti*	
		χ	ΔE	χ	ΔE
C2′*endo*	Ab initio[a]	**60.7°**	**2.8**	**202.6°**	**0.0**
	AMBER	50.8°	5.0	213.4°	0.0
		(*9.9°*)	(*2.2*)	(*10.8°*)	(*0.0*)
	CHARMM22	60.6°	4.3	204.2°	0.0
		(*0.1°*)	(*1.5*)	(*1.6°*)	(*0.0*)
	CHARMM27	60.6°	4.3	204.2°	0.0
		(*0.1°*)	(*1.5*)	(*1.6°*)	(*0.0*)
C3′*endo*	Ab initio[a]	**66.6°**	**3.1**	**192.8°**	**0.0**
	AMBER	40.6°	6.3	210.0°	0.0
		(*26.0°*)	(*3.2*)	(*17.2°*)	(*0.0*)
	CHARMM22	101.5°	5.0	215.2°	0.0
		(*34.9°*)	(*1.9*)	(*22.4°*)	(*0.0*)
	CHARMM27	101.4°	5.2	197.6°	0.0
		(*34.8°*)	(*2.1*)	(*4.8°*)	(*0.0*)

[a] Adapted from Ref. [37].

trans ($\Delta° = 23.5°$), and *gauche*$^-$ ($\Delta° = 6.3°$) for CHARMM22; and at ~30°, ~120° (not observed in ab initio), *trans* ($\Delta° = 13.2°$), and *gauche*$^-$ ($\Delta° = 16.7°$) for CHARMM27 (Table 1). That the minimum at the eclipsed 120° conformation is not observed in the ab initio results is a strong indication that its presence in the CHARMM profiles is artefactual. This conclusion is reinforced by the fact that ϵ does not populated the 120° range in crystal structures of nucleosides or nucleotides [49]. The energy differences between C3′*endo*/*trans* and C3′*endo*/*gauche*$^-$ minima with CHARMM22 (0.1 kcal/mol) and CHARMM27 (0.4 kcal/mol) are in reasonable agreement with their ab initio counterpart (0.0 kcal/mol), as is the barrier between these energy minima.

3.2. Energy difference between B_I and B_{II} conformations

Table 2 shows the values of ϵ, ζ, the pseudorotation angle, and the energy difference between the B_I (ϵ in *trans*; ζ in *gauche*$^-$) and B_{II} (ϵ in *gauche*$^-$; ζ in *trans*) substates of the phosphodiester linkage of B-DNA, obtained ab initio with compound C (Fig. 1). The values of ϵ, ζ and pseudorotation angle for the B_I conformation are 194.3°, 274.2°, and 152.3°; and for the B_{II} conformation are 267.4°, 161.4° and 159.1°, respectively. The B_I conformation was found 1.5 kcal/mol lower in energy than the B_{II}.

Table 9
Energy minima of the torsional profile of HO2′–O2′–C2′–C3′ (Fig. 10) calculated ab initio (in bold), with AMBER, CHARMM22, and CHARMM27, for the C2′*endo* and C3′*endo* conformations of the furanose ring. Numbers in parenthesis show the difference in position ($\Delta°$) relative to ab initio

Method	HO2′–O2′–C2′–C3′ (gauche$^-$)
Ab initio[a]	**330.6°**
AMBER	327.4°
	(*3.2°*)
CHARMM22	333.0°
	(*2.4°*)
CHARMM27	324.6°
	(*6.0°*)

[a] Single point MP2 energies from structures energy minimized at the HF/6-31+G** level.

Table 10
Descriptors related to the structure and energetics of the pyrimidine nucleosides with χ in A and B DNA-like conformations. $\Delta E_{A\text{-}n}$ (kcal/mol) is the energy of the A DNA-like conformation minus the energy of the north energy minimum. $\Delta E_{B\text{-}s}$ (kcal/mol) is the energy of the B DNA-like conformation minus the energy of the north energy minimum. $\Delta E_{B\text{-}A}$ (kcal/mol) is the energy of the B DNA-like conformation minus the energy of the A DNA-like conformation. $\Delta E_{n\text{-}s}$ (kcal/mol) is the north minimum minus the energy of the south minimum. Numbers in parenthesis show the difference in energy ($\Delta\Delta E$) relative to ab initio

Method	Cytosine				Thymine			
	$\Delta E_{A\text{-}n}$	$\Delta E_{B\text{-}s}$	$\Delta E_{B\text{-}A}$	$\Delta E_{n\text{-}s}$	$\Delta E_{A\text{-}n}$	$\Delta E_{B\text{-}s}$	$\Delta E_{B\text{-}A}$	$\Delta E_{n\text{-}s}$
Ab initio[a]	**0.2**	**2.2**	**2.3**	**−0.3**	**<0.1**	**1.2**	**0.3**	**0.9**
AMBER96	0.2	3.3	2.3	0.7	0.3	2.7	1.8	0.5
	(*0.0*)	(*1.1*)	(*0.0*)	(*1.0*)	(*0.3*)	(*1.5*)	(*1.5*)	(*0.4*)
AMBER98	0.5	2.8	1.5	0.7	0.7	2.2	1.0	0.6
	(*0.3*)	(*0.6*)	(*0.8*)	(*1.0*)	(*0.7*)	(*1.0*)	(*0.7*)	(*0.3*)
CHARMM22	<0.1	1.3	1.1	0.2	<0.1	1.2	1.3	−0.2
	(*0.2*)	(*0.9*)	(*1.2*)	(*0.5*)	(*0.0*)	(*0.0*)	(*1.0*)	(*1.1*)
CHARMM27	0.2	2.9	2.9	−0.2	<0.1	1.1	0.9	0.3
	(*0.0*)	(*0.7*)	(*0.6*)	(*0.1*)	(*0.0*)	(*0.1*)	(*0.6*)	(*0.6*)

[a] Adapted from Ref. [40].

3.2.1. AMBER

The deviations relative to ab initio (Table 2) are $\Delta^\circ = 13.0^\circ(\epsilon)$, $\Delta^\circ = 8.6^\circ(\zeta)$, and $\Delta^\circ = 10.0^\circ$ (pseudorotation angle) for the B_I substate; and $\Delta^\circ = 0.9^\circ$ (ϵ), $\Delta^\circ = 5.9^\circ$ (ζ), and $\Delta^\circ = 33.0^\circ$ (pseudorotation angle) for the B_{II} substate. The major deviation corresponds to the pseudorotation angle in the B_{II} conformation. There are no significant differences in the values of ϵ and ζ calculated with AMBER96 and AMBER98 (see Section 2). However, AMBER98 tends to give pseudoration angles (142.3° for B_I and 126.1° and B_{II}) larger than AMBER96 (137.3 and 120.9°). AMBER reproduces the energy difference $\Delta E = 1.5$ kcal/mol between B_I and B_{II}, in favor of the B_I form, in agreement with the ab initio result ($\Delta\Delta E = 0.0$ kcal/mol). The B_{II} conformation is populated in MD simulations with AMBER [18, Table 10].

3.2.2. CHARMM

When energy minimizing compound C with

Table 11
Descriptors related to the structure and energetics of the purine nucleosides with χ in A and B DNA-like conformations. $\Delta E_{A\text{-}n}$ (kcal/mol) is the energy of the A DNA-like conformation minus the energy of the north energy minimum. $\Delta E_{B\text{-}s}$ (kcal/mol) is the energy of the B DNA-like conformation minus the energy of the north energy minimum. $\Delta E_{B\text{-}A}$ (kcal/mol) is the energy of the B DNA-like conformation minus the energy of the A DNA-like conformation. $\Delta E_{n\text{-}s}$ (kcal/mol) is the north minimum minus the energy of the south minimum. Numbers in parenthesis show the difference in energy ($\Delta\Delta E$) relative to ab initio

Method	Adenine				Guanine			
	$\Delta E_{A\text{-}n}$	$\Delta E_{B\text{-}s}$	$\Delta E_{B\text{-}A}$	$\Delta E_{n\text{-}s}$	$\Delta E_{A\text{-}n}$	$\Delta E_{B\text{-}s}$	$\Delta E_{B\text{-}A}$	$\Delta E_{n\text{-}s}$
Ab initio[a]	**<0.1**	**0.8**	**0.4**	**0.4**	**<0.1**	**0.6**	**−0.1**	**0.7**
AMBER96	0.1	1.9	1.1	0.7	0.1	1.4	0.6	0.7
	(*0.1*)	(*1.1*)	(*0.7*)	(*0.3*)	(*0.1*)	(*0.8*)	(*0.7*)	(*0.0*)
AMBER98	0.4	1.3	0.3	0.7	0.5	1.0	−0.2	0.7
	(*0.4*)	(*0.5*)	(*0.1*)	(*0.3*)	(*0.5*)	(*0.4*)	(*0.1*)	(*0.0*)
CHARMM22	0.3	2.5	5.1	−2.9	0.2	0.8	4.4	−3.8
	(*0.3*)	(*1.7*)	(*4.7*)	(*3.3*)	(*0.2*)	(*0.2*)	(*4.5*)	(*4.5*)
CHARMM27	<0.1	0.8	0.2	0.6	<0.1	0.6	−0.4	1.0
	(*0.0*)	(*0.0*)	(*0.2*)	(*0.2*)	(*0.0*)	(*0.0*)	(*0.3*)	(*0.3*)

[a] Adapted from Ref. [40].

