Journal of Biomolecular NMR, **18:** 287–302, 2000.
KLUWER/ESCOM

Conformational dynamics in mixed α/β-oligonucleotides containing polarity reversals: A molecular dynamics study using time-averaged restraints*

James M. Aramini[a], Anwer Mujeeb[b], Nikolai B. Ulyanov[b] & Markus W. Germann[a,**]
[a]*Department of Microbiology and Immunology, Kimmel Cancer Institute, Thomas Jefferson University, Philadelphia, PA 19107, U.S.A.*
[b]*Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143, U.S.A.*

Received 19 June 2000; Accepted 7 September 2000

Key words: AMBER, alpha anomeric, DNA•RNA, molecular dynamics, time-averaged restraints

Abstract

Nucleic acid duplexes featuring a single alpha-anomeric thymidine inserted into each DNA strand via 3′-3′ and 5′-5′ phosphodiester linkages exhibit local conformational dynamics that are not adequately depicted by conventional restrained molecular dynamics (rMD) methods. We have used molecular dynamics with time-averaged NMR restraints (MDtar) to explore its applicability to describing the conformational dynamics of two α-containing duplexes – d(GCGAAT-3′-3′-αT-5′-5′-CGC)$_2$ and d(ATGG-3′-3′-αT-5′-5′-GCTC)•r(gagcaccau). In contrast to rMD, enforcing NOE-based distance restraints over a period of time in MDtar rather than instantaneously results in better agreement with the experimental NOE and *J*-data. This conclusion is based on the dramatic decreases in average distance and coupling constant violations (Δd_{av}, J_{rms}, and ΔJ_{av}) and improvements in sixth-root *R*-factors (R^x). In both duplexes, the deoxyribose ring puckering behavior predicted independently by pseudorotation analysis is portrayed remarkably well using this approach compared to rMD. This indicates that the local dynamic behavior is encoded within the NOE data, although this is not obvious from the local R^x values. In both systems, the backbone torsion angles comprising the 3′-3′ linkage as well as the (high S-) sugars of the α-nucleotide and preceding residue ($\alpha - 1$) are relatively static, while the conformations of the 5′-5′ linkage and the sugar in the neighboring β-nucleotide ($\alpha + 1$) show enhanced flexibility. To reduce the large ensembles generated by MDtar to more manageable clusters we utilized the PDQPRO program. The resulting PDQPRO clusters (in both cases, 13 structures and associated probabilities extracted from a pool of 300 structures) adequately represent the structural and dynamic characteristics predicted by the experimental data.

Abbreviations: CORMA, complete relaxation matrix analysis; EM, energy minimization; MD, molecular dynamics; MDtar, molecular dynamics using time-averaged restraints; NOE, nuclear Overhauser enhancement; ODN, oligodeoxyribonucleotide; PDQPRO, probability distribution by quadratic programming; RANDMARDI, random error MARDIGRAS; rEM, restrained EM; rMD, restrained MD; RMSD, root-mean-square deviation; SANDER, simulated annealing with NMR-derived energy restraints.

Introduction

Elucidating the structure and dynamics of macromolecules at the atomic level is central to gaining an understanding of their action and function in biology. NMR spectroscopy, with its ever-growing technical advances, is especially suited for this purpose. In conventional protocols for NMR structure refinements – including restrained molecular dynamics (rMD), distance geometry, and torsion angle dynamics – the

*Supplementary material pertaining to this article is available from the corresponding author upon request.

**To whom correspondence should be addressed. E-mail: mwg@lac.jci.tju.edu

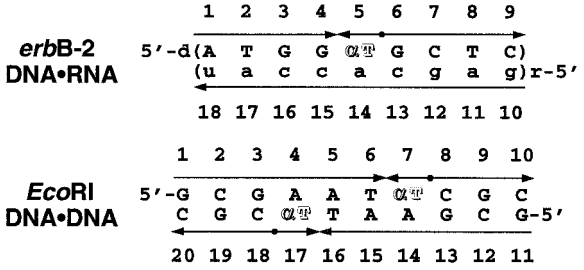

Figure 1. Sequences of the self-complementary *Eco*RI DNA•DNA duplex and *erb*B-2 DNA•RNA nonamer hybrid. The DNA and RNA nucleotides are labeled with upper and lower case letters, respectively; the α-anomeric thymidine is shown in outline type. Arrows indicate the strand polarity in the 5′ → 3′ direction, with the unusual phosphodiester linkages in the DNA strands denoted by head-to-head (3′-3′) and tail-to-tail (5′-5′) junctions. The sequence numbering schemes used throughout this paper (5′ → 3′) are given. Note: to avoid confusion, we have changed the nomenclature for these molecules from that used in our earlier papers (Aramini et al., 1996, 1998a; Aramini and Germann, 1999).

objective is to obtain a structure or pool of structures which best fit all of the experimental (NMR) data (Schmitz and James, 1995; Güntert, 1998). This often masks the presence of conformational flexibility in the molecule by driving a dynamic region into an 'average' structure which may satisfy all of the experimental restraints, but may have little or no physical meaning. In the biomolecular NMR field, this problem has been addressed using a number of strategies. These include (i) time-averaged molecular dynamics, MDtar (Torda et al., 1990), in which restrained molecular dynamics is performed on a single molecule with the requirement that experimental restraints are fulfilled over a period of time rather than instantaneously, and (ii) ensemble approaches in which the experimental restraints are satisfied over several copies of the molecule either generated (Ulyanov et al., 1995; Gyi et al., 1998) or refined *en masse* using multiple-copy refinement (Bonvin and Brünger, 1995; Fennen et al., 1995; Kemmink and Scheek, 1995; Görler et al., 2000) and the promising locally enhanced sampling technique, employed to this point only in unrestrained MD (Simmerling et al., 1998). To date, only a few examples of the successful application of these techniques, primarily MDtar, in nucleic acid NMR structural studies have appeared in the literature (Schmitz et al., 1993; González et al., 1995; Yao et al., 1997; Bachelin et al., 1998). The pool of structures that is generated by MDtar can be further refined by assigning probabilities to selected conformers using the PDQPRO method (Schmitz et al., 1998).

In this study we have employed MDtar followed by PDQPRO analysis to probe the dynamic effects of α-anomeric nucleotides flanked by polarity reversing 3′-3′ and 5′-5′ phosphodiester linkages. The modifications were incorporated into two nucleic acid duplexes: (i) a self-complementary DNA decamer duplex (Aramini et al., 1996, 1998a) and (ii) a DNA•RNA nonamer hybrid (Aramini and Germann, 1999) shown in Figure 1. Both molecules are comprised of DNA strands featuring a single α-thymidine and juxtaposed polarity reversals that switch the orientation of the nucleotide, enabling stable base pair formation with its complement. The DNA duplex corresponds to the core of the Dickerson dodecamer containing a central *Eco*RI recognition sequence, while the target RNA strand in the hybrid corresponds to the initiation region of the *erb*B-2 oncogene, an important marker in breast and ovarian cancers (Slamon et al., 1987, 1989). Interest in the design of chimeric α/β-ODNs containing polarity reversals (Debart et al., 1994; Koga et al., 1995) for potential therapeutic purposes originated from the well-established antisense and nuclease resistance properties of purely α-anomeric ODNs (Bertrand et al., 1989; Lavignon et al., 1992; Vichier-Guerre et al., 1994). Moreover, chimeric α/β-ODNs mediate the cleavage of RNA by RNase H (Boiziau et al., 1995; Aramini et al., 1998b) and have also exhibited *in vivo* antitumor activity (Tan et al., 1998). In our earlier structural work on the α-containing *Eco*RI DNA duplex and *erb*B-2 DNA•RNA hybrid we recognized that these molecules possess a number of localized dynamic 'hot spots' that cannot be adequately described using conventional rMD approaches (Aramini et al., 1998a; Aramini and Germann, 1999). Here we will demonstrate that MDtar coupled with PDQPRO can superbly represent the conformational dynamics in these molecules predicted by NOE and *J*-coupling information. To our knowledge this is the first report that exploits the strengths of MDtar and PDQPRO to probe the dynamics in ODNs containing α-anomeric nucleotides and adjacent polarity reversals. Understanding the impact of these modifications on nucleic acid structure and dynamics is relevant to the design of chimeric α/β-oligonucleotides with optimal antisense properties.

