Dynamic NMR Structures of [Rp]- and [Sp]-Phosphorothioated DNA-RNA Hybrids: Is Flexibility Required for RNase H Recognition?

Marco Tonelli,* Nikolai B. Ulyanov,* Todd M. Billeci,* Boleslaw Karwowski,† Piotr Guga,† Wojciech J. Stec,† and Thomas L. James*

*Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, California 94143-2280 USA; and †Department of Bioorganic Chemistry, Centre of Molecular and Macromolecular Studies, Polish Academy of Sciences, 90-363 Lód'z, Sienkiewicza 112, Poland

ABSTRACT Chemically modified DNA oligonucleotides have been crucial to the development of antisense therapeutics. High-resolution structural studies of pharmaceutically relevant derivatives have been limited to only a few molecules. We have used NMR to elucidate the structure in solution of two DNA-RNA hybrids with the sequence d(CCTATAATCC)·r(GGAUUAUAGG). The two hybrids contain an unmodified RNA target strand, whereas the DNA strand contains one of two different stereoregular sugar-phosphate backbone linkages at each nucleotide: 1), [Rp]-phosphorothioate or 2), [Sp]-phosphorothioate. Homonuclear two-dimensional spectroscopy afforded nearly complete nonlabile proton assignments. Distance bounds, calculated from the nuclear Overhauser effect (NOE) crosspeak intensities via a complete relaxation matrix approach with the program MARDIGRAS, were used to restrain the structure of the two hybrids during simulations of molecular dynamics. Analysis of restrained molecular dynamics trajectories suggests that both hybrids are flexible, requiring the use of molecular dynamics with time-averaged restraints (MDtar) to generate ensembles of structures capable of satisfying the NMR data. In particular, the deoxyribose sugars of the DNA strand show strong evidence of repuckering. Furthermore, deoxyribose sugar repuckering is accompanied by increased flexibility of overall helical geometry. These observations, together with the analysis of the crystal structure of a hybrid duplex in complex with ribonuclease H (RNase H), suggested that this flexibility may be required for recognition by RNase H.

INTRODUCTION

Antisense oligonucleotide inhibitors are designed to disrupt function of specific RNA targets by forming hybrid duplexes with complementary RNA sequences. Mode of action of the first generation antisense oligonucleotides utilized ribonuclease H (RNase H), a ubiquitous sequence-nonspecific endonuclease that cleaves phosphodiester bonds in the RNA strand of RNA/DNA duplexes but leaves the DNA strand intact. Such a convenient property of RNase H allows recycling the oligodeoxyribonucleotides after the RNA strand is hydrolyzed. Antisense oligonucleotides often include phosphorothioate modifications of the backbone, which confer resistance to single-strand-specific nucleases. At the same time, RNA hybrids with phosphorothioated oligodeoxyribonucleotides (PSO) are still recognized by RNase H (Agrawal and Iyer, 1997; Crooke and Bennett, 1996; DeLong et al., 1997; Temsamani and Guinot, 1997). One serious problem with this strategy is a relatively low specificity of both DNA-RNA recognition and recognition of hybrid duplexes by RNase H. Indeed, it has been shown that mismatched hybrid duplexes are still recognized by RNase H, which leads to the cleavage of unintended RNA sequences. Antisense oligonucleotides of newer generations are designed to form very stable hybrid duplexes with complementary RNA due to various base, sugar, or backbone modifications. This comes most often, if not always, at the expense of the hybrid ceasing to be a substrate for the RNase H. Consequently, the mode of action of new antisense drugs is steric block, when formation of very stable hybrid duplexes interferes with normal function of the target RNA (Altmann et al., 1997; Freier and Altmann, 1997; Marquis and Grindel, 2000; Monia, 1997). Nevertheless, recycling of antisense oligonucleotides is a property of the RNase H mode of action too attractive to lose; a promising approach includes ligation of new-generation oligonucleotides with DNA or PSO, thus combining the stability of the hybrid duplex with a window for the RNase H action (Lima and Crooke, 1997; McKay et al., 1999).

Many details of the enzymatic catalysis by RNase H have been elucidated in the past several years (as recently reviewed in Krakowiak et al., 2002). Elucidation of some structural aspects has emanated from the crystal structure of a complex between HIV-1 reverse transcriptase (which contains an RNase H domain) and an RNA-DNA hybrid duplex (Sarafianos et al., 2001). However, many questions remain unanswered, one of the most puzzling being why RNase H cleaves the RNA strand in RNA-DNA or RNA-PSO but not in RNA-RNA duplexes and not in hybrid duplexes of RNA with new antisense drugs. In an attempt to answer these questions, we have been studying a series of complexes of antisense oligonucleotides with a prototype RNA target, which includes the Pribnow box sequence. Previously, we described high-resolution NMR studies in

Submitted May 5, 2003, and accepted for publication July 1, 2003.

Address reprint requests to T. L. James, Dept. of Pharmaceutical Chemistry, University of California, San Francisco, 600 16th St., San Francisco, CA 94143-2280. Tel.: 415-476-1916; Fax: 415-502-8298; E-mail: james@picasso.ucsf.edu.

Marco Tonelli's present address is Dept. of Biochemistry, University of Wisconsin, Madison, WI 53706.

0006-3495/03/10/2525/14 $2.00

solution of hybrid duplexes of RNA with complementary DNA, [Rp]-phosphorothioate, and phosphorodithioate (Gonzàlez et al., 1994, 1995; Furrer et al., 1999). Here we present NMR data for the [Sp]-PSO-RNA hybrid with the same sequence and determine solution structures for both [Sp]-PSO-RNA and [Rp]-PSO-RNA hybrids (Fig. 1). The main conformational feature of these structures, common for both hybrids, is unusually high flexibility of deoxyribose rings in PSO strands, compared with duplex DNA or duplex RNA with the same sequence. A similar behavior was observed before for the DNA strand in a DNA-RNA hybrid (Gonzàlez et al., 1995; Gyi et al., 1998). Such flexibility required a special approach for structure determination, because no single structure can adequately describe the observed NMR data. Here we used a combination of molecular dynamics with time-averaged restraints (MDtar) (Torda et al., 1990) and PDQPRO (Ulyanov et al., 1995) methods to determine an ensemble of conformations consistent with experimental data. Flexibility of deoxyribose sugar rings is accompanied by increased flexibility of overall helical geometry, including the inclination and X-displacement parameters. Comparing our results with the crystal structure of a complex between RNase H and a hybrid duplex, we proffer the hypothesis that this flexibility may be required for the RNase H recognition.

MATERIAL AND METHODS

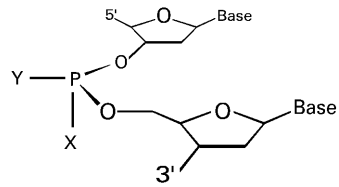
Solid-phase synthesis of PS-oligos

The synthesis of oxathiaphospholane monomers and their chromatographic separation into P-diastereomerically pure species has been published elsewhere (Stec et al., 1998). The monomers were dried at high vacuum (<0.01 mm Hg) overnight before synthesis of oligomers.