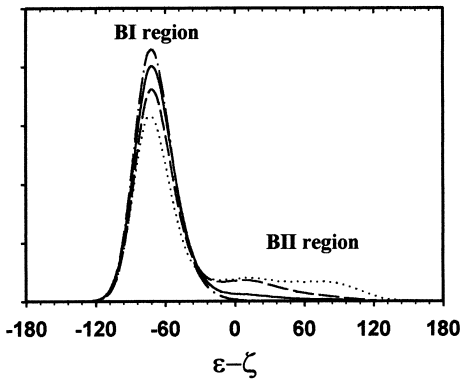

Fig. 3. Distribution curves for the difference $\epsilon - \zeta$ (3454400 values for adenine and thymine and 1604800 values for guanine and cytosine) calculated over several MD trajectories with CHARMM22 for adenine nucleotides (solid line), guanine nucleotides (line-dot-line), thymine nucleotides (dashed line) and cytosine nucleotides (dotted line).

CHARMM22, starting from the B_I conformation, ϵ systematically switches to the *gauche*$^-$ range, while ζ remains in *gauche*$^+$. This outcome was systematically obtained with various energy minimization protocols, even when ϵ was constrained to remain in the *trans* range during the initial stages of the energy minimization. The resulting conformation is neither B_I nor B_{II}. Therefore, the B_I substate is not an energy minimum with CHARMM22 at the model compound level, indicating that the intrinsic energetics associated with this conformation is not properly represented in CHARMM22. Nevertheless, MD simulations of DNA duplexes using CHARMM22 [18] maintains the ϵ and ζ dihedrals in the B_I conformation most of the time (see Ref. [18] and Fig. 3). Fig. 3 shows the distribution for the difference $\epsilon - \zeta$ calculated over several CHARMM22 MD trajectories, and indicates that the backbone essentially remained in the B_I substate, while occasionally sampling the B_{II} substate. Presumably, the B_I conformation remained stable in the CHARMM22 simulations because additional interactions exist in full DNA, that are not present at the model compound level. Also, the $\epsilon =$ *gauche*$^-$/$\zeta =$ *gauche*$^+$ conformation may not be accommodated in duplex DNA. This illustrates how difficult it can be to detect possible problems in the underlying contributions to the force field energetics, only based on simulations results obtained from the full polymer. CHARMM22 yields reasonable condensed phase distributions for the ϵ and ζ torsions [18], although the underlying intrinsic energetics does not appear to be well balanced at the model compound level.

CHARMM27 improves over CHARMM22, by representing the B_I energy minimum in compound C (Table 2). Using this compound, the deviations of CHARMM27 as compared to the ab initio data are $\Delta^\circ = 6.1^\circ$ (ϵ), $\Delta^\circ = 15.2^\circ$ (ζ), and $\Delta^\circ = 1.7^\circ$ (pseudorotation angle) for the B_I substate; and $\Delta^\circ = 6.4^\circ$ (ϵ), $\Delta^\circ = 22.3^\circ$ (ζ), and $\Delta^\circ = 15.9^\circ$ (pseudorotation angle) for the B_{II} substate. These structural deviations are larger for the B_{II} conformation than the B_I. The better representation of the B_I geometries as compared to their B_{II} alternative is consistent with the B_I substate being largely more populated than the B_{II} substate. The CHARMM27 energy difference between B_I and B_{II} ($\Delta E = 1.1$ kcal/mol) deviates slightly more from the ab initio ($\Delta\Delta E = 0.4$ kcal/mol) as compared to AMBER (see above).

3.3. Torsion angle P–O5′–C5′–C4′ (β)

The ab initio torsional energy profiles for β were calculated on model compound B (Fig. 1) in the 90°–330° range (Fig. 4). The ab initio energy profiles (Fig. 4) display only one energy minimum for each furanose pucker, at 238.2° (C2′*endo*) and 240.6° (C3′*endo*).

3.3.1. AMBER

The AMBER β torsional energy profiles (Fig. 4, Table 3) are characterized by a single energy minimum in the *trans* region, with deviations of $\Delta^\circ =$ 51.9° (C2′*endo*) and $\Delta^\circ = 56.8^\circ$ (C3′*endo*) relative to the ab initio data (Table 3). Thus, AMBER positions the β minimum more than 50° lower than in the ab initio calculations. The AMBER energies for β values corresponding to the ab initio minima are 1.7 kcal/mol (C2′*endo*) and 1.9 kcal/mol (C3′*endo*) above the energy minimum in the same energy surface. MD simulations of duplex DNA using the AMBER force filed yield β distributions in the *trans* range [18]), in agreement with experiment [44]. Analysis at the model compound level, however, suggests that the associated intrinsic torsional energetics may be improved. It has already been pointed out [37] that the stabilization of β in the *trans* range

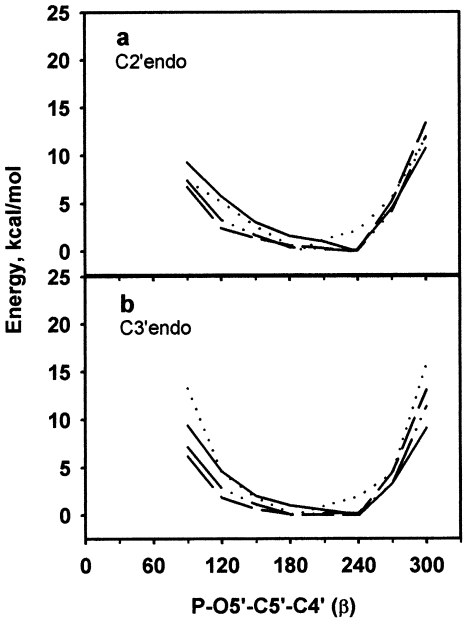

Fig. 4. Potential energy profiles of β in model compound B obtained ab initio (solid line), with AMBER (dotted line), CHARMM22 (dashed line) and CHARMM27 (line-dot-dot-line) algorithms, for (a) C2′*endo* and (b) C3′*endo* conformations of the furanose ring. Each torsional profile was offset relative to its global energy minimum.

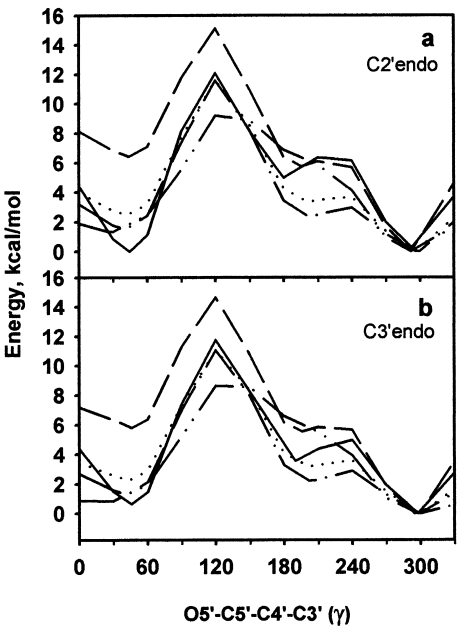

Fig. 5. Potential energy profiles of γ in model compound B obtained ab initio (solid line), with AMBER (dotted line), CHARMM22 (dashed line) and CHARMM27 (line-dot-dot-line), for (a) C2′*endo* and (b) C3′*endo* conformations of the furanose ring. Each torsional profile was offset relative to its global energy minimum.

within the A and B-DNA architectures is expected to result from additional interactions, which are not present at the model compound level. In this context, β should not be forced in the *trans* range by its associated intrinsic torsional energetics.