Table 1. Statistics of NMR restraints and parameters for the rMD and MDtar trajectories and PDQPRO ensembles of *Eco*RI DNA•DNA and *erb*B-2 DNA•RNA

I. Restraints[a]	*Eco*RI DNA•DNA		*erb*B-2 DNA•RNA	
	Number	*k*	Number	*k*
Quantitative distance restraints (RANDMARDI)				
1. Non-exchangeables (total)	374	21	246	20
2. Exchangeables (total)	0	–	22	16
Qualitative distance restraints	72	21	1	20
Watson–Crick distance restraints	26	17.5	23	20
Watson–Crick flat angle restraints	26	7	23	8
Broad backbone torsion angle restraints (α, β, γ, ε, ζ)[b]	70	10	2	8
Other torsion angle restraints	0	–	1	40
Total number of restraints	568		318	

II. Structural parameters						
	*Eco*RI DNA•DNA			*erb*B-2 DNA•RNA		
Parameter[c]	rMD	MDtar	PDQPRO	rMD	MDtar	PDQPRO
$<E_{AMBER}>$	−44785 (60)	−44719 (80)	−44706 (56)	−36665 (56)	−36617 (61)	−36611 (51)
$<E_{constraint}>$	192.9 (8.1)	98.1 (11.1)	90.2 (4.6)	108.3 (6.5)	52.9 (4.7)	53.0 (4.7)
Δd_{av}						
All (D_2O)	0.088	0.067	0.082	0.080	0.064	0.075
DNA	–	–	–	0.083	0.066	0.076
RNA	–	–	–	0.074	0.059	0.073
R^x_{ens}						
Intra	0.0323	0.0279	0.0269	0.0413	0.0397	0.0411
Inter	0.0387	0.0435	0.0382	0.0438	0.0434	0.0416
Total	0.0346	0.0333	0.0308	0.0420	0.0408	0.0412
J_{rms}	1.46	0.90	0.97	1.39	0.74	0.99
ΔJ_{av}	0.56	0.28	0.33	0.64	0.19	0.31
RMSD	0.78 (0.13)	1.48 (0.27)	1.07	0.74 (0.11)	1.09 (0.17)	0.77

[a]The restraints used in the rMD equilibration and final rMD and MDtar production runs are identical to those employed in the respective *in vacuo* structure determinations (Aramini et al., 1998a; Aramini and Germann, 1999) with the following exceptions: all deoxyribose ring (ν_0–ν_4) torsion angles restraints were omitted; well widths for a small number of troublesome restraints (<5) were increased; values given for the force constants (*k*, in units of kcal/(mol · Å^2) and kcal/(mol · rad^2) for the distance and (torsion) angle restraints, respectively) correspond to 70 and 80% of those used in the previous *in vacuo* calculations.

[b]Broad right-handed backbone restraints employed for the DNA duplex are as follows (Mujeeb et al., 1993): α, −90 to −30°; β, 135 to 215°; γ, 30 to 90°; ε, 140 to 300°; ζ, 150 to 315°. Weak γ^+ restraints and an additional broad δ restraint were employed in the DNA•RNA calculation to minimize spurious end effects.

[c]Definitions of structural parameters: $<E_{AMBER}>$, average AMBER (non-constraint) energy (kcal/mol); $<E_{constraint}>$, average constraint energy (kcal/mol); Δd_{av}, third-root average distance bound violation (Å) for non-exchangeable interproton distance restraints; R^x_{ens}, sixth-root CORMA ensemble *R*-factor (intraresidue, interresidue, and total); J_{rms}, average root-mean-square deviation from experimental *J*-coupling constants; ΔJ_{av}, average deviation from the upper and lower experimental *J*-coupling constants; RMSD, all-atom root-mean-square deviation from the start of the trajectory (rMD and MDtar) or within the cluster (PDQPRO). Standard deviations for $<E_{AMBER}>$, $<E_{constraint}>$, and RMSD are shown in parentheses. R^x-factors were calculated using NOE data for the longest mixing time D_2O spectrum and experimental correlation times (*Eco*RI DNA•DNA: τ_m = 300 ms; τ_c = 3.0 ns; *erb*B-2 DNA•RNA: τ_m = 250 ms; τ_c = 3.5 ns). The number of distances and NOE volumes used to obtain the Δd_{av} and R^x-factors are as follows: *Eco*RI DNA•DNA: Δd_{av}: all, 374; R^x: all, 394; *erb*B-2 DNA•RNA: Δd_{av}; all, 246, DNA, 162, RNA, 82; R^x: all, 256. Coupling constant data for all deoxyribose residues in the DNA duplex except C10/C20 and the entire DNA strand of the hybrid were used in the calculations of J_{rms} and ΔJ_{av}.

Materials and methods

In vacuo *structure determinations*

Techniques and methodologies used in the synthesis, NMR spectroscopy and analysis, and solution structure determinations of both the *Eco*RI DNA duplex and *erb*B-2 DNA•RNA hybrid were described in our earlier reports (Aramini et al., 1996, 1998a; Aramini and Germann, 1999). Briefly, both initial structures were elucidated *in vacuo* using the SANDER program within the AMBER 4.1 suite and the 1994 all-atom nucleic acid parameterization (Pearlman et al., 1995). All rMD and rEM calculations were driven by empirical interproton distance and deoxyribose ring torsion angle restraints obtained by RANDMARDI (Liu et al., 1995) and pseudorotation (van Wijk et al., 1992) analysis of homonuclear NOE and *J*-coupling data, respectively. The deposited final structures [PDB accession numbers: 1BX5 and 1C2Q], used as starting models in this study, were derived by coordinate averaging of the pool of structures in the final stage of the protocol followed by rEM.

Structural restraints

Structural restraints used in this study, enforced within SANDER in the form of flat-well pseudo energy terms, are based on those employed in our earlier *in vacuo* structure calculations (Aramini et al., 1998a; Aramini and Germann, 1999); see Table 1 for detailed restraint lists for both duplexes. Deoxyribose torsion angle restraints derived from coupling constant (*J*) data were *not* included, hence any conformational fluctuations observed in the dynamics runs ultimately are due to the NOE data. We opted not to use *J*- or torsion angle restraints in this work for several reasons: (i) it has been documented that the use of *J*- and/or torsion angle restraints in MDtar calculations can lead to unnaturally large structural fluctuations (Nanzer et al., 1997; Scott et al., 1998); (ii) the SANDER program is only equipped to internally convert *J*-restraints into the corresponding torsion angles using a simple Karplus relationship, not the more sophisticated generalized Karplus formalism of Altona (see Appendix I); (iii) the *J*-data serve as an independent experimental parameter for evaluation of the calculated structure ensembles. In the DNA duplex case only, broad right-handed backbone torsion angle restraints (α–ζ except δ; Mujeeb et al., 1993) were applied to all standard 3′-5′ phosphodiester bonds, on the basis of ^{31}P chemical shift and the fine structure of the characteristically weak H5′/H5″-^{31}P correlation cross peaks (Pikkemaat and Altona, 1996). The modified 3′-3′ and 5′-5′ linkages were left totally unconstrained in all calculations. A number of control MDtar runs on the DNA•RNA system established that the backbone restraints do not impinge upon the dynamic motions (i.e., sugar repuckering) in the molecule, and the optimal force constants (in terms of the balance between the energy of the system and the agreement with experimental restraints) to be used in the final production runs (*vide infra*).

rMD and MDtar calculations

The rMD and MDtar calculations presented here were conducted with explicit water molecules and sodium counterions using the SANDER program (AMBER 4.1) on single processor SGI Indigo2 R10000 and 500 MHz dual PentiumIII computers. In all computations, electrostatic interactions were treated by the particle mesh Ewald (PME) method (Darden et al., 1993) using cubic B-spline interpolation and $\approx$ 1 Å grid spacing in each dimension. All dynamics calculations were performed at constant pressure as single runs using the following typical parameters: a 9 Å cut-off for non-bonded interactions, SHAKE (Ryckaert et al., 1977) on bonds involving hydrogen atoms, Berendsen temperature coupling (Berendsen et al., 1984), a 0.2 fs time step (0.1 fs in the MDtar runs), and non-bonded pair list updates every 10 steps. Atomic coordinates were saved every 500 steps, resulting in 150 and 300 written frames for each 150 ps rMD and MDtar run, respectively.

Preparation and equilibration for rMD/MDtar

In preparation for the dynamics runs, the final *Eco*RI DNA•DNA and *erb*B-2 DNA•RNA *in vacuo* structures were solvated and equilibrated following standard protocols (Cheatham and Kollman, 1997). Using the sub-program LEaP within AMBER, both final *in vacuo* structures were placed in periodic solvent boxes containing 10 Å of TIP3P water molecules surrounding the solute in each direction, and subsequently neutralized with Na^+ counterions. With the position of the nucleic acid kept fixed, the counterions and solvent molecules were equilibrated using a 1000 step steepest descent EM followed by a 25 ps MD at 300 K ($T = 100 \rightarrow 300$ K in the first ps), and then a series of 1000 step EM with decreasing positional restraints culminating in a 1000 step rEM. Next, the entire system was equilibrated over the course of a 25 ps rMD at 300 K ($T = 100 \rightarrow 300$ K in the