The manual syntheses were performed on 2-μmol scale using a solid support (CPG with sarcosinyl linker) loaded with 35 μmol nucleoside per gram. Standard solutions of dichloroacetic acid (DCA) in methylene chloride and 4-(dimethylamino)pyridine (DMAP)/acetic anhydride (Ac_2O)/lutidine in tetrahydrofuran (THF) were used for detritylation and capping steps, respectively. Gas-tight syringes were used for delivery of all reagents. The solvents were stored under an atmosphere of dry argon, in vials (or bottles) capped with rubber septa. An appropriate monomer (20 mg) was dissolved in dry acetonitrile (450 μL) just before each condensation step. To that solution, 45 μL of a solution of 5.8 M 1,4-diazabicyclo[5.4.0]undec-7-ene (DBU) in acetonitrile was added, yielding 0.53 M final concentration of the activator. The mixture was applied to the column, followed by intensive swirling over the whole coupling time, ~15 min. The reagents were expelled and the support was washed with dry methylene chloride (5 mL) and dry acetonitrile (7 mL). Other steps, such as detritylation (5 ml) and capping, were followed by exhaustive washing with 5 mL dry acetonitrile and drying in vacuum. Dimethoxytrityl cation (DMT^+) assay showed a repetitive yield of ~93%.

Purification and characterization of PS-oligonucleotides

When the synthesis was complete, the oligomer (bearing 5′-*O*-DMT groups) was cleaved from the support under standard conditions (25% NH_4OH, 2 h), and the protecting groups from nucleobases were removed at 55°C over 12 h.



	X	Y
phosphate	O	O
[Rp]-phosphorothioate	O	S
[Sp]-phosphorothioate	S	O
phosphorodithioate	S	S

FIGURE 1 Scheme showing definition of [Rp] and [Sp] oxygens.

The sample was concentrated under reduced pressure in a Speed-Vac concentrator (Thermo Savant, Holbrook NY), and two-step RP-HPLC (DMT-on and DMT-off) was used to isolate the product (a column ODS Hypersil, RP-C_{18}, 250 × 4.6 mm, buffer A: 0.1 M triethylammonium bicarbonate (TEAB), buffer B: 40% acetonitrile in 0.1 M triethylammonium bicarbonate, flow rate 1 mL/min). Retention times of ~31.3 min (gradient 0–100% B over 30 min) were recorded for the DMT-on analysis. After detritylation with 50% acetic acid and evaporation to dryness, the product was isolated by HPLC (retention time ~24.9 min, gradient 0–70% B over 25 min). The appropriate fractions were collected and concentrated under reduced pressure in a Speed-Vac concentrator. The samples collected from consecutive syntheses were dissolved in water and placed on a Dowex-25 (Na^+ form) column. The product was analyzed by ^{31}P NMR spectra acquired on a Bruker AC-200 instrument (200 MHz; 85% H_3PO_4 as the external standard for ^{31}P). The spectrum contained six resonance signals at δ = 56.58, 56.55, 56.41, 56.37, 56.25, and 55.99 ppm (D_2O), with the correct ratio of integrals (Fig. 2). The spectrum revealed also the presence of resonances of phosphate groups at δ ranging from 0.2 to −0.5 ppm of total intensity 2.4%. The PAGE analysis of the sample labeled with ^{32}P-phosphate group at the 5′-end of the oligonucleotide confirmed close to 100% purity of the product.

The FAB-MS spectra (13 keV, CS^+) were recorded on a Finnigan MAT 95 spectrometer, and negative ion MALDI mass spectra were recorded on a Voyager-Elite instrument (PerSeptive Biosystems, Framingham, MA). Ultraviolet (UV) spectra were recorded on a GBC 916 UV/VIS spectrophotometer. Densitometry of autoradiograms was performed on an LKB Ultroscan XL densitometer. HPLC analyses were done on a ThermoSeparation system (a pump SpectraSystem P4000, a UV detector SpectroMonitor SM5000). Acetonitrile and DBU were supplied by Merck (Darmstadt,

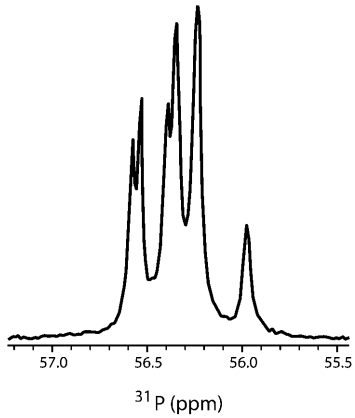


FIGURE 2 1D ^{31}P NMR spectrum of the [Sp]-phosphorothioated deoxyribonucleotide strand after purification.

Germany). Acetonitrile to be used as solvent for DBU and oxathiaphospholane monomers was dried over P_2O_5 (5 g/L) and distilled through a 20-cm Vigreux column in the atmosphere of dry argon.

Sample preparation and NMR spectroscopy

The hybrid duplex was prepared for NMR by titration of a solution of the modified [Sp]-phosphorothioate DNA strand with a similar solution of the RNA strand. The 1:1 stoichiometry was ensured by monitoring the disappearance of the DNA single strand peaks in a 1D NMR spectrum. After each addition of RNA strand, the sample was reannealed directly in the NMR tube to ensure proper forming of the duplex. The final concentration of hybrid in the NMR sample is ~1 mM dissolved in 0.1 mM EDTA, 100 mM NaCl, and 20 mM sodium phosphate (pH 6.5) buffer. The sample was then lyophilized twice to exchange water for ^2H_2O and finally placed in a 5-mm Shigemi tube.

Two 2D NOESY, a 2D TOCSY and a 2D DQF-COSY, spectra were collected for the [Sp]-phosphorothioate hybrid sample in ^2H_2O at 600 MHz on a Varian INOVA spectrometer (Varian, Palo Alto, CA). All measurements were taken at 293 K using a spectral width of 6600 Hz in both dimensions and 4096 × 2048 data points along ω_2 and ω_1 dimensions, respectively. These spectra were then processed using the program NMRPipe (Delaglio et al., 1995) to yield a final 2 K × 2 K data set; a combination of a Gaussian and sine-bell window functions was applied for resolution enhancement in each dimension. Subsequently, the processed spectra were imported into our locally developed NMR data analysis program SPARKY (Goddard and Kneller, 1998) for assignment and integration. The two 2D NOESY spectra were recorded in hypercomplex mode with mixing times of 100 and 200 ms to allow for buildup of NOE crosspeaks. These spectra were acquired with 32 scans for each t_1 fid and a delay of 8 s between scans. The pure absorption 2D DQF-COSY was also acquired with 32 scans for each t_1 fid, but a repetition delay of 3 s between scans. A somewhat stronger apodization function was applied during processing to reduce cancellation of the antiphase components of the DQF-COSY crosspeaks. Finally, the 2D TOCSY spectrum was collected using a mixing period of 50 ms, and only eight scans for each t_1 fid were sufficient. Two-dimensional NOESY spectra were also acquired with the sample dissolved in H_2O, but they were only used for the assignments of exchangeable resonances.