3.3.2. CHARMM

Fig. 4 shows the CHARMM22 and CHARMM27 β potential energy profiles, with a single energy minimum in every profile. With a C2′*endo* pucker, the β minimum deviates by $\Delta^\circ = 5.7^\circ$ (CHARMM22) and $\Delta^\circ = 3.2^\circ$ (CHARMM27) relative to the ab initio. When the furanose pucker is C3′*endo*, the differences are $\Delta^\circ = 8.4^\circ$ (CHARMM22) and $\Delta^\circ = 52.4^\circ$ (CHARMM27). Thus, the only significant discrepancy with the ab initio profiles is the C3′*endo* CHARMM27 minimum being more than 50° lower than its ab initio counterpart. However, the CHARMM27 energy well in the 180–240° range is broad and shallow, with the higher energy relative to the minimum (at 188.2°) being ~0.6 kcal/mol. The β profiles obtained with CHARMM22 present are in better agreement with ab initio calculations than CHARMM27.

3.4. Torsion angle O5′–C5′–C4′–C3′ (γ)

Fig. 5 shows the γ ab initio potential energy profiles, using model compound B (Fig. 1). The three energy minima in these energy profiles are at 44.7° (*gauche*$^+$), 182.6° (*trans*), and 292.6° (*gauche*$^-$) for C2′*endo*; and 45.9° (*gauche*$^+$), 189.9° (*trans*), and 297.4° (*gauche*$^-$) for C3′*endo* (Table 4). The minimum of lowest energy is in the *gauche*$^+$ region for C2′*endo*, and in the *gauche*$^-$ region for C3′*endo*. However, the energy differences between these *gauche*$^+$ and *gauche*$^-$ minima are less than 1.0 kcal/mol with both puckers (Table 4). The *trans* conformation is significantly higher in energy relative to *gauche*$^+$ ($\Delta E = 5.0$ kcal/mol) C2′*endo* and relative to *gauche*$^-$ ($\Delta E = 3.6$ kcal/mol) for C3′*endo*.

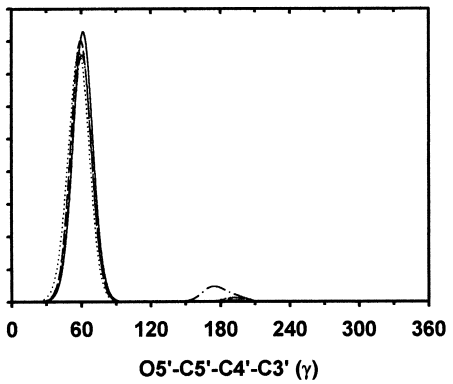

Fig. 6. Distribution curves for γ (3454400 values for adenine and thymine and 1604800 values for guanine and cytosine) calculated over several MD trajectories with CHARMM22 for adenine nucleotides (solid line), guanine nucleotides (line-dot-line), thymine nucleotides (dashed line) and cytosine nucleotides (dotted line).

3.4.1. AMBER

The AMBER γ torsional energy profiles (Fig. 5) also display three energy minima at *gauche*$^+$ ($\Delta^\circ = 3.6^\circ$), *trans* ($\Delta^\circ = 18.9^\circ$), and *gauche*$^-$ ($\Delta^\circ = 3.3^\circ$) for the C2′*endo* conformation; and at *gauche*$^+$ ($\Delta^\circ = 3.8^\circ$), *trans* ($\Delta^\circ = 12.2^\circ$), and *gauche*$^-$ ($\Delta^\circ = 1.6^\circ$) for the C3′*endo* conformation (Table 4). Thus, there is concordance between AMBER and the ab initio regarding the location of the minima, especially in *gauche*$^+$ and *gauche*$^-$ regions. The main difference between the AMBER and ab initio results resides in the relative energies among their respective minima. In contrast to the ab initio results, the minimum of lowest energy in AMBER is in the *gauche*$^-$ region, for both furanose conformations (Table 4). The difference in energy between the *gauche*$^-$ and the *gauche*$^+$ minima are $\Delta E = 1.5$ kcal/mol (C2′*endo*) and $\Delta E = 1.3$ kcal/mol (C3′*endo*). The *trans* minimum is 2.4 kcal/mol (C2′*endo*) and 2.2 kcal/mol (C3′*endo*) above the *gauche*$^-$ minimum. It is thus closer in energy to the *gauche*$^+$ and *gauche*$^-$ minima than in the ab initio calculations (Table 4). Despite these deviations, the AMBER γ energy profiles closely parallel their ab initio counterpart. This is remarkable, given that the γ ab initio energy profiles were not part of the training set of data used to parametrize AMBER.

3.4.2. CHARMM

Fig. 5 shows the CHARMM22 and CHARMM27 profiles for the γ torsion. The CHARMM22 energy profile has three energy minima at *gauche*$^+$ ($\Delta^\circ = 0.6^\circ$), *trans* ($\Delta^\circ = 13.0^\circ$), and *gauche*$^-$ ($\Delta^\circ = 0.3^\circ$) for the C2′*endo* pucker; and at *gauche*$^+$ ($\Delta^\circ = 0.3^\circ$), *trans* ($\Delta^\circ = 6.0^\circ$), and *gauche*$^-$ ($\Delta^\circ = 2.8^\circ$) for the C3′*endo* pucker (Table 4). Thus, CHARMM22 reproduces the number and location of the ab initio minima. In contrast, CHARMM27 γ energy profiles display only the two energy minima of lowest energy, at *gauche*$^+$ ($\Delta^\circ = 19.0^\circ$) and *gauche*$^-$ ($\Delta^\circ = 0.5^\circ$) for the C2′*endo* pucker; and *gauche*$^+$ ($\Delta^\circ = 24.9^\circ$), and *gauche*$^-$ ($\Delta^\circ = 0.3^\circ$) for the C3′*endo* pucker (Table 4). CHARMM27 does not represent the *trans* minimum, found in Z-DNA duplexes [44] and nucleosides and nucleotides [49]. As in AMBER, the *gauche*$^-$ minimum is the lowest in energy for both furanose puckers in both versions of CHARMM (Table 3). In CHARMM22, the energy level of *gauche*$^+$ minimum grossly departs from the ab initio, and is the highest in energy (6.4 kcal/mol above *gauche*$^-$ and 0.6 kcal/mol above *trans* for C2′*endo*; and 5.8 kcal/mol above *gauche*$^-$ and 0.3 kcal/mol above *trans* for C3′*endo*). Fig. 6 shows the distribution curves for γ calculated over several CHARMM22 MD trajectories. Despite this poor representation of the γ torsional energetics, CHARMM22 simulations of DNA duplexes yielded γ distributions that remained in the *gauche*$^+$ experimental range. Therefore, the poor representation of the γ torsional energetics could not have been detected solely from the MD results. This is one more illustration of the importance of the insights gained by using model compounds when assessing the quality of a force field. CHARMM27 dramatically improves over CHARMM22 in its representation of the relative energies of the *gauche*$^+$ and *gauche*$^-$ minima. In CHARMM27 the *gauche*$^+$ minimum is only 1.3 kcal/mol (C2′*endo*) and 0.8 kcal/mol (C3′*endo*) above the *gauche*$^-$ minimum.