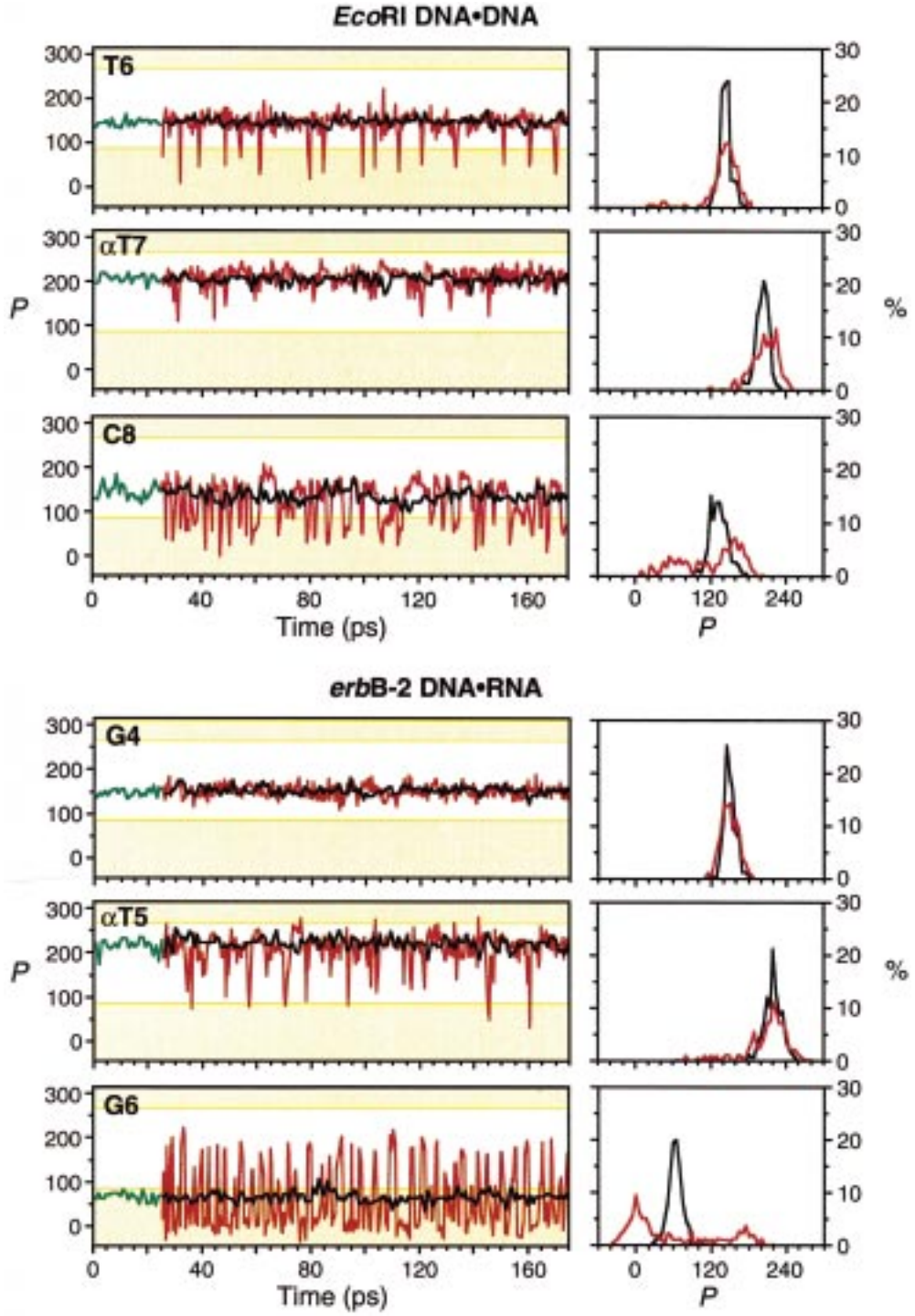

Figure 2. Time courses and relative populations of the pseudorotation phase angle (*P*) for selected nucleotides from the 150 ps rMD (black) and MDtar (red) calculations on *Eco*RI DNA●DNA and *erb*B-2 DNA●RNA in water, following a 25 ps rMD equilibration (green). Values of *P* corresponding to the N-hemisphere of the pseudorotation wheel are shaded yellow. Relative pseudorotation percentages represent 5° windows from 0° to 360°.

first 2 ps), in which the restraint force constants were identical to those used *in vacuo*.

rMD and MDtar production runs

Following the 25 ps rMD equilibration, final rMD and MDtar runs of 150 ps were performed on the *Eco*RI DNA duplex and *erb*B-2 DNA•RNA hybrid using attenuated force constants (70% and 80% of the *in vacuo* values, respectively; Table 1). In the time-averaging runs, the force constants were ramped up in the first 2 ps from a starting value of 10% to prevent anomalous heating at the beginning of the calculations (Yao et al., 1997). Only the non-exchangeable quantitative (i.e., RANDMARDI) distance restraints were 'loosened' by third-root time-averaging using a damping constant of 10 ps and the 'pseudoforce' option (Pearlman and Kollman, 1991; Schmitz et al., 1993) within AMBER 4.1;[1] all other restraints (Table 1) were applied continuously.

Analysis

Values of several structural parameters, including backbone torsion angles, pseudorotation phase angles, endocyclic torsion angles, and atomic RMSDs were computed from snapshots of atomic coordinates recorded along the dynamics trajectory using the CARNAL program within AMBER. Ensemble sixth-root R^{x} factors (Equation 1) (Thomas et al., 1991), which reflect the difference between experimental (a_o) NOE cross-peak volumes and calculated values (a_c), were obtained from the dynamics trajectories assuming an isotropic (experimental) correlation time, rapid molecular interconversion, and equal probability of each snapshot using the complete relaxation matrix analysis program CORMA (Keepers and James, 1984; Borgias and James, 1988).

$$R^{x} = \frac{\Sigma|a_o(i)^{1/6} - a_c(i)^{1/6}|}{\Sigma|a_o(i)^{1/6}|} \tag{1}$$

Global helical parameters were extracted from each dynamics time course using the Dials and Windows package (Ravishanker et al., 1989). The average root-mean-square deviation, J_{rms} (Equation 2) and average J violation, ΔJ_{av} (Equation 3), are used as indicators of the agreement between experimental deoxyribose ^{1}H-^{1}H J-coupling constants (J_{exp}) determined in our earlier work (Aramini et al., 1998a; Aramini and Germann, 1999) and those calculated from the model structures (J_{calc}).

$$J_{rms} = \sqrt{\Sigma(J_{calc} - J_{exp})^2}/N, \tag{2}$$

$$\Delta J_{av} = \Sigma \begin{Bmatrix} J_{calc} - J_{up} & \text{if } J_{calc} > J_{up} \\ 0 & \text{if } J_{lo} \leq J_{calc} \leq J_{up} \\ J_{lo} - J_{calc} & \text{if } J_{calc} < J_{lo} \end{Bmatrix} /N, \tag{3}$$

where J_{up} and J_{lo} denote the upper and lower experimental coupling constants, respectively.[2] For all dynamics runs, values of four vicinal coupling constants – $J_{1'2'}$, $J_{1'2''}$, $J_{2'3'}$, and $J_{2''3'}$ – were calculated per deoxyribose nucleotide at each snapshot along the trajectory on the basis of the corresponding endocyclic torsion angles (ν_1 and ν_2) using the generalized Karplus equations and parameterization developed by Altona and co-workers (van Wijk et al., 1992, 1995; Donders et al., 1989; Altona et al., 1994); the average values of every J-coupling over the entire run were used in the above expressions to obtain J_{rms} and ΔJ_{av} (see Appendix I).

PDQPRO calculations

Representative structural ensembles were selected from the entire 150 ps MDtar trajectories using the PDQPRO algorithm (Ulyanov et al., 1995; Schmitz et al., 1998). Briefly, the program finds a sub-set of structures and their probabilities, p^{α}, that yield the best agreement between theoretical, T_k, and experimental, E_k, dipolar relaxation rates by minimizing the following quadratic function (Equation 4):

$$Q^{r}(\{p^{\alpha}\}) = \Sigma w_k(T_k - E_k)^2, \tag{4}$$

where w_k is the weighting for each relaxation rate (assumed to be equal in this work). Experimental homonuclear cross-relaxation rates, E_k, were derived from the NOE data using RANDMARDI. Three 2D NOESY data sets were used for each molecule, and the calculations were performed assuming isotropic

[1] Longer MDtar trial runs on the DNA•RNA system using longer damping constants (i.e., $\tau = 20$ ps) produced very similar results, meaning that the value chosen for this study (10 ps) does not result in aberrant conformational behavior (Schmitz and James, 1995).

[2] Of the four ^{1}H-^{1}H coupling constants used in our work, $J_{1'2'}$ and $J_{2''3'}$ ($J_{1'2''}$ and $J_{2''3'}$ for α-anomeric nucleotides) are quite sensitive to changes in pseudorotation phase angle, and hence have the most impact on J_{rms} and ΔJ_{av}.

correlation times of 3.0 and 3.5 ns for the DNA•DNA duplex and DNA•RNA hybrid, respectively. For each rate, the 20% lowest and highest values were discarded, and the remaining data were averaged to yield the final E_k values used in the PDQPRO calculations. The ensemble-averaged theoretical relaxation rates, T_k, were computed using Equation 5 from the individual cross-relaxation rates, R_k, and probabilities, p^α. Individual rates, R_k, for each structure in the ensemble were calculated with CORMA using the highest mixing time NOE data sets.

$$T_k = \Sigma p^\alpha R_k^\alpha. \quad (5)$$

In all PDQPRO calculations, data for interproton distances that are insensitive to conformational changes (i.e., H5-H6 and H1′-H2′) were omitted. For both molecules an initial round of PDQPRO calculations reduced the pool of structures from the MDtar trajectory to ca. 40. Structures with $p^\alpha < 3\%$ were discarded and PDQPRO was run again to yield the final ensemble of 13 in each case. Structural figures of merit were calculated as discussed above taking into account the probability of each structure in the ensemble. All molecular graphics figures of the PDQPRO ensembles were prepared using the MidasPlus 2.1 program (Ferrin et al., 1988).

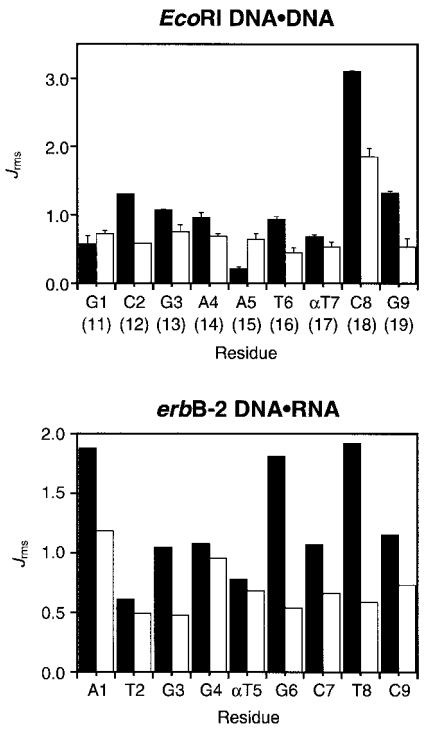

Figure 3. Plots of J_{rms} per nucleotide from the 150 ps rMD (black) and MDtar (white) calculations on *Eco*RI DNA•DNA and *erb*B-2 DNA•RNA in water. For the self-complementary duplex each value plus upper bound shown was obtained by averaging the J_{rms} for the symmetry-related nucleotides in the two strands. The two 3′-terminal nucleotides in the DNA duplex (C10/C20) are omitted because a complete pseudorotation analysis was not performed on these nucleotides due to H2′/H2″ chemical shift degeneracy.