Assignments

The 2D NOESY spectra acquired in ^2H_2O were used for assigning nonexchangeable protons by identifying base-base, base-H1′, base-H2′/H2′ (for the DNA strand), and base-H3′ sequential walks (Broido et al., 1984; Wüthrich, 1986). Intraresidue connectivities provided by 2D TOCSY and COSY spectra were helpful to distinguish among overlapped resonances. Proton assignments for the [Sp]-hybrid are reported in Table S1 of the Supplementary Material.

Extracting interproton distance restraints

In 2D NOESY spectra, the crosspeaks intensities were measured with SPARKY by line fitting of the peaks to a Gaussian function and integration of the theoretical curve. These intensities were then used to calculate accurate distance restraints by using the complete relaxation matrix approach implemented in MARDIGRAS (Borgias and James, 1990). Two of the three starting models that were used for MARDIGRAS calculations were generated by the miniCarlo program (Zhurkin et al., 1991) using standard helical parameters for A-form and B-form nucleic acid duplexes. The third starting model, H-form, was also generated by miniCarlo; it has C3′-endo sugar puckers for the RNA strand and C2′-endo for the PSO strand. All three structures were energy minimized by the same program. Several values of correlation time were used to represent the isotropic motion of the molecule as a whole but, in the end, only distances calculated using 5, 6, and 7 ns were used to estimate the bounds for structure calculation. These values of τ_c were chosen by analyzing their ability to accurately reproduce certain intrasugar distances that are known to change only within a narrow range, independently of the sugar conformation. MARDIGRAS was run with the RANDMARDI option set to 200 (Liu et al., 1995). Thus, for each of the two sets' experimental NOE intensities, 200 new sets were generated by adding noise, randomly calculated within user-chosen limits, directly to the intensities before calculation of the distances. In conclusion, three starting models and three values of isotropic τ_c were used to run MARDIGRAS on the two experimental NOE intensity sets measured at 100- and 200-ms mixing times, for a total of 18 MARDIGRAS calculations. Each one of these calculations was then repeated 200 times on randomly modified intensity sets generated by the RANDMARDI option selected in MARDIGRAS. All these distances, generated by MARDIGRAS, were combined to calculate average values and standard deviations. Two sets of upper and lower bounds were calculated for each proton pair. One was calculated as average MARDIGRAS values plus or minus standard deviation. The other set of bounds was generated by building a distribution of the distances and eliminating 10% from each end of the distribution. The wider of the two constituted the final distance bounds used for structure calculation.

Sample preparation and NMR spectroscopy for the [Rp]- hybrid were performed following a strategy similar to that described here for the [Sp]-hybrid and have already been reported elsewhere (Furrer et al., 1999). However, even though the distances for the [Rp]-hybrid were calculated previously, we repeated the procedure for both hybrids, to use compatible computational schemes for both molecules. The main difference between the two systems was in the concentration of hybrid in the NMR sample, being higher for the [Rp]-hybrid. However, given that all spectra for the [Rp]-hybrid were also collected at a higher temperature (30°C vs. 20°C), the values of τ_c used in MARDIGRAS calculations were the same for both hybrids (5, 6, and 7 ns); see Table 1 for a summary of MARDIGRAS-calculated distance restraints for both [Rp]- and [Sp]-hybrids.

Finally, together with the NOE-derived distance restraints, some empirical restraints were also included for structure refinement. These include typical hydrogen bond distance and flat angle restraints to ensure proper pairing of the bases. Loose distance restraints (3.0-Å–8.0-Å upper and lower bounds) were imposed between each sodium counterion and two phosphorous atoms to maintain a sodium atom in the space between two phosphate groups. These restraints account for the position of 18 ions, 9 for each strand. The remaining two ions, required to neutralize the system, were let free to move in the bulk solvent. Furthermore, later in our attempt to refine the hybrids, non-NOE cross-strand distance restraints (4.5-Å lower bound) were added between H1′ protons: H1′(i)-H1′(j + 2), where (j) is the nucleotide basepaired to (i), yielding a total of eight distance restraints.

TABLE 1 Number and average width of NOE-derived distance bounds for the PSO hybrids

	[Sp]-hybrid		[Rp]-hybrid	
	Number of distances	Average width, Å	Number of distances	Average width, Å
Total	208	1.51	196	1.66
Intraresidue	77	1.29	65	1.74
Interresidue	118	1.66	117	1.70
Cross-strand	13	1.36	14	0.98
DNA strand	118	1.53	100	2.17
Intraresidue	50	1.39	39	2.17
Interresidue	68	1.64	61	2.17
RNA strand	77	1.49	82	1.16
Intraresidue	27	1.12	26	1.10
Interresidue	50	1.70	56	1.18

Distance bounds calculated based on NOE intensities as described in text.

These bounds were needed to keep the two strands from coming too close in the minor groove of the duplex during the MD simulations; use of these non-NOE restraints is justified by the absence of the corresponding crosspeaks in NOESY spectra. A summary of experimental and empirical restraints used in structure refinement is reported in Table 2.

Structure calculations

Upper and lower bounds generated by MARDIGRAS were used to refine the structure of both hybrids by restrained molecular dynamics simulations (rMD) or by MD with time averaging. The program AMBER 6 was used (Case et al., 2000).

The initial models used at the start of AMBER calculations for the two PSO hybrids were generated with miniCarlo by restrained minimization of a typical A-form hybrid duplex. These initial structures were further energy minimized with AMBER in vacuo and, then, placed in periodic solvent boxes containing 10 Å of TIP3P water molecules surrounding the hybrid in each direction. Subsequently, the system was neutralized by adding sodium counterions. The solvated systems were then equilibrated with standard protocols. Briefly, after an initial minimization, 50 ps of MD simulations were run at constant volume, followed by 150 ps of MD simulations at constant pressure. The water molecules were liberated to equilibrate from the beginning of the simulation, whereas the hybrid duplex was kept fixed by positional restraints that were gradually reduced toward the end of the equilibration. The sodium counter ions were also initially fixed by positional restraints, but let free to equilibrate after 40 ps of MD simulations at constant pressure, their position being then restrained in the neighborhood of phosphate groups by the empirical bounds described above. In fact, the empirical restraints, as well as the NOE-derived distance bounds, were gradually built up during the equilibration period.