3.5. Torsion angle O4′–C1′–N1–C2 in pyrimidines or O4′–C1′–N9–C4 in purines (χ)

Figs. 7 and 8 show the potential energy profiles of the torsion angle χ in compound E (Fig. 1) for adenine, guanine, cytosine and thymine bases, respectively. In all cases the ab initio profiles mirrored the known bimodal distribution of χ [37]. The minimum in the 0°–100° range corresponds to the *syn* orientation

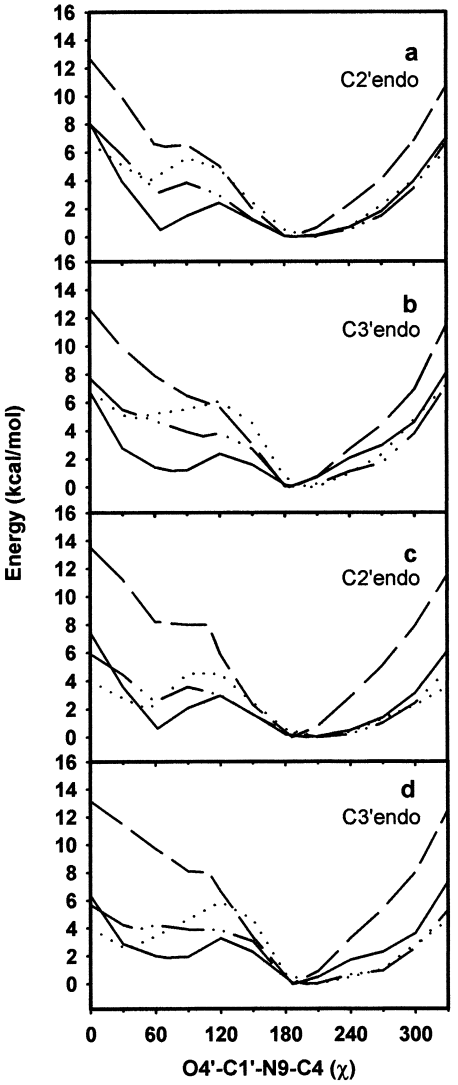

Fig. 7. Potential energy profiles of χ in model compound E with purine bases obtained ab initio (solid line), with AMBER (dotted line), CHARMM22 (dashed line) and CHARMM27 (line-dot-dot-line). Panels (a) and (b) correspond to adenine with sugar in C2′*endo* and C3′*endo* conformations of the furanose ring, respectively. Panels (c) and (d) correspond to guanine with sugar in C2′*endo* and C3′*endo* conformations of the furanose ring, respectively. Each torsional profile was offset relative to its global energy minimum.

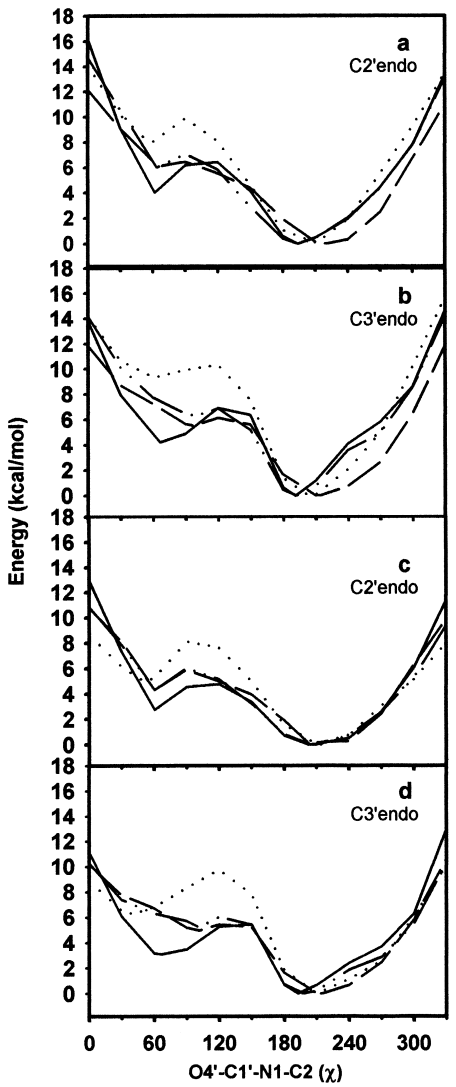

Fig. 8. Potential energy profiles of χ in model compound E with pyrimidine bases obtained ab initio (solid line), with AMBER (dotted line), CHARMM22 (dashed line) and CHARMM27 (line-dot-dot-line). Panels (a) and (b) correspond to cytosine with sugar in C2′*endo* and C3′*endo* conformations of the furanose ring, respectively. Panels (c) and (d) correspond to thymine with sugar in C2′*endo* and C3′*endo* conformations of the furanose ring, respectively. Each torsional profile was offset relative to its global energy minimum.

of the base, whereas the minimum in the 170°–280° range corresponds to the *anti* orientation [39]. Tables 5–8 give the locations of the ab initio minima and the energy differences between them. When the furanose ring is C2′*endo*, the *syn* minimum is located at 65.5, 62.1, 61.6, and 60.7° for adenine, guanine, cytosine and thymine, respectively (same order throughout); and the *anti* minimum at 189.9, 198.5, 194.1, and

202.6°. The C2′*endo*/*anti* minimum is always lower in energy than the C2′*endo*/*syn* minimum ($\Delta = 0.5$, 0.6, 4.0, and 2.8 kcal/mol, respectively). The C3′*endo*/*syn* minimum is located at 75.4, 70.4, 66.6, and 66.6°, and the C3′*endo*/*anti* minimum at 184.1, 187.7, 191.5, and 192.8°. Similarly, the C3′*endo*/*anti* minimum is always lower in energy than the C3′*endo*/*syn* minimum ($\Delta E = 1.2$, 1.9, 4.2 and 3.1 kcal/mol).

3.5.1. AMBER

The recent reparametrization of the AMBER force field for nucleic acids involved a modification of the torsional parameters associated with the glycosyl linkage [23]. However, the χ energy profiles obtained with compound E in the present work, with both the AMBER96 and AMBER98 parameter sets are very similar. The B-DNA range was systematically marginally more stable relative to the A-DNA range with AMBER98, as compared to AMBER96. Besides, the *anti* minimum in the χ torsional profile is consistently displaced towards larger values with AMBER98 than with AMBER96 (average of 6.2°). Thus, AMBER98 moves the *anti* minimum away from the value obtained its ab initio counterpart. These differences, however, were not compelling enough to warrant a presentation of the energy profiles obtained with both AMBER parameter sets. In the following, only the profiles obtained with AMBER98 are shown and discussed.

Figs. 7 and 8 show the AMBER energy profiles for the χ torsion, with adenine, guanine, cytosine and thymine, respectively. The AMBER energy profiles have an overall shape similar to their ab initio counterpart, with two energy minima. The first energy minimum is in *syn* (C2′*endo*: $\Delta° = 10.7$, 11.4, 5.3, and 9.9°; C3′*endo*: $\Delta° = 33.7$, 35.0, 4.6, and 26.0° for adenine, guanine, cytosine and thymine, respectively). The second in *anti* (C2′*endo*: $\Delta° = 15.3$, 12.5, 16.0, and 10.8°; C3′*endo*: $\Delta° = 19.6$, 19.5, 12.2, and 17.2°). AMBER properly represents the locations the *syn* and *anti* minima when the furanose is in C2′*endo*, but larger deviations relative to the ab initio results are observed for the C3′*endo* pucker. These deviations mainly occur in the *syn* minimum with adenine, guanine, and thymine ($\Delta°$ in the 26.0–35.0° range).

The energy differences between *syn* and *anti* minima with the C2′*endo* pucker are $\Delta E = 3.9$ kcal/mol ($\Delta\Delta E = 3.4$ kcal/mol), $\Delta E = 2.1$ kcal/mol ($\Delta\Delta E = 1.5$ kcal/mol), ($\Delta E = 8.0$ kcal/mol), ($\Delta\Delta E = 4.0$ kcal/mol), and ($\Delta E = 5.0$ kcal/mol) ($\Delta\Delta E = 2.2$ kcal/mol) for adenine, guanine, cytosine and thymine, respectively. For the C3′*endo* pucker, the AMBER *syn* minima are also higher in energy as compared to *anti*, with $\Delta E = 4.9$ kcal/mol ($\Delta\Delta E = 3.7$ kcal/mol) for adenine, ($\Delta E = 2.6$ kcal/mol) ($\Delta\Delta E = 1.5$ kcal/mol) for guanine, $\Delta E = 9.4$ kcal/mol ($\Delta\Delta E = 5.2$ kcal/mol) for cystosine, and $\Delta E = 6.3$ kcal/mol ($\Delta\Delta E = 3.4$ kcal/mol) for thymine. Therefore, the AMBER profiles agree with the ab initio in that the *anti* minima are lower in energy than the associated *syn* minima. However, the energy differences between these minima are always significantly larger than in the ab initio. Previous analysis of the ab initio structures [37] suggested that, with a C3′*endo* pucker, there may be an energetically favorable interaction between the C2′–H2′ group and the N3 atom of purines or the O2 atom of pyrimidines. It was suggested that this interaction may present some hydrogen bond character, although additional work is needed to clarify this point (Foloppe and MacKerell, in preparation). If this interaction is favorable, it will be represented as such in a quantum mechanical framework, however, it is likely that it is not the case in the current biomolecular force fields. This point was already commented upon when discussing the CHARMM27 force field [22], where the presumably favorable interaction between the C3′–H3′ group and N3/O2 is not represented. The present data suggest that the same limitation exist in the current AMBER force field.