Results

rMD vs MDtar calculations

For both *Eco*RI DNA•DNA duplex and *erb*B-2 DNA•RNA hybrid, molecular dynamics runs using standard and time-averaged restraints were typically quite stable, as judged by the modest fluctuations in various system variables; several parameters that are indicative of the behavior of the solute molecule in these trajectories are listed in Table 1. In both cases, allowing the non-exchangeable RANDMARDI-derived distance restraints to be satisfied over a period of time results in larger sampling of conformational space in the MDtar calculations; this is reflected by the increased atomic RMSDs, although there is little change in the overall energy of the system. However, both the average distance and *J*-coupling constant violations markedly decrease when time-averaged restraints are employed compared to conventional rMD, along with some improvement in the sixth-root R^x-factors. Thus, in global terms the enhanced conformational sampling afforded by MDtar results in an improved fit to both NOE-based distance restraints and the experimental *J*-data, although the latter were not applied as experimental restraints and serve as independent structural monitors. In the following sections we explore the nature of the conformational dynamics at the level of (i) the sugar puckering, (ii) backbone, and (iii) helical parameters, with particular emphasis on the regions spanning the α-anomeric nucleotide and unnatural phosphodiester linkages in both molecules. In general, we will refer to only one strand in the self-complementary DNA duplex, since the behavior of both strands is essentially identical.

Sugar pucker dynamics

The rMD and MDtar time courses and population distributions for the pseudorotation phase angles of the α-thymidines and flanking β-anomeric residues in *Eco*RI DNA•DNA and *erb*B-2 DNA•RNA are shown in Figure 2. In both molecules the αT and the residue

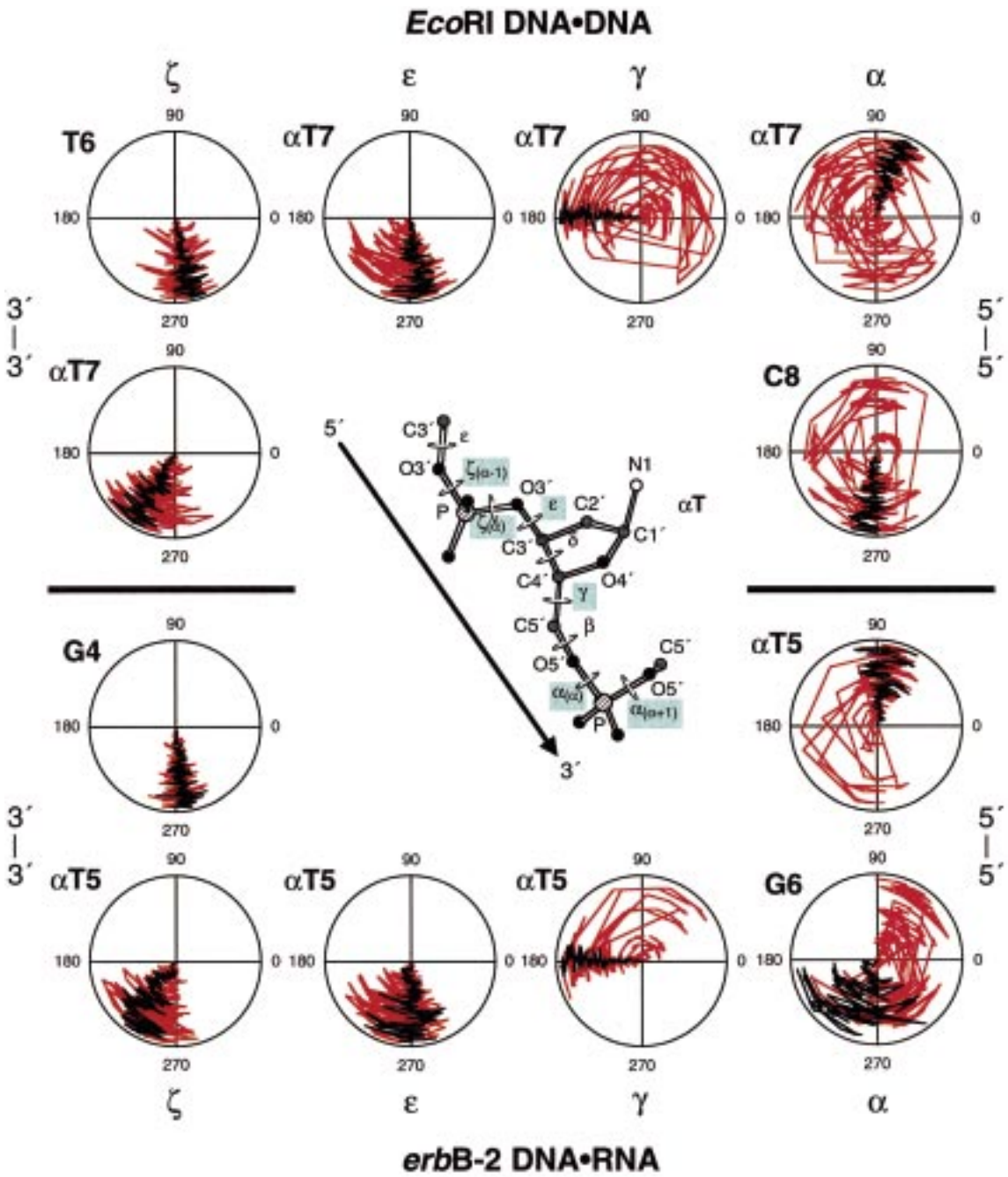

Figure 4. Polar plots for selected backbone torsion angles from the 150 ps rMD (black) and MDtar (red) calculations on *Eco*RI DNA●DNA and *erb*B-2 DNA●RNA in water. The time axis radiates out from the origin. The corresponding torsion angles are highlighted in the schematic diagram of the α-nucleotide and flanking 3′-3′ and 5′-5′ phosphodiester linkages; see Aramini et al. (1998a) and Aramini and Germann (1999) for definitions of the backbone torsion angles within the unnatural linkages.

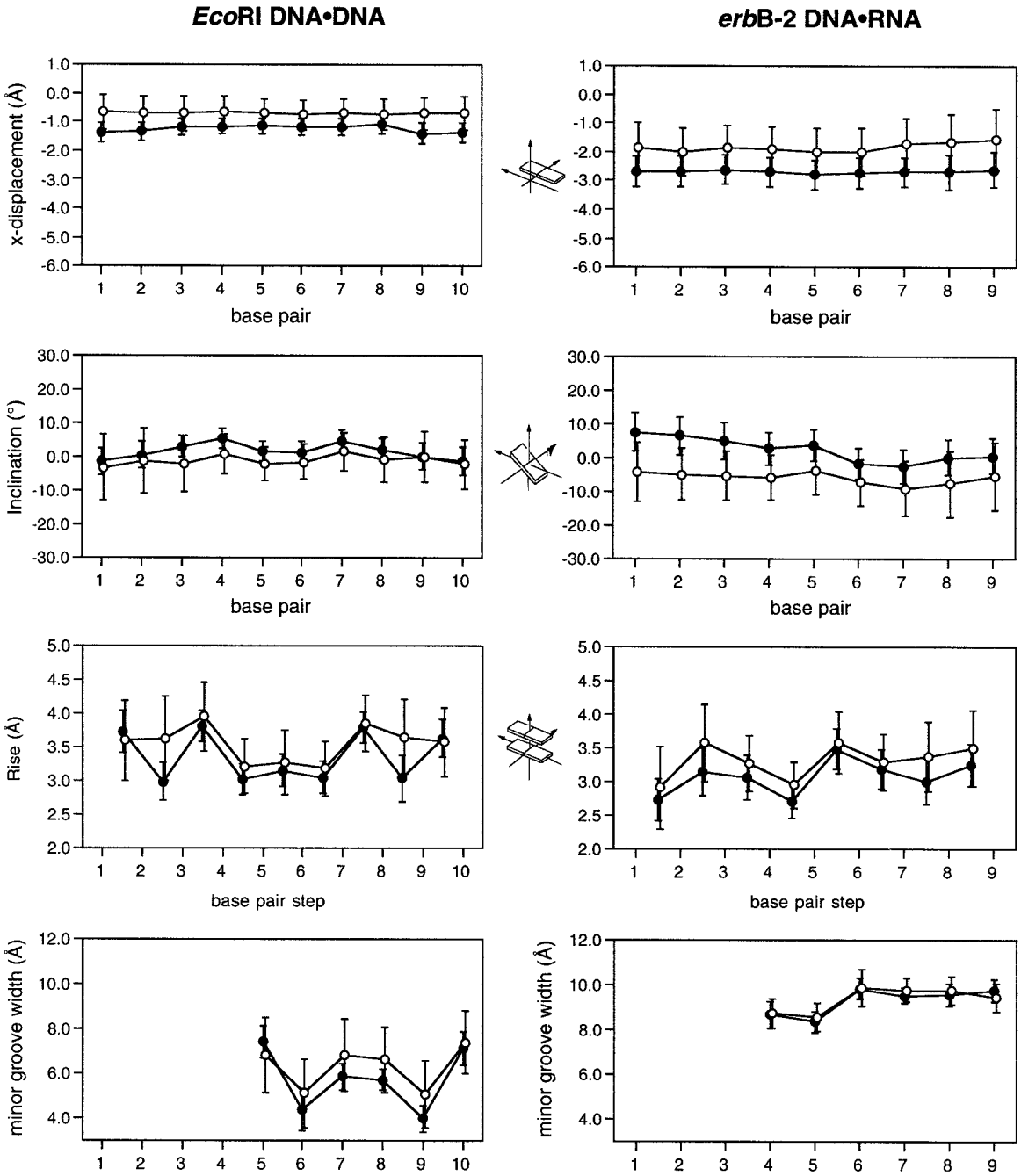

Figure 5. Selected helicoid parameters (dx, η, Dz) and minor groove widths from the 150 ps rMD (●) and MDtar (○) calculations on *Eco*RI DNA•DNA and *erb*B-2 DNA•RNA in water. For clarity, the MDtar points in each plot are shifted slightly. Minor groove widths correspond to the average minimum cross-strand P-P distance minus 5.8 Å (Bhattacharyya and Bansal, 1992) obtained as follows: (i) DNA•DNA: P5-P20, P6-P19, P7-P18, P8-P17, P9-P16, P10-P15; (ii) DNA•RNA: P4-P18, P5-P17, P6-P16, P7-P15, P8-P14, P9-P13.