Production runs for regular restrained MD and MDtar calculations were started independently from the equilibrated system. Each production calculation was run for a period of 1 ns. Throughout the equilibration and production runs, the particle mesh Ewald (PME) method (Essmann et al., 1995) was used to treat electrostatic interactions (default in AMBER 6 for solvated systems), and the temperature was kept constant at 300 K by using the Berendsen coupling algorithm (Berendsen et al., 1984) with a time constant of 1 ps. Other parameters were set to typical values for MD: a 9-Å cutoff for the nonbonded interactions, SHAKE option on bonds involving hydrogen atoms (Ryckaert et al., 1977), and a 1-fs time step. The atomic coordinates were saved every 1 ps (1000 structures for each rMD or MDtar run). All trajectories were stripped of the water molecules and analyzed by visual inspection using the program MOIL-View (Simmerling et al., 1995). In the end, four simulations were selected for further analysis: two rMD and two MDtar runs, one of each for the two [Sp]- and [Rp]-phosphorothioate hybrids. A total of 1000 pdb files were generated for each trajectory, and the program Curves (Lavery and Sklenar, 1996) was run on each saved duplex structure to calculate the helical parameters. The programs MIDAS (Ferrin et al., 1988), Chimera (Huang et al., 1996), and Molmol (Koradi et al., 1996) were used to visualize and further analyze the coordinates. Representative structural ensembles for each 1-ns production run were generated using the locally developed PDQPRO algorithm (Ulyanov et al., 1995). CORMA (Keepers and James, 1984) was then run for each simulation on the whole ensemble of 1000 structures (with equal probability for each structure), on the ensemble of PDQPRO selected structures (using PDQPRO probabilities for each model), and separately on each pdb file, to back calculate theoretical NOESY spectra. R factors were measured by comparison of the simulated NOE intensities with the corresponding experimental ones.

The complete list of experimental and empirical restraints used for structure refinement and the force constants used to enforce them is reported in Table 2. During MDtar runs, only the NOE-derived distance bounds were subjected to the exponentially weighted third-root time averaging with an exponential time constant of 20 ps and the ''pseudo-force'' option (Torda et al., 1990; Pearlman and Kollman, 1991); all other restraints were applied continuously.

TABLE 2 Types of restraints used in MD simulations of the [Rp]- and [Sp]-hybrids

	[Sp]-hybrid		[Rp]-hybrid	
	Number of bounds	k*	Number of bounds	k*
NOE distance restraints	208	2.5	196	5.0
H-bond distance restraints	24	3.75	24	7.5
H-bond flat angle restraints	24	5.0	24	10.0
Na^+-P distance restraints†	32	0.5	32	1.0
H1′(i)-H1′(j + 2) restraints‡	8	25.0	8	50.0

*Force constant, kcal/(mol × Å^2).
†Distance bounds of 3.0–8.0 Å (see text).
‡Distance bounds of 4.5–40.0 Å (see text).

PDQPRO calculations

Representative structural ensembles were selected from the entire 1-ns MDtar trajectories using the PDQPRO algorithm (Ulyanov et al., 1995). The program finds a subset of structures and their probabilities so that the resulting ensemble produces the best fit of observed NMR parameters. Experimental homonuclear cross-relaxation rates were derived from the NOE data using MARDIGRAS simultaneously with calculating distance restraints (see above).

RESULTS

NMR spectroscopy and interproton distance restraints

Proton assignments were performed following the procedure described earlier. We were able to identify all base protons and most sugar protons, except for H4′ of residues A13, U17, G19, and G20. Some H5′/H5″ protons were also tentatively assigned based on their NOE and TOCSY contacts, but not used for structure refinement. Not surprisingly, resonances for the RNA strand are similar to those already assigned for the other hybrid duplexes with the same sequence (Furrer et al., 1999), although resonances for the modified DNA duplex strand are somewhat different. DNA and RNA strand sequential walks for the fingerprint region are shown in Fig. 4 *A*, and the proton assignments for the [Sp]-hybrid are given in Table S1. Interproton distance restraints were calculated from experimental NOE data as described in Materials and Methods, using a complete relaxation matrix analysis (MARDIGRAS) with the random error perturbation procedure (RANDMARDI). Even though the distances were calculated for the [Rp]-hybrid previously (Furrer et al., 1999), we repeated the procedure for both hybrids, to use compatible computational schemes for both molecules. The statistics of distance restraints is similar for both hybrids, with ~10 restraints per residue and an average flat-well width of 1.7 Å and 1.5 Å for the [Rp]- and [Sp]-hybrids, respectively (Table 1). These numbers do not include interproton distances with low variation, such as

H1′-H2′, which do not carry any structural information. The distribution of the calculated distance restraints is shown in Fig. 3. Distance bounds for the RNA residues are largely consistent with the helical A conformation and with C3′-endo sugar pucker of riboses. In contrast, distances involving deoxyribose protons in [Sp]- or [Rp]-PSO strands are not consistent with either C3′-endo or C2′-endo sugar conformations. Intraresidue H2′-H6/H8 distances are consistent with the B-like sugar conformations (C2′-endo or C1′-exo), but they are too short for the C3′-endo conformations (Fig. S1 in Supplementary Material). At the same time, intraresidue H3′-H6/H8 distances are consistent with the A-like C3′-endo conformations, but they tend to be too short for B-like conformations. Similar sugar flexibility has been observed previously for the DNA residues in regular DNA-RNA hybrids (Gonzàlez et al., 1994; Gyi et al., 1998) and, to a lesser extent, for DNA duplexes (Tonelli and James, 1998; Schmitz et al., 1990, 1992b; Weisz et al., 1992; Mujeeb et al., 1992; Rinkel and Altona, 1987). This is qualitatively confirmed by the DQF-COSY spectrum of the [Sp]-hybrid (Fig. 4 *B*): the presence of H2′-H1′ peaks rules out pure C3′-endo sugar pucker for the PSO residues. In principle, this situation may arise either due to some nonstandard conformations of sugar rings in PSO residues, or by flexibility of these residues when both C2′-endo and C3′-endo conformers are present.

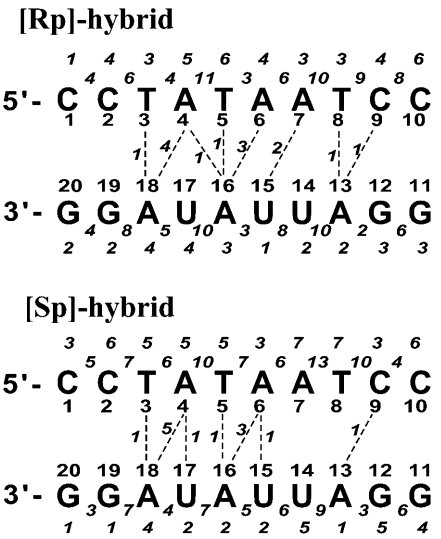


FIGURE 3 Sequence numbering (*bold*) and distance restraint distributions (intraresidue, sequential, and cross-strand) for the [Rp]- and [Sp]-hybrids. Number of distance restraints in each category is shown in italic.