3.5.2. CHARMM

The comparison between CHARMM22, CHARMM27 and ab initio χ torsional energy profiles (Figs. 7–8) will be carried out separately for the purine and pyrimidines bases.

For pyrimidines, CHARMM22 does not differ significantly from CHARMM27 concerning the location of the minima, and the energy differences between them, for both furanose puckers (Tables 7–8 and Fig. 8). All these energy profiles reproduce the two energy minima, at *syn* (C2′*endo*: $\Delta° = 3.5/2.3°$ and 0.1/0.1°; C3′*endo*: $\Delta° = 32.3/33.2°$ and 34.4/34.8° for CHARMM22/CHARMM27 and cytosine and thymine, respectively) and *anti* (C2′*endo*:

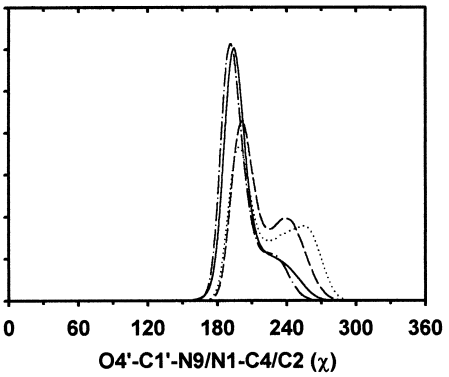

Fig. 9. Distribution curves for χ (3454400 values for adenine and thymine and 1604800 values for guanine and cytosine) calculated over several MD trajectories with CHARMM22 for adenine nucleotides (solid line), guanine nucleotides (line-dot-line), thymine nucleotides (dashed line) and cytosine nucleotides (dotted line).

Δ° 27.5/0.2° and 1.6/1.6°; C3′*endo*: $\Delta^\circ = 23.0/0.6$ and 22.4/4.8°). The largest deviations between CHARMM22 and CHARMM27 occur in the *anti* region. For both furanose puckers, CHARMM27 better reproduces the ab initio location of the *anti* minimum than CHARMM22, by shifting this minimum to smaller values (Tables 7–8). Fig. 9 shows the distribution curves for χ calculated over several CHARMM22 MD trajectories. These simulations of DNA duplexes yielded χ distributions in both the A- and B-DNA conformations (the average values of χ for crystal structures of A-DNA and B-DNA are 199 and 249°, respectively [44]). However, the values of χ corresponding to the A-DNA form are more populated. The differences in CHARMM27 energy between *syn* and *anti* minima for C2′*endo* are $\Delta E =$ 6.0 kcal/mol ($\Delta\Delta E = 2.0$ kcal/mol) for cytosine, and $\Delta E = 4.3$ kcal/mol ($\Delta\Delta E = 1.5$ kcal/mol) for thymine; and for C3′*endo* are $\Delta E = 6.3$ kcal/mol ($\Delta\Delta E = 2.1$ kcal/mol) for cytosine, and $\Delta E =$ 5.2 kcal/mol ($\Delta\Delta E = 2.1$ kcal/mol) for thymine. The *anti* minimum is lower in energy than the *syn* minimum, in agreement with ab initio results. Therefore, the properties of the CHARMM27 minima with a C2′*endo* pucker are closer to the ab initio than with a C3′*endo* pucker, in both location and energy difference. In both CHARMM22 and CHARMM27, the C3′*endo* potential energy profiles do not present a deep enough *syn* minimum, as compared to the ab initio results. These findings may be attributed to the formation of a weak hydrogen bond interaction between the C3′–H3′ group and the O2 atom of the base (see discussion above with AMBER results), which would be poorly represented in the force fields.

For purine residues, CHARMM22 (Figs. 7) does not reproduce the two-energy minima present in the ab initio profiles. CHARMM22 does not present any *syn* minimum for adenine and guanine in C3′*endo*, nor for guanine in C2′*endo*; the *syn* minimum is practically inexistent for adenine in C2′*endo*. CHARMM22 departs from the locations of the ab initio *anti* minima by $\Delta^\circ = 3.5^\circ$ for C2′*endo*/adenine, $\Delta^\circ = 12.2^\circ$ for C2′*endo*/guanine, $\Delta^\circ = 2.3^\circ$ for C3′*endo*/adenine and $\Delta^\circ = 1.4^\circ$ for C3′*endo*/guanine (Tables 5–6). Also, the CHARMM22 energy profiles strongly deviate from the ab initio in the χ range corresponding to the B-DNA region, for the purines. In this region, CHARMM22 strongly destabilizes the conformations corresponding to B-DNA, relative to A-DNA. This is likely to be one of the major reasons why CHARMM22 systematically stabilized the A form of DNA as compared to the B form, in MD of duplex DNA in aqueous solution [4]. Fig. 9 shows the χ distributions for each base separately, calculated over several CHARMM22 MD trajectories [15]. These simulations yielded χ distributions essentially in the A-DNA conformation. In the present context, it is useful to note that the χ distributions associated with the pyrimidines populate the B-DNA range more than the purines (Fig. 9). These differences in simulated χ distributions mirror the differences in the χ energetics in compound E. This illustrates how the energetics derived at the model compound level may help to understand, and correct, the simulated properties of the full polymer.

The general shape of the torsional energy profiles with purines is significantly improved with CHARMM27 versus CHARMM22. Indeed, with the C2′*endo* pucker, CHARMM27 allows for the existence of the *syn* minima. But even more importantly, the energy wells corresponding to the *anti* orientation are significantly shallower in CHARMM27 than in CHARMM22. The CHARMM27 deviations relative to ab initio, for the C2′*endo*/*syn* minima are $\Delta^\circ = 2.4^\circ$ for adenine and $\Delta^\circ = 1.6^\circ$ for guanine; and of the C2′*endo*/*anti* minima are $\Delta^\circ = 4.4^\circ$ for adenine and $\Delta^\circ = 14.8^\circ$ for guanine (Tables 5–6). The difference in energy between the *syn* and *anti* minima with the

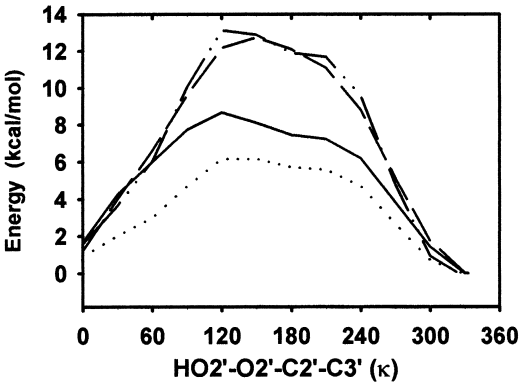

Fig. 10. Potential energy profiles for the rotation around the C2′–O2′ bond for the typical C3′*endo* conformation of RNA, obtained ab initio (solid line), with AMBER (dotted line), CHARMM22 (dashed line) and CHARMM27 (line-dot-dot-line). Each torsional profile was offset relative to its global energy minimum.