Table 2. Mole fraction of S-pucker for *Eco*RI DNA•DNA and *erb*B-2 DNA•RNA: MDtar *vs.* pseudorotation theory[a]

*Eco*RI DNA•DNA			*erb*B-2 DNA•RNA				
Residue	f_S MDtar[b]	f_S *J*-data[c]	Residue	f_S MDtar	f_S *J*-data[c]	Residue	f_S MDtar
G(1)1	0.94	0.85–0.98	A1	0.44	0.54–0.66	g10	0.26
C(1)2	0.80	0.80–0.85	T2	0.96	0.79–0.93	a11	0
G(1)3	0.98–0.99	0.90–1.00	G3	0.88	0.81–0.95	g12	0
A(1)4	0.98–1.00	0.89–1.00	G4	1.00	0.85–0.98	c13	0
A(1)5	0.84–0.85	0.89–1.00	αT5	0.97	0.80–1.00	a14	0
T(1)6	0.88–0.93	0.85–1.00	G6	0.33	0.23–0.35	c15	0
αT(1)7	1.00	0.80–1.00	C7	0.09	0.10–0.24	c16	0
C(1)8	0.69–0.72	0.44–0.55	T8	0.45	0.46–0.56	a17	0
G(1)9	0.84–0.87	0.84–0.94	C9	0.40	0.49–0.57	u18	0.21

[a] f_S values shown for all MDtar runs were obtained by summation of all conformations in the entire S-range ($P = 90$ to 270°).
[b] Range of f_S for symmetry related residues in the complementary strands.
[c] Calculated from the experimental *J*-coupling constants – $J_{1'2'}$, $J_{1'2''}$, $J_{2'3'}$, $J_{2''3'}$ – based on pseudorotation theory assuming a two-state model (Aramini et al., 1998a; Aramini and Germann, 1999).

preceding it via a 3′-3′ phosphodiester bond (T6 in the DNA duplex and G4 in the hybrid) exhibit a high percentage of S-puckering, although the α-nucleotide resides in a different window of the pseudorotation wheel ($P \approx 215°$; C3′-*exo*/C4′-*endo*) compared to that normally observed for β-anomeric deoxyribose sugars ($P \approx 155°$; C2′-*endo*). However, for the nucleotide following the 5′-5′ linkages (C8 in the DNA duplex and G6 in the hybrid), frequent interconversions between the N- and S-conformers occur over the course of the MDtar trajectory, resulting in an approximately bimodal distribution of pseudorotation phase angles (Figure 2). This is a striking departure from the rMD calculations in which the continuous enforcement of the restraints restricts the conformation of these sugars to an envelope somewhere in between the N- and S-forms. Moreover, the enhanced conformational variability of the deoxyribose rings in these residues seen in the MDtar results in a significant drop in the J_{rms} values compared to those resulting from the relatively static rMDs, indicative of better agreement with the experimental coupling constants (Figure 3). Overall, enhanced sugar repuckering and concomitant decreases in J_{rms} are observed for certain deoxyribonucleotides within these sequences, most notably G1 and T8 in the hybrid, and to a lesser extent C2 and G9 in the DNA duplex. In addition, the fraction of S-pucker, f_S, observed in the MDtar runs generally correlates remarkably well with values independently predicted by pseudorotation analysis of the experimental coupling constants, assuming a simple two-state equilibrium (Table 2). In the DNA duplex, all residues are generally highly S-type, and f_S for C8 is reduced, although not to the extent expected on the basis of the *J*-coupling data. In the DNA•RNA hybrid, excellent agreement is observed across the DNA strand, which exhibits a wide range of puckering behavior; high S-type sugars preceding the 3′-3′ linkage (T2, G3, G4), high N-type sugars following the 5′-5′ linkage (G6, C7), and ≈ 50:50 cases at the ends (A1, T8, C9).[3] Also, aside from the terminal residues, all ribose sugars exclusively adopt an N-pucker, in accordance with the minute $J_{1'2'}$ couplings observed for these residues (Aramini and Germann, 1999). In all cases, the glycosidic and backbone torsion angles χ and δ are intimately correlated with the dynamics of the sugar ring (data not shown).

Backbone dynamics

The dynamic behavior of selected backbone torsion angles in the MDtar runs of *Eco*RI DNA•DNA and *erb*B-2 DNA•RNA is presented in Figure 4. In both molecules, the time dependence of the ζ torsion angles of the α − 1 and α-nucleotides suggests that the movement of the 3′-3′ linkage is somewhat restricted. Conversely, large oscillations in the α-γ torsion angles of the α and α + 1 residues reflect the enhanced mobility of the 5′-5′ linkage. Although the deletion

[3] The unusual high S-character of both G3 and G4 in this hybrid is a direct consequence of the modifications, and is not observed in the corresponding unmodified control hybrid (Aramini and Germann, 1999).

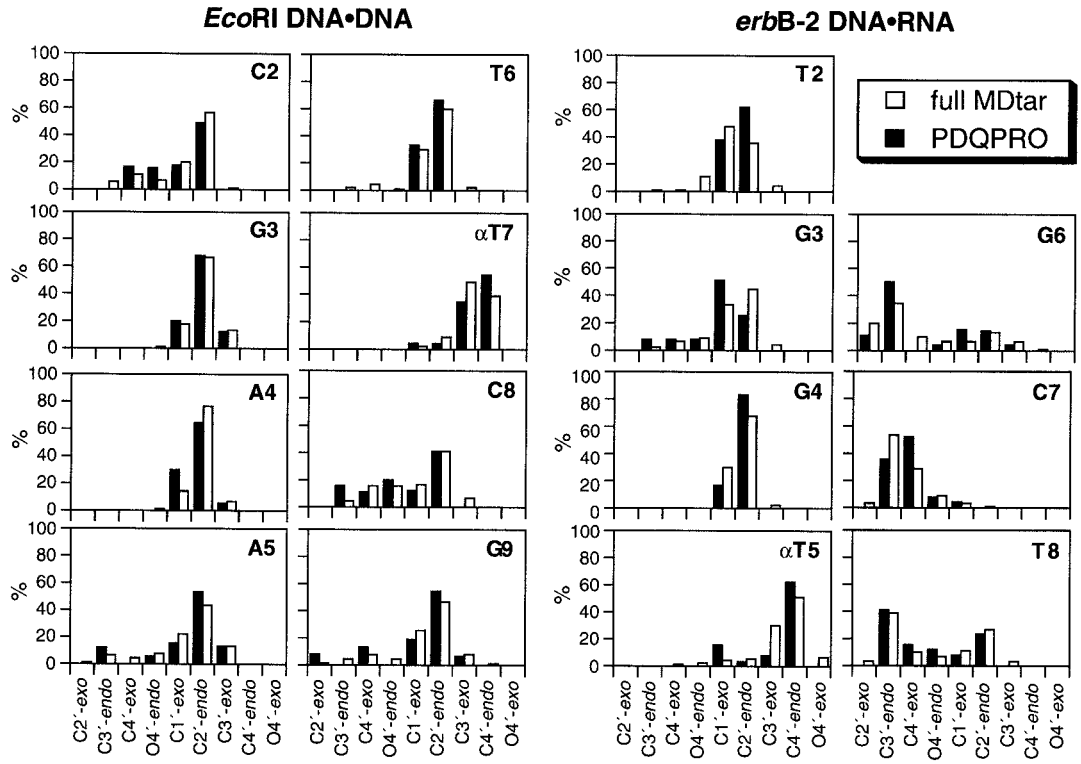

Figure 6. Pseudorotation phase angle histograms of the non-terminal deoxyribose residues from the full MDtar trajectories (white) and PDQPRO ensemble (black) of *Eco*RI DNA•DNA and *erb*B-2 DNA•RNA. The puckering states shown represent *P* windows of 36° (i.e., C3′-*endo*: 0 to 36°; C2′-*endo*: 144 to 180°).

or insertion of a methylene group conferring rigidity or added flexibility to a phosphodiester linkage seems intuitively obvious, this behavior is not manifested in the rMD calculations, nor for that matter in long (1 ns) MD production runs (data not shown). The ε torsion angle for each α-nucleotide resides in a rather odd window centered at $\approx -100°$, yet consistent with the increased $^{3}J_{H3'-P}$ couplings extracted from 2D ^{31}P-^{1}H correlation spectra (Aramini et al., 1998a; Aramini and Germann, 1999). Also, the ca. 50:50 population of γ^{+} and γ^{t} rotamers predicted for the α-nucleotides in both molecules on the basis of $^{3}J_{4'5''}$ couplings is very well represented in the MDtar results for the DNA duplex; a somewhat higher percentage of γ^{t} was observed for the hybrid. The remainder of the unconstrained backbone in the hybrid MDtar run is relatively static, and we rarely observed isolated α/γ crankshafting events ($\alpha^{-}/\gamma^{+} \rightarrow \alpha^{t}/\gamma^{t}$) reported in other hybrid studies (González et al., 1995); this result is, however, consistent with the picture of a more rigid backbone predicted for such molecules described in free MD calculations (Cheatham and Kollman, 1997).