Structure calculation for the [Sp]-hybrid

Initially, we attempted to refine the [Sp]-hybrid by rMD in vacuo using a distance-dependent dielectric constant and big hydrated sodium ions to neutralize the charges in the

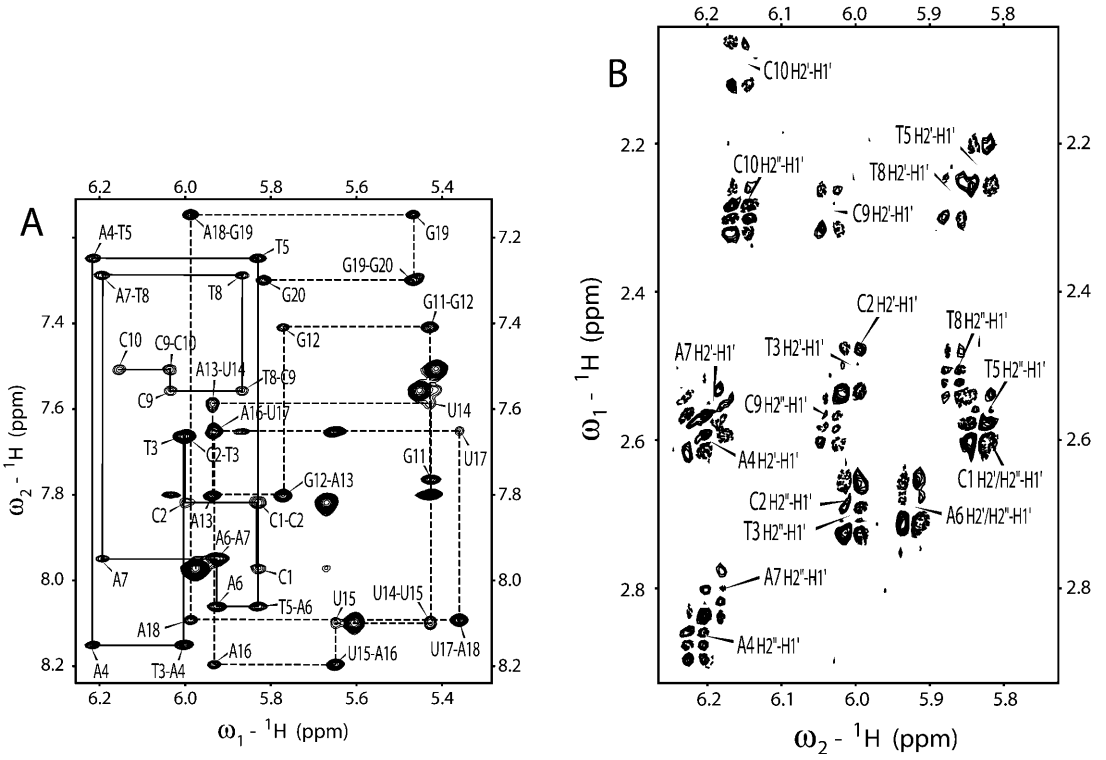


FIGURE 4 (*A*) Fingerprint portion of the 200-ms NOESY spectrum for the [Sp]-hybrid showing sequential base-to-H1′ walks for the PSO (*solid line*) and RNA (*dashed line*) strands. Only the crosspeaks along the sequential walks are labeled. (*B*) Portion of the DQF-COSY spectrum for the [Sp]-hybrid showing H1′-H2′ and H1′-H2″ crosspeaks in the PSO strand.

phosphate groups. The structure appears to be stable at the beginning of the simulation, when no experimental restraints are enforced. However, as soon as we start ramping up the force constant for the NOE restraints, the sodium ions fly away and the duplex conformation rapidly deteriorates and keeps fluctuating among bad geometries, indicating that, at these conditions, the restraints we apply are not able to define any reasonable structure.

rMD refinement of fully solvated [Sp]-hybrid with particle mesh Ewald model for calculation of electrostatic interactions failed to converge; average atomic RMSD deviation for snapshots taken each picosecond is 2.4 Å (Table 4). The structure continues drifting during 1 ns of simulation toward conformations with negative inclination (as low as −40.0 for T5-A16, A6-U15, and A7-U14 basepairs), positive propeller twist (with peaks higher than 50.0 for A6-U15 and 40.0 for T5-A16) and zero X-displacement (Fig. 5). In the course of the trajectory, the same basepairs are weakened, despite the use of hydrogen bond restraints. This is accompanied by sugar repuckering in some residues (e.g., *T3*, *black solid lines* in the *left panels* of Fig. 6 *A*). At the end of the 1-ns simulation, the duplex appears elongated with very narrow minor groove and short cross-strand contacts (e.g., H1′(*i*)-H1′($j + 2$) close to 3 Å, where j is the nucleotide basepaired to i). Surprisingly, this conformational change is accompanied by an increase in the distance restraint violation and, hence, of the constraint energy (from 40 to 80 kcal/mol Å). All our attempts to stabilize the rMD simulations by changing the conditions, such as force constant for the NMR restraints, adding and modifying empirical restraints, have only been partially successful in slowing down the drifting of the molecule. In the end, the molecule is irreversibly driven toward structures with bad geometry. The average distance restraint deviation $\langle Rdev \rangle$ is quite high for these structures (0.14 Å). This is clearly a failed trajectory; we show some data for this trajectory only for the sake of comparison (Table 3).

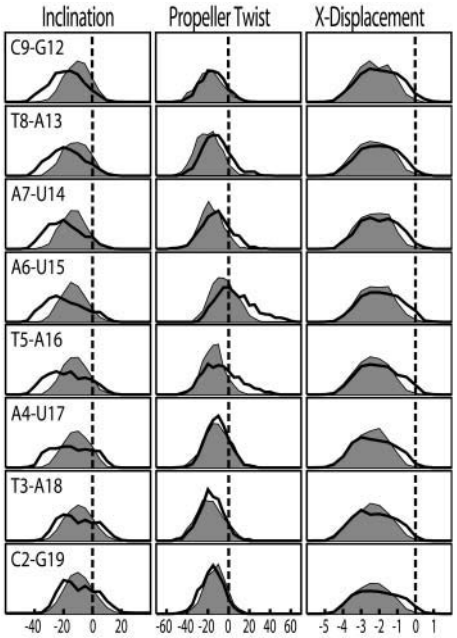


FIGURE 5 Distribution of Inclination, Propeller Twist, and X-displacement parameters for internal basepairs of the [Sp]-hybrid calculated from MDtar (*gray histograms*) and rMD (*black solid lines*) ensembles using CURVES.

The MDtar trajectory of the fully solvated [Sp]-hybrid was run with PME electrostatics as described in Materials and Methods, starting with an independent equilibration process. In contrast to rMD calculations, basepairs stayed intact during the course of the trajectory. Unlike rMD calculations, it is not required in MDtar simulations that all restraints are satisfied simultaneously for each snapshot. Instead, the distances are third-root-averaged, with more recent snapshots weighted exponentially more heavily. Effectively, it is required that restraints are satisfied for the distances averaged during each fragment of trajectory equal in length to the exponential time constant for distance averaging (20 ps for current simulations). Correspondingly, to assess the quality of such a trajectory, the third-root-averaged distances must be compared with the experimental restraints. Indeed, the corresponding distance deviation index, $dev\langle r^{-3} \rangle$ is very low for the trajectory (0.08 Å), whereas the distance deviations are high for individual structures, with the average $\langle Rdev \rangle$ of 0.16 Å (Table 3). Another index assessing the quality of the refinement is sixth-root-weighted R^x-factor, which directly compares calculated and experimentally measured NOE intensities. R^x-factors were calculated using CORMA for the two NOESY data sets, and their average values and standard deviations are reported in Table 3. These R^x-factors characterize how well each individual structure fits experimental data. In addition, ensemble R^x-factors (Schmitz et al., 1992a) (*ens*R^x in Table 3) were calculated for each ensemble of structures, assuming fast exchange between individual conformers. Ensemble R^x-factor assesses quality of an ensemble as a whole, rather than each individual member of the ensemble. Based on these indices, we consider this refinement successful and the MDtar trajectory a representation of [Sp]-hybrid in solution.