C2′*endo* pucker is $\Delta E = 3.1$ kcal/mol ($\Delta\Delta E = 2.6$ kcal/mol) for adenine, and $\Delta E = 2.5$ kcal/mol ($\Delta\Delta E = 1.9$ kcal/mol) for guanine. Thus, the *anti* minimum is always lower in energy than the *syn* minimum in agreement with ab initio results, but the energy differences are always larger than in the ab initio. The CHARMM27 deviations relative to the ab initio for the C3′*endo*/*syn* minima are $\Delta^\circ = 29.5^\circ$ for adenine and $\Delta^\circ = 27.0^\circ$ for guanine; and of the C3′*endo*/*anti* minima are $\Delta^\circ = 7.0^\circ$ for adenine and $\Delta^\circ = 10.8^\circ$ for guanine (Tables 5–6). The energy difference between the *syn* and *anti* minima is $\Delta E = 3.6$ kcal/mol ($\Delta\Delta E = 2.4$ kcal/mol) for adenine, and $\Delta E = 3.9$ kcal/mol ($\Delta\Delta E = 2.0$ kcal/mol) for guanine. The location and relative energies of the *syn* minima with the purines are poorly reproduced in CHARMM27 as compared to ab initio. The same feature has been observed throughout the present work, with all force fields and all bases. This reinforces the notion that it is due to a systematic shortcoming in the analyzed force fields. This shortcoming may be related to the poor treatment, if any, in current biomolecular force fields, of possible hydrogen bonds involving a C–H donor.

3.6. Torsion angle C3′–C2′–O2′–HO2′

The presence of the 2′-OH group in RNA plays a major role in the structure [39,50,51], stability [52], and hydration [53,54] of RNA relative to DNA. The orientation of the O2′–HO2′ bond in RNA is therefore an important structural feature. Experimental methods alone, however, are generally unable to unambigously locate the positions of the HO2′ hydrogen in RNA, and MD simulations can complement experiments by providing direct insights regarding the orientation of the O2′–HO2′ bond [12]. It is therefore important to ensure that the torsional potential associated with the C3′–C2′–O2′–HO2′ torsion is properly treated by the RNA force fields. Fig. 10 shows the ab initio potential energy profile for the C3′–C2′–O2′–HO2′ torsion in model compound D (Fig. 1). The only energy minimum is at *gauche*$^-$ (Table 9) and the energy maximum is approximately at 120° (7.8 kcal/mol higher in energy relative the minimum).

3.6.1. AMBER

The AMBER C3′–C2′–O2′–HO2′ torsional energy profile displays a single energy minimum in the *gauche*$^-$ range ($\Delta^\circ = 3.2^\circ$, see Table 9) and a maximum in the 120–150° range (~6.1 kcal/mol), in agreement with ab initio values. This agreement provides support to the previously published analysis of the 2′-OH orientation in RNA, based on MD simulation using the AMBER force field [12].

3.6.2. CHARMM

CHARMM22 and CHARMM27 produce comparable C3′–C2′–O2′–HO2′ torsional energy profiles (Fig. 10). Both energy profiles display a single energy minimum in the *gauche*$^-$ range ($\Delta^\circ = 2.4^\circ$ for CHARMM22; $\Delta^\circ = 6.0^\circ$ for CHARMM27), and an energy maximum at approximately 120°. The energy maximum is higher by ~4.5 kcal/mol in both CHARMM versions as compared to the ab initio values. This suggests that the current versions of the CHARMM force field may not provide an optimal representation of the dynamics associated with the 2′-OH group orientation. The height of the energy maximum is also the major difference between the CHARMM and AMBER energy profiles, this height being significantly close to the ab initio in AMBER than in CHARMM. This feature should be improved in future versions of the CHARMM nucleic acids force field.

3.7. Energy differences between A-DNA and B-DNA like conformations

Tables 10 and 11 show the difference in energy at the nucleoside level between the north, south, A-DNA, and B-DNA conformations (see Section 2). In the ab initio, restricting χ to a B-DNA like conformation (see Section 2) increases the energy relative to the south energy minimum ($\Delta E_{\text{B-s}}$ in Tables 9–10) by 0.8, 0.6, 2.2, and 1.2 kcal/mol for adenine, guanine, cytosine, and thymine, respectively. In contrast, the energy difference between the A-DNA like conformation and the north energy minimum ($\Delta E_{\text{A-n}}$) is significantly smaller: <0.1, <0.1, <0.1, and 0.2 kcal/mol, respectively. The energy difference between B- and A-DNA like conformations ($\Delta E_{\text{B-A}}$) for adenine, guanine, and thymine are less than 0.5 kcal/mol (0.4, -0.1, and 0.3 kcal/mol), suggesting that neither the A- or B-DNA form is favored. For deoxy-cytidine, however, $\Delta E_{\text{B-A}}$ is 2.3 kcal/mol, suggesting that the A-DNA form is intrinsically favored [40]. Also, deoxy-cytidine has a lower energy in the north minimum as compared to the south (-0.3 kcal/mol), in contrast with what was obtained with the other bases (0.4, 0.7, 0.9 kcal/mol for adenine, guanine, and thymine, respectively).

3.7.1. AMBER

AMBER96 properly reproduces the energy difference between A-DNA and north conformation, $\Delta E_{\text{A-n}}$, relative to ab initio. AMBER98 uniformly increases these energy differences, leading to a $\Delta\Delta E_{\text{A-n}}$ average of 0.5 kcal/mol. The values of $\Delta E_{\text{B-s}}$ obtained with AMBER96 are systematically larger than in ab initio: $\Delta\Delta E_{\text{B-s}}$ of 1.1, 0.8, 1.1, and 1.5 kcal/mol for adenine, guanine, cytosine, and thymine, respectively. AMBER98 diminishes this difference $\Delta\Delta E_{\text{B-s}}$ to 0.5, 0.4, 0.6, and 1.0 kcal/mol, respectively. Thus, AMBER98 reduces 0.5 kcal/mol in average of the energy difference between B- and south minima. Both $\Delta E_{\text{A-n}}$ and $\Delta E_{\text{B-s}}$ reflects the main scope of AMBER98 of obtaining higher χ values for the south conformation (Tables 10 and 11). The energy differences between A- and B-DNA conformations, $\Delta E_{\text{B-A}}$, are 1.1, 0.6, 2.3 and 1.8 kcal/mol for adenine, guanine, cytosine, and thymine, respectively. Thus, AMBER96 favoured the A form of DNA for all deoxy-nucleosides. The average increment in energy relative to ab initio, $\Delta\Delta E_{\text{B-A}}$, is 0.7 kcal/mol. AMBER98 clearly decreased this value to an average of 0.4 kcal/mol. It is important to note that in the case of deoxy purines the energy difference relative to ab initio, $\Delta\Delta E_{\text{B-A}}$, is distinctly reduced to 0.1 kcal/mol. With the exception of deoxy-guanosine, both AMBER96 and AMBER98 have tendency to favor the A-DNA conformation at the nucleoside level. Nevertheless AMBER98 better reproduces the ab initio results, as has already been reported in the original contribution [23]. The values reported, with both AMBER96 and AMBER98, for the energy difference between north and south conformations, $\Delta E_{\text{n-s}}$, leads to a lower energy in the south conformation than the north conformation for all nucleosides. This is in contrast to ab initio calculations where the south conformation is lower in energy for adenine, guanine and thymine, but not for cytosine (Tables 10 and 11). The average energy difference relative to ab initio, $\Delta\Delta E_{\text{n-s}}$, is 0.4 kcal/mol in either AMBER96 or AMBER98. Thus, with the exception of $\Delta E_{\text{A-n}}$, AMBER98 is in better agreement with the intrinsic energetics of the nucleosides than AMBER96, considering the ab initio calculations as the reference.

3.7.2. CHARMM

Some properties of the nucleosides in the context of the CHARMM27 force field were already reported [22], however their relative energies in the A- and B-DNA like conformations with CHARMM27 were obtained as part of the present work (Tables 10 and 11). The $\Delta E_{\text{B-A}}$ values in CHARMM27 follow the trends observed in their ab initio counterpart. In particular, the destabilization of the B form of DNA by deoxy-cytidine is represented in CHARMM27. This confirms that CHARMM27 provides a reasonable representation of the nucleosides intrinsic conformational energetics, and significantly improves over CHARMM22 in this respect. Indeed, $\Delta E_{\text{B-A}}$ values with the purines in CHARMM22 are on the order of 4.5 kcal/mol higher than in the ab initio data. This provides a convincing explanation as to why CHARMM22 systematically stabilized the A form of DNA over the B form. The destabilization of the B form of DNA in CHARMM22 is much less with the

pyrimidines than with the purines, as already observed with the χ energy profiles in compound E (Figs. 7 and 8). This specifically pinpoints to the parametrization of the glycosyl torsion energetics in purines as a major shortcoming of CHARMM22. Ironically, $\Delta E_{\text{B-A}}$ was the lowest for deoxy-cytidine in CHARMM22, in contrast to what is expected from the ab initio calculations. The comparison of the CHARMM22 and CHARMM27 treatments of the deoxynucleosides intrinsic energetics demonstrates the usefulness of this type of analysis to understand and calibrate the properties of a nucleic acid force field.