Helicoid parameters and minor groove width

In general, the average values for the 16 global helical parameters extracted from the rMD and MDtar runs are highly superimposable for the respective molecules, although the standard errors observed in the MDtar runs are larger, consistent with increased conformational searching around the mean values under time-averaged conditions. Plots of the average rMD and MDtar values of three global helical parameters sensitive to secondary structure in nucleic acid duplexes – *x*-displacement (dx), inclination (η), and rise (Dz) – for *Eco*RI DNA•DNA and *erb*B-2 DNA•RNA are shown in Figure 5. In both dynamics runs, the DNA duplex displays values consistent with a B-motif, while in the hybrid dx and η are shifted away from A-like values in the MDtar run compared to the rMD; these phenomena have been reported previously for DNA•RNA hybrids (González et al., 1995). A prominent feature of the helical parameters of both molecules is a pronounced positive roll angle in all pyrimidine-purine (Y-R) steps (data not shown). Positive roll is associated with localized compression

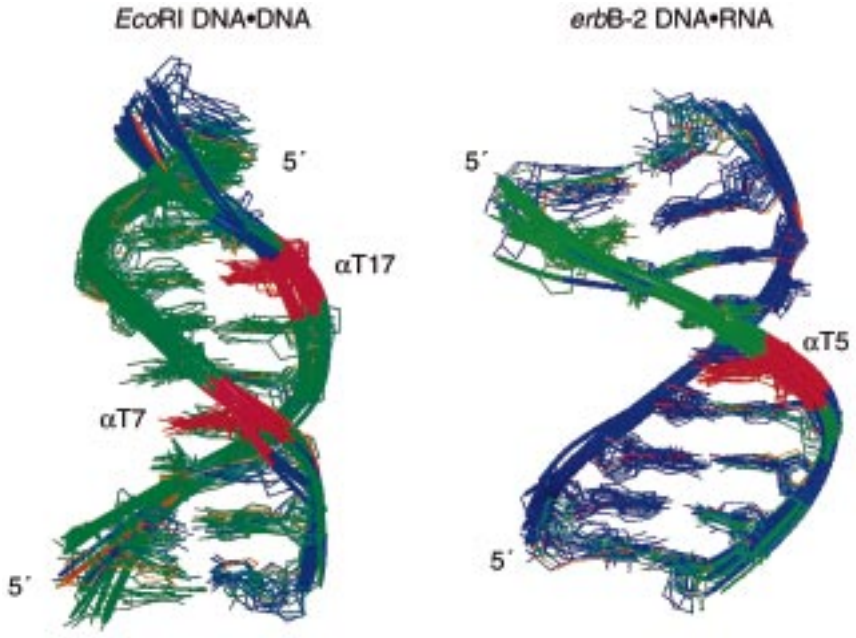

Figure 7. Heavy atom superpositions of the PDQPRO structures extracted from the MDtar trajectories of *Eco*RI DNA•DNA and *erb*B-2 DNA•RNA and the corresponding *in vacuo* structures (orange). Residues adopting N- and S-type sugar puckering are shown in blue and green, respectively; α-nucleotides are depicted in red. The backbone thickness of each PDQPRO structure is proportional to its probability.

of the major groove within a double helix, and is a common feature of Y-R dinucleotides (Ulyanov and Zhurkin, 1984; Gorin et al., 1995; Ulyanov and James, 1995). In the case of *Eco*RI DNA•DNA, a pair of positive (Y-R) and flanking negative roll steps separated by a helical half-turn results in an essentially straight duplex.

On the basis of interstrand P-P distances, the average minor groove width of the DNA duplex is consistent with a B-like motif (narrow minor groove). In agreement with our *in vacuo* structure (Aramini and Germann, 1999), the minor groove width of the hybrid is intermediate between canonical A-RNA and B-DNA, and shows a distinct reduction preceding the 3′-3′ linkage. Notice that the backbone of the B-DNA duplex is inherently more flexible than that of the A-like DNA•RNA hybrid, as indicated by the larger error bars in the minor groove values for the DNA duplex, consistent with MD calculations on such molecules (i.e., Cheatham and Kollman, 1997).

PDQPRO ensembles

By using the PDQPRO approach, we have distilled the MDtar trajectories of both duplexes to a handful of structures with associated probabilities. In contrast to rMD and MDtar, which attempt to satisfy distance bounds derived from NOE data, PDQPRO selects structures from a large calculated pool on the basis of the original NOE data (relaxation rates). For both the DNA duplex and DNA•RNA hybrid, the program yields a pool of 13 structures that collectively maintain good agreement with the experimental data in terms of R^x and J_{rms} (Table 1). Comparing the intra- and intermolecular contributions to R^x, we note that R^x_{inter} is significantly lower in the PDQPRO subset compared to the entire MDtar trajectory, whereas R^x_{intra} is essentially the same; R^x_{inter} is more sensitive to the helical structure of the duplex. Moreover, the dynamics of the (deoxy)ribose rings and backbone torsion angles observed across the complete trajectories are captured by the reduced ensembles. By way of an example, histograms comparing the pseudorotation phase angles of DNA strands in *Eco*RI DNA•DNA and *erb*B-2 DNA•RNA clearly show that the PDQPRO ensembles correctly represent the puckering behavior of both molecules across the entire MDtar runs (Figure 6). However, note that for the residues that exhibit significant puckering dynamics, the behavior is generally not consistent with a simple two-state N ↔ S equilibrium since, as noted previously (Ulyanov et al., 1995), there is a small predicted population of numerous intermediate conformations (i.e., O4′-*endo*). Also, in some

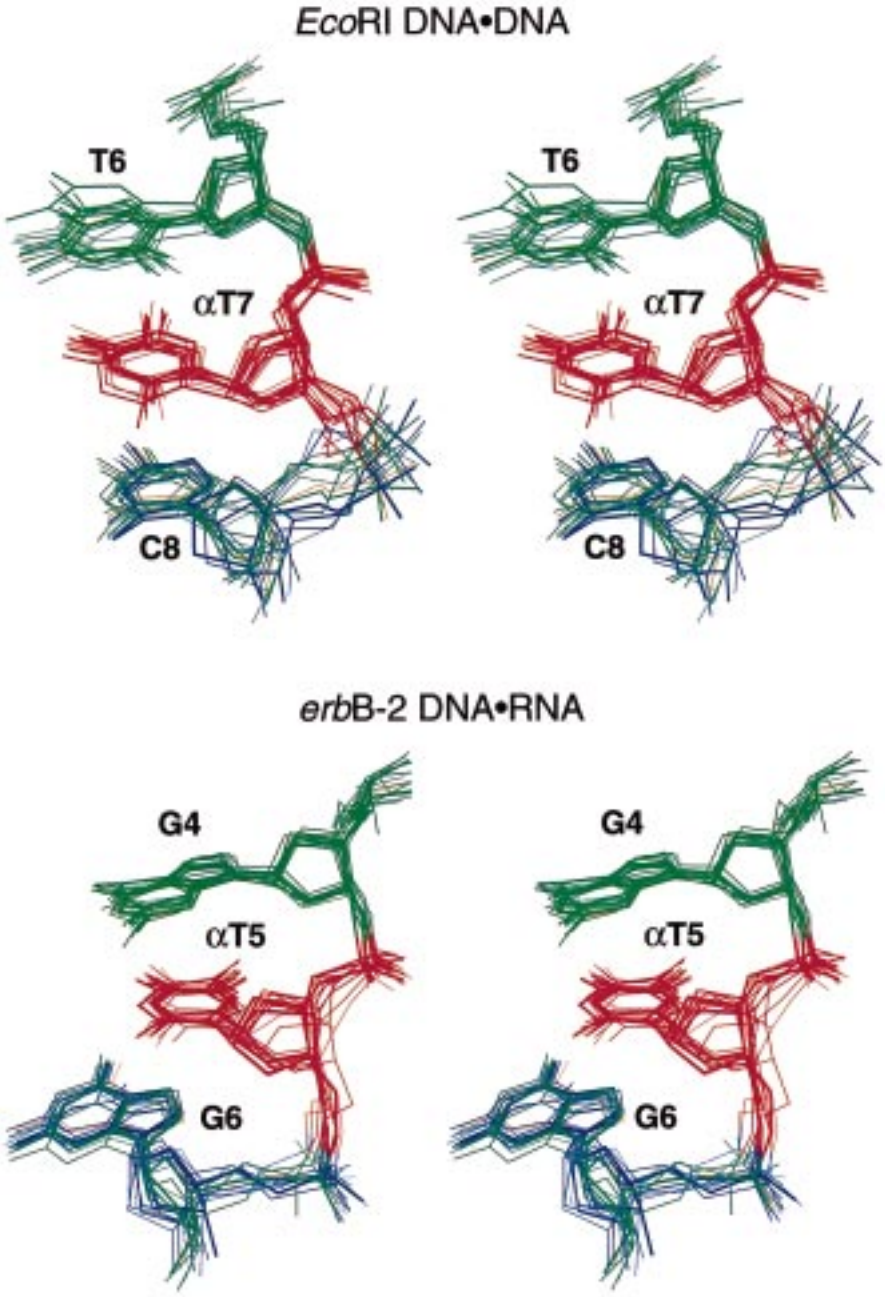

Figure 8. Stereoviews of the overlayed PDQPRO structures of *Eco*RI DNA•DNA and *erb*B-2 DNA•RNA and *in vacuo* structures showing trinucleotide segments encompassing the α-anomeric thymidine and neighboring β-anomeric nucleotides. Same color scheme as in Figure 7.

cases (i.e., C8 in the DNA duplex) the pseudorotation phase angle of the N-conformer is somewhat higher than what is generally assumed in a two-state model (i.e., C3′-*endo*). The smaller pools of structures afford a better means of visualizing the structural and dynamic features of these molecules; superimposed PDQPRO structures extracted from the two MDtar runs showing the entire molecules and trinucleotide stretches encompassing the modifications are shown in Figures 7 and 8. Again we see that enhanced conformational flexibility is primarily manifested in the sugar-phosphate backbone, particularly the 5′-5′ link-

age and the sugar ring of the $\alpha + 1$ nucleotide, and not in the core (bases) of the duplexes.