During MDtar calculations, the molecule drifts toward geometries with negative inclination (up to −20.0°), but to a lesser extent than in rMD and not irreversibly. In fact, during MDtar simulations, the molecule fluctuates between structures with negative and near-zero inclination, for an average value of −8.7° for the whole trajectory (vs. −13.1° for rMD; Table 5). Furthermore, propeller twist becomes consistently positive only for A6-U15 basepair (up to ~20.0°), although X-displacement remains negative (Fig. 5). Whereas for the RNA residues, sugars stayed in the C3′-endo conformations (Fig. 6 *A*, *right panels*), sugars for the PSO residues adopted multiple conformations (Fig. 6 *A*, *gray trace* in the *left panels* and *gray histograms* in the *middle panels*). Fig. 7 *A* shows distribution of Twist, Slide, and Roll helical parameters calculated with the CURVES program; the distributions are mostly monomodal for the [Sp]-hybrid.

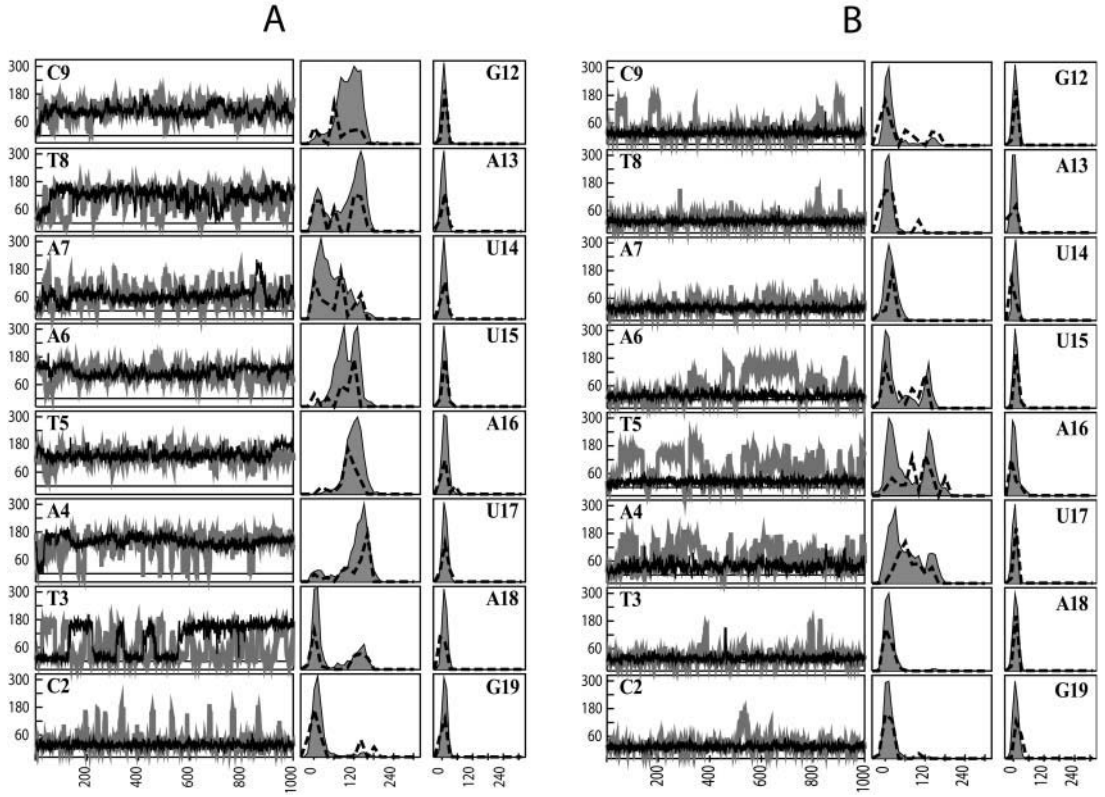


FIGURE 6 Sugar pucker analysis for the nonterminal residues in the [Sp]-hybrid (*A*) and [Rp]-hybrid (*B*). For each hybrid, left panels show the time course of the pseudorotation phase angle (***P***) in the PSO strand during the rMD (*black solid lines*) and MDtar (*gray lines*) production runs. Center panels show distributions of ***P***-values in the PSO strand in the MDtar ensemble (*gray histograms*) and in the PDQPRO ensemble (*black dashed lines*). The PDQPRO distributions are calculated using the PDQPRO probabilities. Right panels show similar distributions of ***P***-values for the RNA strand.

TABLE 3 Figures of merit for ensembles of the [Rp]- and [Sp]-hybrids

					$\langle R^x \rangle$ ($\times 100$)$^{\|}$			$ensR^x$ ($\times 100$)**	
Ensemble	Number*	$\langle E_1 \rangle^{\dagger}$	$\langle E_2 \rangle^{\ddagger}$	$Rdev^{\S}$	(1)	(2)	$dev\langle r^{-3} \rangle^{\P}$	(1)	(2)
[Sp]-hybrid									
rMD	1000	−42196 (60)	−1147 (20)	0.141 (0.02)	8.83 (0.42)	8.49 (0.40)	0.095	7.31	7.04
MDtar	1000	−42199 (57)	−1153 (12)	0.161 (0.02)	8.99 (0.39)	8.64 (0.35)	0.081	6.84	6.72
PDQPRO	26††	−42208 (58)	−1153 (10)				0.075	6.36	6.38
[Rp]-hybrid									
rMD	1000	−42248 (58)	−1168 (12)	0.095 (0.01)	8.34 (0.34)	7.03 (0.26)	0.076	7.42	6.20
MDtar	1000	−42238 (59)	−1151 (16)	0.116 (0.01)	8.64 (0.39)	7.43 (0.32)	0.063	7.05	5.91
PDQPRO	15‡‡	−42233 (44)	−1148 (18)				0.065	6.79	5.80

Where relevant, the standard deviation from the average is written within parentheses below the corresponding average value. For the definitions of the various ensembles, see text.

*Number of structures in an ensemble.

†Average AMBER energy (kcal/mol) calculated for the solvated system including water molecules and counter ions.

‡Average AMBER energy (kcal/mol) calculated after stripping the water molecules and counter ions, adding big hydrated sodium ions to neutralize the hybrid and optimizing their position by energy minimization while keeping the hybrid fixed.

§Average deviation of distances from the experimental NOE bounds, Å.

¶Average deviation of the third-root-weighted ensemble-average distances from the experimental NOE bounds, Å.