4. Conclusions

The latest parametrizations of CHARMM and AMBER represent the state of the art in nucleic acid force field development. The CHARMM27 and AMBER98 empirical torsional energy profiles agree to a large extent with their ab initio counterpart, suggesting that these force field are well balanced in this respect. It would be interesting to assess the performances of other nucleic acids force fields along the same lines. Overall, AMBER98 performs as well as CHARMM27 regarding the fit of the empirical torsional potentials to their ab initio counterpart, although CHARMM27 relied more explicitely and extensively than AMBER98 on these ab initio data to parametrize this energetics. This is in part due to the fit to the ab initio data being occasionally sacrificed when parametrizing CHARMM27, to improve the representation of the full polymer condensed-phase properties. The uncertainties involved in this compromise were highlighted in the Introduction. Also, the results obtained with CHARMM22 illustrate how model compounds may be used to dissect the various contributions to a force field, and pinpoint to specific problems which may or may not be apparent in the simulated properties of the full polymer. For instance, the present results indicate that a major shortcoming of CHARMM22 was the glycosidic torsional energetics associated with the purines. The latest GROMOS parameter set for nucleic acids also has a tendency to destabilize B-DNA [55]; the analysis performed in this work might be of use to localize the interactions responsible for this behavior.

Despite the improvements implemented in CHARMM27 and AMBER98, our comparison shows that the ab initio energy profiles are not reproduced perfectly by any of the analyzed versions of CHARMM or AMBER. Assuming that the level of ab initio calculations used in the parametrization is adequate to describe the energetics of torsional motion, the differences between the force fields and the ab initio calculations may reflect possible limitations inherent in performing energetic analyses of one conformational variable at a time. This may be alleviated with the development of more sophisticated potential functions with an explicit coupling between the different nucleic acid dihedrals. This notion is further stressed by the fact that the deepest ab initio minima do not always correspond to the most populated conformation found in the available structural databases or in molecular dynamics simulations (see for example the profiles for β and γ).

While the overall lowest energy structure of nucleic acid oligomers seems to be reasonably well represented in the latest versions of CHARMM and AMBER, backbone dynamics are also likely to be revised in the future, as the location and height of the energy barriers separating different conformers are not quantitatively reproduced in general.

Fine details regarding conformational transitions and alternative conformations in DNA oligomers, and the associated energetics, still show room for improvement. For example, the CHARMM and AMBER energy profiles for χ require revision, especially if the A-DNA to B-DNA and B-DNA to Z-DNA transitions are to be described in a quantitatively accurate way. The location of the minimum in the *syn* conformation and the barrier heights are poorly described, probably because the possibility of interactions between the base rims and the sugar appears not to be properly reproduced by either force field. While CHARMM27 has corrected the main deficiency in CHARMM22 regarding B-DNA destabilization, the shallow and A-DNA centered *anti* minima are likely to require very fine tuning, together with sugar pucker preferences, in order to produce accurate barriers for the A-DNA to B-DNA transition.

Model compounds have proven to be a valuable tool in the analysis of the performance of force fields, and also in their refinement. It is to be expected that careful choices in the selection of model compounds will help to further improve current force fields to

resolve the discrepancies detailed in this work. It should also be stressed that the validation of the force fields requires the simulation of a collection of different DNA or RNA sequences, as some of the dihedral angles (notably, the B_I to B_{II} transition and χ) show qualitatively distinct behavior for purines and pyrimidines.

Acknowledgements

MC wants to dedicate her contribution in this Special Issue in honor of Professor Fraga on his 70th birthday, because he introduced her to the fascinating world of computer simulation of biomolecules. This work was supported in part by a grant from DGES: PB98-0907 (LP and MC) and the Training and Mobility of Researchers program of the European Community: ERB FMGE CT95 0062 (NF). Some of the simulations were run at the Centre de Computació i Comunicacions de Catalunya and at the Dirección General de Servicios de Cómputo Académico (UNAM-México).

Appendix A

Atomic RESP partial charges derived for model compounds A, B, C and D. Constrained RESP atomic partial charges are shown in italics (see Section 2 for the charge constraints used). All partial charges are in atomic units.

Atom name	A	B	C	D
C1′	*−0.0342*	*0.0120*	*0.1105*	*0.0177*
H1′1	*0.0779*	*0.0548*	*0.0055*	*0.0519*
H1′2	*0.0779*	*0.0548*	*0.0055*	*0.0519*
C2′	−0.0854	−0.0631	0.0867	0.2622
H2′1	0.0757	0.0374	−0.0260	−0.0218
H2′2	0.0757	0.0374	−0.0260	–
O2′	–	–	–	−0.5836
HO2′	–	–	–	0.3158
C3′	0.0713	−0.0767	0.5714	0.3448
H3′1	−0.0400	0.0325	−0.0921	0.0021
H3′2	–	0.0325	–	–
C4′	0.1629	0.2977	0.0527	0.0811
H4′1	0.0239	0.0266	0.0159	0.0464
H4′2	0.0239	–	0.0159	0.0464
O4′	−0.4691	−0.4546	−0.4726	−0.4531
O3′	−0.5521	−0.6221	−0.6303	−0.5687
H3T	–	0.3727	–	–
P	1.6839	1.1115	1.2584	1.2172
O1P	−0.8599	−0.7501	−0.7643	−0.7858
O2P	−0.8599	−0.7501	−0.7643	−0.7858
O5′	−0.6276	−0.4589	−0.4813	−0.3730
H5T	0.2552	–	–	–
C5′	–	−0.1987	–	–
H5′1	–	0.1522	–	–
H5′2	–	0.1522	–	–
C5T	–	–	0.0090	−0.0433
H51	–	–	0.0418	0.0593
H52	–	–	0.0418	0.0593
H53	–	–	0.0418	0.0593

Appendix B

Atomic RESP partial charges derived for model compound E with pyrimidine bases. RESP atomic partial charges frozen to AMBER partial charges values are shown in italics (see Section 2 for the charge constraints used). All partial charges are in atomic units.

Atom name	Cytosine	Thymine
O4′	*−0.3691*	*−0.3691*
C1′	*−0.0116*	*0.0680*
H1′	*0.1963*	*0.1804*
N1	*−0.0339*	*−0.0239*
C2	*0.7959*	*0.5677*
O2	*−0.6548*	*−0.5881*
N3	*−0.7748*	*−0.4340*
H3	–	*0.3420*
C4	*0.8439*	*0.5194*
N4	*−0.9773*	–
H41	*0.4314*	–
H42	*0.4314*	–
O4	–	*−0.5563*
C5	*−0.5222*	*0.0025*
H5	*0.1863*	–
C5M	–	*−0.2269*
H51	–	*0.0770*
H52	–	*0.0770*
H53	–	*0.0770*
C6	*−0.0183*	*−0.2209*
H6	*0.2293*	*0.2607*
C2′	*−0.0854*	*−0.0854*
H2′1	*0.0718*	*0.0718*
H2′2	*0.0718*	*0.0718*

(*continued*)

Atom name	Cytosine	Thymine
C3′	0.2124	−0.1884
H3′1	−0.0446	0.0554
H3′2	−0.0446	0.0554
C4′	−0.3768	0.1526
H4′1	0.2215	0.0571
H4′2	0.2215	0.0571

Appendix C

Atomic RESP partial charges derived for model compound E with purine bases. RESP atomic partial charges frozen to AMBER partial charges values are shown in italics (see Section 2 for the charge constraints used). All partial charges are in atomic units.