Conclusions

We have used molecular dynamics driven by time-averaged restraints (MDtar) combined with PDQPRO to shed light on the conformational dynamics within a DNA duplex and a DNA•RNA hybrid containing 3′-3′-αT-5′-5′ units. For both the DNA duplex and the hybrid relaxing only the distance restraints by the time-averaging condition results in dynamics trajectories that significantly improve the agreement with both the NOE and *J*-coupling data; this in spite of the different dependence of these parameters on conformational averaging. In both cases, while rMD provides the correct global structural picture, the MDtar approach illuminates the regions exhibiting local conformational dynamics that were predicted by the experimental data, while, in general, not conferring unwarranted added mobility to other parts of the molecule, as noted previously (González et al., 1995). Finally, as recently demonstrated for an RNA hairpin (Schmitz et al., 1998), PDQPRO is an effective tool for reducing large MDtar trajectories into presentable ensembles which retain the dynamic properties of the entire run.

Concerning the effects of the α-nucleotide and polarity reversals, this technique reveals that the 3′-3′ linkage is relatively rigid in terms of both backbone torsion angles as well as the deoxyribose moieties of the $\alpha - 1$ and α-nucleotides, which are highly S in both molecules. Conversely, the backbone torsion angles across the 5′-5′ linkage are quite fluxional and significant N $\leftrightarrow$ S sugar repuckering occurs in the $\alpha + 1$ residue. The global helical properties of both molecules are generally preserved in spite of the modifications, with the DNA duplex quite B-like and the hybrid somewhat intermediate between A- and B-motifs. In the hybrid, the minor groove upstream of the 3′-3′ linkage is on average reduced, consistent with our final *in vacuo* structure (Aramini and Germann, 1999). In terms of the physical properties of these molecules and ODNs featuring such modifications, we feel that the rigid 3′-3′ linkage is a point that could be altered to recover the small but real thermodynamic penalty resulting from the insertion of these modifications. In addition, while there is controversy about the importance of minor groove width in RNase H activity (Fedoroff et al., 1993; Szyperski et al., 1999) we have observed preferential cleavage of RNA strands opposite the 3′-3′ linkage (manuscript in preparation), evidently in a region possessing a narrower minor groove.

In spite of the results presented here, the MDtar approach is not without its shortcomings. For instance, the time course of the trajectory is essentially compressed, the snapshots from the trajectory will invariably have a range of conformational energies, certain restraint types (i.e., *J*) can be problematic when applied in a time-averaged fashion (*vide supra*), and the technique as a whole is a computationally intensive, lengthy process. Hence, it would be interesting to compare our MDtar findings for these systems to other approaches; however, to date attempts at using the novel multiple-copy refinement strategy based on dipolar relaxation rates (Görler et al., 2000) have been unsuccessful in our hands due to the unnatural linkages. Moreover, we feel it appropriate to propose some criteria for assessing when the use of techniques such as MDtar and multiple-copy refinement is warranted. The most fundamental requirement, of course, is the presence of multiple conformations. Evidence for multiple conformers may come from the existence of NOE cross peaks between residues that are obviously mutually exclusive, and/or (deoxy)ribose *J*-coupling analysis. In general, the first case represents larger scale conformational changes that can be correlated with locally higher R^x values (per residue). In the second case the situation is more subtle and does not necessarily lead to markedly different local R^x values. For example, examination of the R^x values for the *in vacuo* DNA•DNA structure did not reveal any obvious hotspot although the local R^x for C8/C18 was somewhat higher, and the ensemble R^x values do not change markedly between the rMD and MDtar calculations for both molecules investigated here. It is important to note that the dynamics of the system is encoded in the NOE data even though this is not readily evident from the local R^x data. The use of J_{rms} as an independent monitor provides evidence that in the MDtar trajectories the conformational averaging was largely confined to those residues (sugars), which based on their coupling constants were *a priori* thought to be dynamic. We emphasize that the dynamic information is inherent in our NOE data, since no *J*-based restraints were applied in our MDtar calculations. Finally, while PDQPRO was used to winnow the MDtar ensembles exclusively on the basis of relaxation rates, an extension of this work is

to incorporate an additional source of experimental information, namely J_{rms}, into the selection process.

Acknowledgements

We thank Eric Pettersen (UCSF Computer Graphics Lab) for assistance in preparing the ribbon figures presented here. This work was supported by a grant from the National Institutes of Health GM OD55404.

References

Altona, C., Francke, R., de Haan, R., Ippel, J.H., Daalmans, G.J., Westra Hoekzema, A.J.A. and van Wijk, J. (1994) *Magn. Reson. Chem.*, **32**, 670–678.

Aramini, J.M. and Germann, M.W. (1999) *Biochemistry*, **38**, 15448–15458.

Aramini, J.M., Kalisch, B.W., Pon, R.T., van de Sande, J.H. and Germann, M.W. (1996) *Biochemistry*, **35**, 9355–9365.

Aramini, J.M., Mujeeb, A. and Germann, M.W. (1998a) *Nucleic Acids Res.*, **26**, 5644–5654.

Aramini, J.M., van de Sande, J.H. and Germann, M.W. (1998b) In *ACS Symposium Series 682: Molecular Modeling of Nucleic Acids* (Eds., Leontis, N.B. and SantaLucia J., Jr.), American Chemical Society, Washington, DC, pp. 92–105.

Bachelin, M., Hessler, G., Kurz, G., Hacia, J.G., Dervan, P.B. and Kessler, H. (1998) *Nat. Struct. Biol.*, **5**, 271–276.

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A. and Haak, J.R. (1984) *J. Chem. Phys.*, **81**, 3684–3690.

Bertrand, J.-R., Imbach, J.-L., Paoletti, C. and Malvy, C. (1989) *Biochem. Biophys. Res. Commun.*, **164**, 311–318.

Bhattacharyya, D. and Bansal, M. (1992) *J. Biomol. Struct. Dyn.*, **10**, 213–226.

Boiziau, C., Debart, F., Rayner, B., Imbach, J.-L. and Toulmé, J.-J. (1995) *FEBS Lett.*, **361**, 41–45.

Bonvin, A.M.J.J. and Brünger, A.T. (1995) *J. Mol. Biol.*, **250**, 80–93.

Borgias, B.A. and James, T.L. (1988) *J. Magn. Reson.*, **79**, 493–512.

Cheatham III, T.E. and Kollman, P.A. (1997) *J. Am. Chem. Soc.*, **119**, 4805–4825.

Darden, T.A., York, D.M. and Pedersen, L.G. (1993) *J. Chem. Phys.*, **98**, 10089–10092.

Debart, F., Tosquellas, G., Rayner, B. and Imbach, J.-L. (1994) *Bioorg. Med. Chem. Lett.*, **4**, 1041–1046.

Donders, L.A., de Leeuw, F.A.A.M. and Altona, C. (1989) *Magn. Reson. Chem.*, **27**, 556–563.

Fedoroff, O.Y., Salazar, M. and Reid, B.R. (1993) *J. Mol. Biol.*, **233**, 509–523.

Fennen, J., Torda, A.E. and van Gunsteren, W.F. (1995) *J. Biomol. NMR*, **6**, 163–170.

Ferrin, T.E., Huang, C.C., Jarvis, L.E. and Langridge, R. (1988) *J. Mol. Graphics*, **6**, 13–27.

González, C., Stec, W., Reynolds, M.A. and James, T.L. (1995) *Biochemistry*, **34**, 4969–4982.

Gorin, A.A., Zhurkin, V.B. and Olson, W.K. (1995) *J. Mol. Biol.*, **247**, 34–48.

Görler, A., Ulyanov, N.B. and James, T.L. (2000) *J. Biomol. NMR*, **16**, 147–164.

Güntert, P. (1998) *Q. Rev. Biophys.*, **31**, 145–237.

Gyi, J.I., Lane, A.N., Conn, G.L. and Brown, T. (1998) *Biochemistry*, **37**, 73–80.

Keepers, J.W. and James, T.L. (1984) *J. Magn. Reson.*, **57**, 404–426.

Kemmink, J. and Scheek, R.M. (1995) *J. Biomol. NMR*, **6**, 33–40.

Koga, M., Wilk, A., Moore, M.F., Scremin, C.L., Zhou, L. and Beaucage, S.L. (1995) *J. Org. Chem.*, **60**, 1520–1530.