‖Average NOE-based CORMA R^x-factor calculated independently for each structure of an ensemble. (1) and (2) refer to the two NOESY data sets (see text).

**NOE-based ensemble-CORMA R^x-factor (see text). (1) and (2) refer to the two NOESY data sets.

††Sorted probabilities of 13 PDPQRO structures with probabilities above 3%: 10.8, 9.0, 8.7, 8.7, 6.7, 6.2, 5.3, 5.1, 4.1, 3.8, 3.6, 3.2, and 3.1%. Thirteen more structures have probabilities between 0.1% and 2.7%.

‡‡Sorted probabilities of 15 PDQPRO structures: 19.0, 15.0, 12.9, 10.7, 7.2, 6.5, 5.5, 5.2, 4.5, 4.3, 4.2, 3.0, 1.0, and 0.6%.

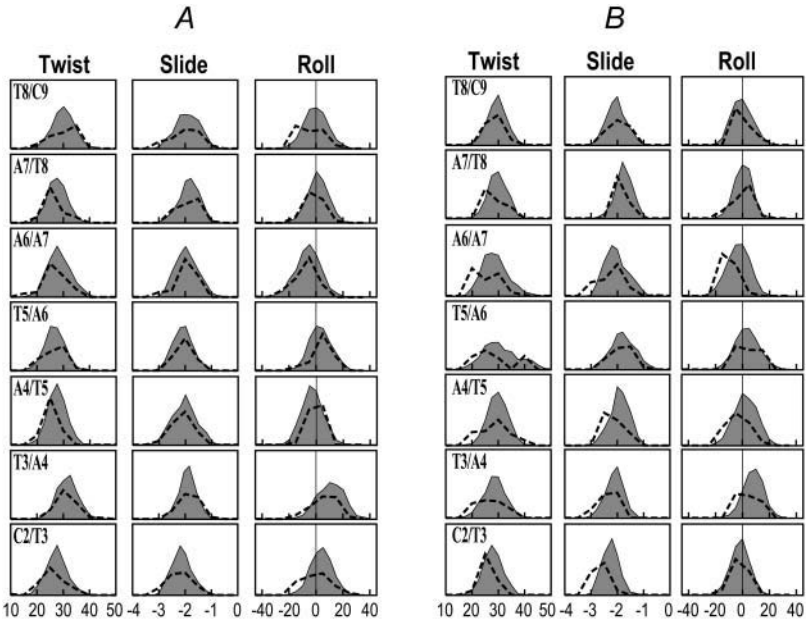


FIGURE 7 Distributions of local CURVES parameters Twist, Slide, and Roll for internal dinucleotide steps calculated from the MDtar (*gray histograms*) and PDQPRO (*black dashed lines*) ensembles: (*A*) [Sp]-hybrid; (*B*) [Rp]-hybrid. The PDQPRO distributions are weighted based on PDQPRO probabilities.

One thousand saved pdb files from the MDtar trajectory were subjected to the PDQPRO analysis. PDPQRO finds a subset of structures and their probabilities so that the resulting ensemble produces the best fit of experimental NMR parameters, NOE-derived cross-relaxation rates. PDQPRO selected 26 structures out of 1000 with the probabilities ranging from 0.1% to 10.8%. AMBER energies and various figures of merit are reported for the PDQPRO

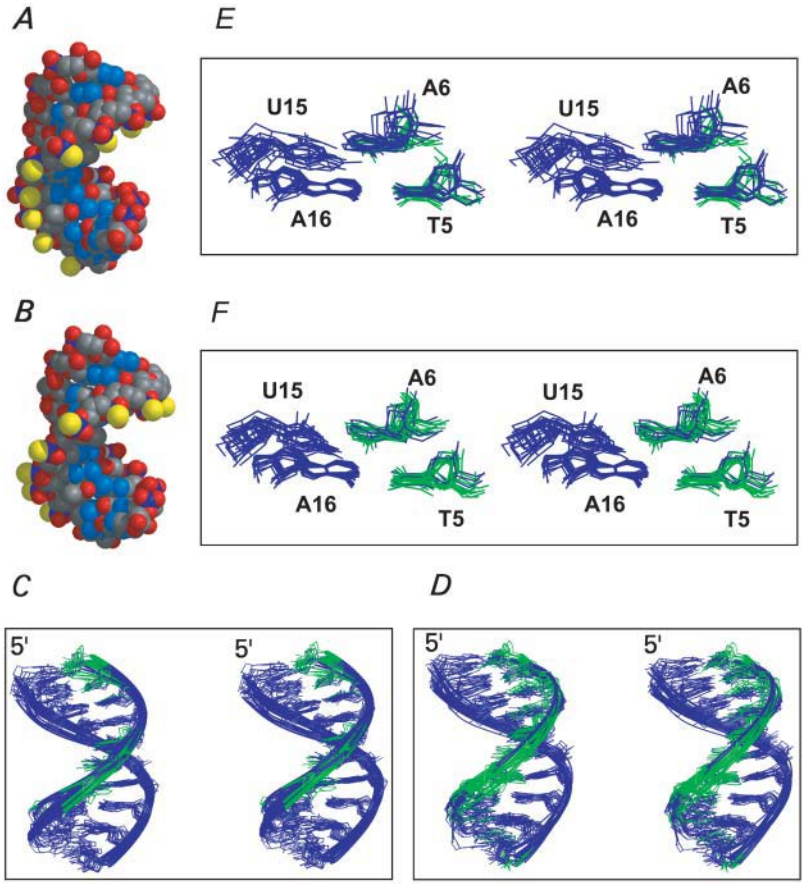


FIGURE 8 Molecular graphics showing (*A*) most populated PDQPRO conformer for the [Rp]-hybrid (19%); (*B*) most populated PDQPRO conformer for the [Sp]-hybrid (11%); (*C* and *E*) superposition of 11 PDQPRO conformers with populations >3% for the [Rp]-hybrid; (*D* and *F*) superposition of 13 PDQPRO conformers with populations >3% for the [Sp]-hybrid. In *A* and *B*, atoms are colored according to the atom type. In *C–F*, residues with C3′-endo sugar pucker are colored blue, and residues with C2′-endo sugar pucker are colored green. In *E* and *F*, bases T5 and A16 are superimposed. *C–F* are in stereo. This figure was generated by using the molecular graphics program MIDAS (Ferrin et al., 1988).

ensemble in Table 3; average values and standard deviations were calculated by taking into account the probability of each structure. Members of PDQPRO ensemble with probabilities >3% are shown in Fig. 8.

Structure calculation for the [Rp]-hybrid

Calculations for this molecule were carried out similarly to the [Sp]-hybrid. The [Rp]-hybrid seems to behave better during PME rMD and MDtar production runs. Both simulations yielded reasonable duplex geometries, allowing us to use a stronger force constant for NMR restraints (Table 2). Consequently, the refined structure for the [Rp]-hybrid features lower distance deviation values and lower R^x-factors compared to the [Sp]-hybrid.