Atom name	Adenine	Guanine
O4′	*−0.3691*	*−0.3691*
C1′	*0.0431*	*0.0358*
H1′	*0.1838*	*0.1746*
N9	*−0.0268*	*0.0577*
C4	*0.3800*	*0.1814*
N3	*−0.7417*	*−0.6636*
C2	*0.5716*	*0.7432*
H2	*0.0598*	–
N2	–	*−0.9230*
H21	–	*0.4235*
H22	–	*0.4235*
N1	*−0.7624*	*−0.5053*
H1	–	*0.3520*
C6	*0.6897*	*0.4918*
N6	*−0.9123*	–
H61	0.4167	–
H62	0.4167	–
O6	–	*−0.5699*
C5	0.0725	*0.1991*
N7	*−0.6175*	*−0.5725*
C8	*0.1607*	*0.0736*
H8	*0.1877*	*0.1997*
C2′	*−0.0854*	*−0.0854*
H2′1	*0.0718*	*0.0718*
H2′2	*0.0718*	*0.0718*
C3′	−0.0695	−0.2283
H3′1	0.0603	0.0914
H3′2	0.0603	0.0914
C4′	−0.1410	0.0740
H4′1	0.1396	0.0803
H4′2	0.1396	0.0803

References

[1] B. Hartmann, D. Piazzola, R. Lavery, Nucleic Acids Res. 21 (1993) 561.
[2] S. Sen, L. Nilsson, Biophys. J. 77 (1999) 1801.
[3] S. Sen, L. Nilsson, Biophys. J. 77 (1999) 1782.
[4] L. Yang, B.M. Pettitt, J. Phys. Chem. 100 (1996) 2564.
[5] T.E. Cheatham, P.A. Kollman, Structure 5 (1997) 1297.
[6] M.A. Young, G. Ravishanker, D.L. Beveridge, Biophys. J. 73 (1997) 2313.
[7] T.E. Cheatham, P.A. Kollman, J. Mol. Biol. 259 (1996) 434.
[8] D. Flatters, M. Young, D.L. Beveridge, R. Lavery, J. Biomol. Struct. Dyn. 14 (1997) 757.
[9] M. Feig, B.M. Pettitt, J. Mol. Biol. 286 (1999) 1075.
[10] M. Feig, B.M. Pettitt, Structure 6 (1998) 1351.
[11] A.D.J. MacKerell, J. Phys. Chem. B. 101 (1997) 646.
[12] P. Auffinger, E. Westhof, J. Mol. Biol. 274 (1997) 54.
[13] L. Pardo, N. Pastor, H. Weinstein, Biophys. J. 74 (1998) 2191.
[14] L. Pardo, N. Pastor, H. Weinstein, Biophys. J. 75 (1998) 2411.
[15] N. Pastor, L. Pardo, H. Weinstein, Biophys. J. 73 (1997) 640.
[16] L. Pardo, M. Campillo, D. Bosch, N. Pastor, H. Weinstein, Biophys. J. 78 (2000) 1988.
[17] A.R. Srinivasan, T.E. Cheatham, P. Cieplak, P.A. Kollman, D.A. Case, J. Am. Chem. Soc. 120 (1998) 9401.
[18] M. Feig, B.M. Pettitt, Biophys. J. 75 (1998) 134.
[19] M. Feig, B.M. Pettitt, J. Phys. Chem. B. 101 (1997) 7361.
[20] A.D.J. MacKerell, J. Wiórkiewicz-Kuczera, M. Karplus, J. Am. Chem. Soc. 117 (1995) 11 946.
[21] W.D. Cornell, P. Cieplak, C. Bayly, I.R. Gould, K.M. Merz, D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, J. Am. Chem. Soc. 117 (1995) 5179.
[22] N. Foloppe, A.D.J. MacKerell, J. Comp. Chem. 21 (2000) 86.
[23] T.E. Cheatham, P. Cieplak, P.A. Kollman, J. Biomol. Struct. Dyn. 16 (1999) 845.
[24] D.R. Langley, J. Biomol. Struct. Dyn. 16 (1998) 487.
[25] S. Neidle, Nat. Struct. Biol. 5 (1998) 754.
[26] X. Shui, L. McFail-Isom, G.G. Hu, L.D. Williams, Biochemistry 37 (1998) 8341.
[27] S. Jain, M. Sundaralingam, J. Biol. Chem. 264 (1989) 12780.
[28] Z. Shakked, G. Guerstein-Guzikevich, M. Eisenstein, F. Frolow, D. Rabinovich, Nature 342 (1989) 456.
[29] A. Lipanov, M.L. Kopka, M. Kaczor-Grzeskowiak, J. Quintana, R.E. Dickerson, Biochemistry 32 (1993) 1373.
[30] R.E. Dickerson, D.S. Goodsell, S. Neidle, Proc. Natl. Acad. Sci. USA 91 (1994) 3579.
[31] N.B. Ulyanov, U. Schmitz, A. Kumar, T.L. James, Biophys. J. 68 (1995) 13.
[32] U. Schmitz, T.L. James, Methods Enzymol. 261 (1995) 3.
[33] J.P. Rife, S.C. Stallings, C.C. Correll, A. Dallas, T.A. Steitz, P.B. Moore, Biophys. J. 76 (1999) 65.
[34] W.J. Metzler, C. Wang, D.B. Kitchen, R.M. Levy, A. Pardi, J. Mol. Biol. 214 (1990) 711.
[35] F.H. Allain, G. Varani, J. Mol. Biol. 267 (1997) 338.
[36] H.L. Ng, M.L. Kopka, R.E. Dickerson, Proc. Natl. Acad. Sci. USA 97 (2000) 2035.
[37] N. Foloppe, A.D.J. MacKerell, J. Phys. Chem. B. 103 (1999) 10955.

[38] W.F. van Gunsteren, S.R. Billeter, A.A. Eising, P.H. Hünenberger, P. Krüger, A.E. Mark, W.R.P. Scott, I.G. Tironi, Groningen Molecular Simulation (GROMOS) Library Manual, Biomos, Zurich, 1996.
[39] W. Saenger, Principles of Nucleic Acid Structure, Springer, New York, 1984.
[40] N. Foloppe, A.D.J. MacKerell, Biophys. J. 76 (1999) 3206.
[41] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery, T.A. Keith, G.A. Petersson, K. Raghavachari, A. Al-Laham, R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Menucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. González, M. Challacombe, P.M.W. Gill, B.G. Johnson, W. Chen, W. Wong, J.L. Andres, M. Head-Gordon, E.S. Replogle, J.A. Pople, GAUSSIAN 98, Gaussian Inc., Pittsburgh, PA, 1998.
[42] D.A. Case, D.A. Pearlman, J.W. Caldwell, T.E. Cheatham, W.S. Ross, C. Simmerling, T.A. Darden, K.M. Merz, R.B. Stanton, A. Cheng, J.J. Vincent, M. Crowley, D.M. Ferguson, R.J. Radmer, G.L. Seibel, U.C. Singh, P. Weiner, P.A. Kollman, AMBER 5.0, University of California, San Francisco, 1997.
[43] B.B. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, J. Comp. Chem. 4 (1983) 187.
[44] B. Schneider, S. Neidle, H.M. Berman, Biopolymers 42 (1997) 113.
[45] Y. Gu, T. Kar, S. Scheiner, J. Am. Chem. Soc. 121 (1999) 9411.
[46] R. Taylor, O. Kennard, J. Am. Chem. Soc. 104 (1982) 5063.
[47] G.R. Desiraju, Acc. Chem. Res. 24 (1991) 290.
[48] M.C. Wahl, M. Sundaralingam, Trends Biochem. Sci. 22 (1997) 97.
[49] A. Gelbin, B. Schneider, L. Clowney, S.-H. Hsieh, W.K. Olson, H.M. Berman, J. Am. Chem. Soc. 118 (1996) 519.
[50] M. Egli, N. Usman, A. Rich, Biochemistry 32 (1993) 3221.
[51] N. Foloppe, A.D.J. MacKerell, J. Phys. Chem. B. 102 (2000) 6669.
[52] S. Wang, E.T. Kool, Biochemistry 34 (1995) 4125.
[53] M. Egli, S. Portmann, N. Usman, Biochemistry 35 (1996) 8489.
[54] J.I. Gyi, A.N. Lane, G.L. Conn, T. Brown, Nucleic Acids Res. 26 (1998) 3104.
[55] A.M. Bonvin, M. Sunnerhagen, G. Otting, W.F. van Gunsteren, J. Mol. Biol. 282 (1998) 859.