Lavignon, M., Tounekti, N., Rayner, B., Imbach, J.-L., Keith, G., Paoletti, J. and Malvy, C. (1992) *Antisense Res. Dev.*, **2**, 315–324.

Liu, H., Spielmann, H.P., Ulyanov, N.B., Wemmer, D.E. and James, T.L. (1995) *J. Biomol. NMR*, **6**, 390–402.

Mujeeb, A., Kerwin, S.M., Kenyon, G.L. and James, T.L. (1993) *Biochemistry*, **32**, 13419–13431.

Nanzer, A.P., Torda, A.E., Bisang, C., Weber, C., Robinson, J.A. and van Gunsteren, W.F. (1997) *J. Mol. Biol.*, **267**, 1012–1025.

Pearlman, D.A. and Kollman, P.A. (1991) *J. Biol. Mol.*, **220**, 457–479.

Pearlman, D.A., Case, D.A., Caldwell, J.W., Ross, W.S., Cheatham III, T.E., Ferguson, D.M., Seibel, G.L., Singh, U.C., Weiner, P.K. and Kollman, P.A. (1995) *Amber 4.1*, University of California, San Francisco, CA.

Pikkemaat, J.A. and Altona, C. (1996) *Magn. Reson. Chem.*, **34**, S33–S39.

Ravishanker, G., Swaminathan, S., Beveridge, D.L., Lavery, R. and Sklenar, H. (1989) *J. Biomol. Struct. Dyn.*, **6**, 669–699.

Ryckaert, J.P., Ciccotti, G. and Berendsen, H.J.C. (1977) *J. Comput. Phys.*, **23**, 327–341.

Schmitz, U. and James, T.L. (1995) *Methods Enzymol.*, **261**, 3–44.

Schmitz, U., Ulyanov, N.B., Kumar, A. and James, T.L. (1993) *J. Biol. Mol.*, **234**, 373–389.

Schmitz, U., Donati, A., James, T.L., Ulyanov, N.B. and Yao, L. (1998) *Biopolymers*, **46**, 329–342.

Scott, W.R.P., Mark, A.E. and van Gunsteren, W.F. (1998) *J. Biomol. NMR*, **12**, 501–508.

Simmerling, C., Miller, J.L. and Kollman, P.A. (1998) *J. Am. Chem. Soc.*, **120**, 7149–7155.

Slamon, D.J., Clark, G.M., Wong, S.G., Levin, W.J., Ullrich, A. and McGuire, W.L. (1987) *Science*, **235**, 177–182.

Slamon, D.J., Godolphin, W., Jones, L.A., Holt, J.A., Wong, S.G., Keith, D.E., Levin, W.J., Stuart, S.G., Udove, J., Ullrich, A. and Press, M.F. (1989) *Science*, **244**, 707–712.

Szyperski, T., Götte, M., Billeter, M., Perola, E., Cellai, L., Heumann, H. and Wüthrich, K. (1999) *J. Biomol. NMR*, **13**, 343–355.

Tan, T.M.C., Kalisch, B.W., van de Sande, J.H., Ting, R.C.Y. and Tan, Y.H. (1998) *Antisense Nucleic Acid Drug Dev.*, **8**, 95–101.

Thomas, P.D., Basus, V.J. and James, T.L. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 1237–1241.

Torda, A.E., Scheek, R.M. and van Gunsteren, W.F. (1990) *J. Biol. Mol.*, **214**, 223–235.

Ulyanov, N.B. and Zhurkin, V.B. (1984) *J. Biomol. Struct. Dyn.*, **2**, 361–385.

Ulyanov, N.B. and James, T.L. (1995) *Methods Enzymol.*, **261**, 90–120.

Ulyanov, N.B., Schmitz, U., Kumar, A. and James, T.L. (1995) *Biophys. J.*, **68**, 13–24.

van Wijk, J., Huckriede, B.D., Ippel, J.H. and Altona, C. (1992) *Methods Enzymol.*, **211**, 286–306.

van Wijk, J., Haasnoot, K., de Leeuw, F., Huckriede, D. and Altona, C. (1995) *PSEUROT 6.2. A Program for the Conformational Analysis of Five Membered Rings*. University of Leiden, The Netherlands.

Vichier-Guerre, S., Pompon, A., Lefebvre, I. and Imbach, J.-L. (1994) *Antisense Res. Dev.*, **4**, 9–18.

Yao, L., James, T.L., Kealey, J.T., Santi, D.V. and Schmitz, U. (1997) *J. Biomol. NMR*, **9**, 229–244.

Appendix I. Calculation of theoretical coupling constants from models/pseudorotation theory

Theoretical ^{1}H-^{1}H J-coupling constants – $J_{1'2'}$, $J_{1'2''}$, $J_{2'3'}$, and $J_{2''3'}$ – were calculated from structural coordinates on the basis of pseudorotation theory in a two-step procedure. First, the endocyclic torsion angles ν_1 (O4′-C1′-C2′-C3′) and ν_2 (C1′-C2′-C3′-C4′) were converted to exocyclic torsion angles, Φ_{HH}, using Equation A-1.

$$\Phi_{HH} = B_i + A_i\nu_i \tag{A-1}$$

where the coefficients A and B are dependent upon the protons defining the torsion angle, as well as the type of five-membered ring (Table A-1) (van Wijk et al., 1992, 1995).

Table A-1. *A* and *B* coefficients for β- and α-D-deoxyribose ring systems

β-D-deoxyribose				α-D-deoxyribose			
Φ_{HH}	A	B	ν_i	Φ_{HH}	A	B	ν_i
$\Phi_{1'2'}$	1.030	121.4	ν_1	$\Phi_{1'2'}$	1.030	1.4	ν_1
$\Phi_{1'2''}$	1.020	0.9	ν_1	$\Phi_{1'2''}$	1.020	−119.1	ν_1
$\Phi_{2'3'}$	1.060	2.4	ν_2	$\Phi_{2'3'}$	1.060	2.4	ν_2
$\Phi_{2''3'}$	1.060	122.9	ν_2	$\Phi_{2''3'}$	1.060	122.9	ν_2

Second, a set of generalized Karplus equations (Equation A-2) relates the individual J-coupling constants and corresponding exocyclic torsion angles, taking into account the significant electronegativity effects of substituents (λ_i) and their orientations (ξ_i) with respect to a given pair of protons involved in the coupling (Table A-2) (van Wijk et al., 1992, 1995; Donders et al., 1989; Altona et al., 1994).

$$\begin{aligned} J_{HH} \;=\; & C_0 + C_1\cos\Phi_{HH} + C_2\cos 2\Phi_{HH} + C_3\cos 3\Phi_{HH} + S_2\sin 2\Phi_{HH} \\ & C_0 = 6.97 - 0.58\Sigma\lambda_i - 0.24(\lambda_1\lambda_2 + \lambda_3\lambda_4) \\ & C_1 = -1.06 \\ & C_2 = 6.55 - 0.82\Sigma\lambda_i + 0.20(\lambda_1\lambda_4 + \lambda_2\lambda_3) \\ & C_3 = -0.54 \\ & S_2 = 0.68\Sigma\xi_i\lambda_i^2 \end{aligned} \tag{A-2}$$

Table A-2. λ_i values for β- and α-D-deoxyribose ring systems[a]

β-D-deoxyribose					α-D-deoxyribose				
J_{HH}	λ_1	λ_2	λ_3	λ_4	J_{HH}	λ_1	λ_2	λ_3	λ_4
$J_{1'2'}$	0.56	1.26	0	0.62	$J_{1'2'}$	1.26	0.56	0	0.62
$J_{1'2''}$	0.56	1.26	0.62	0	$J_{1'2''}$	1.26	0.56	0.62	0
$J_{2'3'}$	0.62	0	1.26	0.62	$J_{2'3'}$	0.62	0	1.26	0.62
$J_{2''3'}$	0	0.62	1.26	0.62	$J_{2''3'}$	0	0.62	1.26	0.62

[a]In all cases the substituent orientation factors are: $\xi_1 = \xi_3 = +1$; $\xi_2 = \xi_4 = -1$.

Thus, inserting the appropriate λ_i and ξ_i values into Equation A-2 results in the following Karplus relations for $J_{1'2'}$, $J_{1'2''}$, $J_{2'3'}$, and $J_{2''3'}$ in a β-D-deoxyribose ring (A-3):

$$J_{1'2'} = 5.385 - 1.06\cos\Phi_{1'2'} + 4.619\cos 2\Phi_{1'2'} - 0.54\cos 3\Phi_{1'2'} - 1.128\sin 2\Phi_{1'2'}, \tag{A-3a}$$

$$J_{1'2''} = 5.385 - 1.06\cos\Phi_{1'2''} + 4.705\cos 2\Phi_{1'2''} - 0.54\cos 3\Phi_{1'2''} - 0.605\sin 2\Phi_{1'2''}, \tag{A-3b}$$

$$J_{2'3'} = 5.333 - 1.06\cos\Phi_{2'3'} + 4.577\cos 2\Phi_{2'3'} - 0.54\cos 3\Phi_{2'3'} + 1.080\sin 2\Phi_{2'3'}, \tag{A-3c}$$

$$J_{2''3'} = 5.333 - 1.06\cos\Phi_{2''3'} + 4.656\cos 2\Phi_{2''3'} - 0.54\cos 3\Phi_{2''3'} + 0.557\sin 2\Phi_{2''3'}. \tag{A-3d}$$

In the case of dynamics runs, individual J-couplings for each deoxyribose residue were computed in this way at each point in the trajectory, and subsequently averaged prior to calculating J_{rms} and ΔJ_{av}.