The $\langle Rdev \rangle$ index for rMD on the [Rp]-hybrid is 0.10 Å, which could be considered acceptable. Still, the MDtar ensemble description of the molecules offers a significant improvement in figures of merit. Index $dev\langle r^{-3} \rangle$ for MDtar is 34% lower than $\langle Rdev \rangle$ for rMD, and ensemble R^x-factor for MDtar is 16–18% lower than the average R^x-factor for rMD (the range is given for the two data sets, at 150 and 300 ms). Thus, we consider MDtar trajectory a more accurate description of the [Rp]-hybrid in solution than a conventional rMD description. PDQPRO calculations selected 15 conformers out of the MDtar trajectory; probabilities of individual conformers varied from below 1% to 19%. Parameters of the PDQPRO ensemble are reported in Table 3. Figs. 6 *B* and 7 *B* show distributions of sugar conformations and helical parameters for the MDtar and PDQPRO ensembles for the [Rp]-hybrid.

DISCUSSION

rMD versus MDtar calculations

Molecular dynamics with distance restraints is a robust refinement method of NMR structures for nucleic acids when a molecule possesses a well-defined structure in solution (Schmitz et al., 1992b; Mujeeb et al., 1993; Weisz et al., 1994; Jucker and Pardi, 1995; Conte et al., 1997; Tonelli et al., 1998; Thiviyanathan et al., 1999). A successful rMD refinement leads to a set of similar structures; quality of refinement can be assessed by calculating residual violation of distance restraints (*Rdev*) and by comparing experimental and predicted NOE intensities (using, e.g., sixth-root weighted R^x-factor). Often it is necessary to find an optimal balance between empirical force field and experimental restraints, so that satisfying restraints does not compromise conformational energy. Other factors contributing to successful refinement include accurate experimental restraints and accurate empirical force field. To define a nucleic acid conformation, especially in the absence of residual dipolar couplings, distance bounds must be as tight as warranted by experimental data; such bounds can be determined by using full relaxation matrix analysis of NOE data, e.g., with MARDIGRAS. From the perspective of the empirical force field, it may be critical to use explicit solvent with the PME scheme, especially for RNA molecules, where hydroxyl groups of riboses may form spurious hydrogen bonds during in vacuo calculations (K.-S. Ryu, unpublished data). On the other hand, when rapidly exchanging multiple conformations contribute to the observed time-averaged NMR signal, the derived experimental restraints correspond to some virtual structure, which may or may not be significantly populated in solution. Practically, enforcing such restraints during conventional rMD refinement may lead to a number of different scenarios. If the virtual structure has in fact a relatively low energy, i.e., represents one of the solution conformers, or if the empiric force field is not accurate enough to reflect the differences in the populations of solution conformers, then the refinement may produce a well-defined structure, and we will just not detect the existence of multiple conformers. If the experimental restraints are not precise enough (we are assuming that they are accurate though), e.g., if distance bounds are not very tight, then the resulting structure will not be very well defined, or it will be defined more by the empiric force field than by experimental data. The only scenario when we can infer the existence of multiple conformers based on conventional refinement is when the experimental restraints are sufficiently tight and the virtual structure has high conformational energy. Practically, this will be manifest by high conformational energy, high restraint energy, or both; increasing the weight of restraints would only deteriorate the quality of the structure. By this token, NMR data for the [Sp]-hybrid clearly cannot be fit by a single conformation. The rMD structures have high residual distance deviation *Rdev* and relatively high NOE-based R^x-factors for both data sets at 100 and 200 ms (Table 3). At the same time, as explained in Results, rMD produces structures of poor quality for the [Sp]-hybrid. It is noteworthy though that it is difficult to correlate the quality of the structure as assessed by visual inspection and analysis of helical parameters with the AMBER conformational energy (Table 3). Indeed, although the energy difference between feasible and unfeasible conformations may be of the order of 10° kcal/mole, the total AMBER energy is of the order of 10^3–10^4 kcal/mol.

There are several approaches proposed for NMR refinement of systems with multiple conformers (Torda et al., 1990; Ulyanov et al., 1995; Bonvin and Brünger, 1995; Fennen et al., 1995; Kemmink and Scheek, 1995; Pearlman, 1996; Gyi et al., 1998; Görler et al., 2000). We have successfully used the MDtar method of Torda et al. in the past (Schmitz et al., 1993; Gonzàlez et al., 1995; Yao et al., 1997). Instead of enforcing all restraints simultaneously for each snapshot, it is only required in MDtar simulations that the restraints are satisfied over the course of trajectory for the appropriately averaged parameters, third-root averaged distances in this case. Individual MDtar snapshots have

very high residual distance deviations for the [Sp] hybrid (average $\langle Rdev \rangle$ is 0.16 Å; Table 3), even higher than individual snapshots of rMD trajectory. However the distances third-root averaged over the whole MDtar trajectory have a low residual deviation of 0.08 Å ($dev\langle r^{-3} \rangle$ in Table 3). Comparing the MDtar trajectory to individual snapshots of rMD with the average $\langle Rdev \rangle$ of 0.14 Å (rMD assumes that there is a single conformer), this is a 42% improvement. Ultimately, the quality of refinement must be judged by how well the refined structures predict experimentally measured parameters, NOE intensities in this case. Ensemble R^x-factors, $ensR^x$, show a better than 20% improvement for the MDtar trajectory compared to average R^x-factors of rMD (Table 3). Interestingly, ensemble figures of merit, $dev\langle r^{-3} \rangle$ and $ensR^x$, calculated for the rMD trajectory, also show some improvement compared to average values of $\langle Rdev \rangle$ and R^x for the same trajectory, although the restraints were imposed on each individual time frame and not on the distances time averaged over the rMD trajectory. Apparently, this happened because even rMD simulations captured some degree of flexibility of [Sp]-hybrid, such as sugar flexibility in residues T3, A7, and T8 (Fig. 6 *A*). However, the improvement is not as dramatic as in the case of MDtar, and the overall quality of structures was poor (see Results).

The situation is somewhat less clear in the case of the [Rp]-hybrid. To begin with, snapshots of the rMD trajectory do not have such high residual distance violation, 0.095 Å on average, and the overall quality of structures is better than in the case of the [Sp]-hybrid (see Results). This could be because [Rp]-hybrid is indeed somewhat less flexible, or because the MARDIGRAS-calculated distance restraints are less tight (average flat-well width of 1.66 Å compared to 1.51 Å for the [Sp]-hybrid). Nevertheless, the MDtar trajectory shows a significant amount of flexibility for the [Rp]-hybrid as well (Figs. 6 *B* and 7 *B*). Most importantly, ensemble figures of merit are still significantly improved for the MDtar compared to average individual figures of merit for the rMD: distance deviation by 34%, and NOE-based R^x-factors by 15% (Table 3).

It must be clear of course, that we do not have enough experimental data to define each conformer to a high resolution, especially for very flexible molecules like the [Sp]-hybrid or [Rp]-hybrid, where much of the flexibility occurs at the nucleotide level with sugar repuckering of deoxyribose rings (see below). Instead, the MDtar trajectory must be considered as a representative ensemble of structures capturing elements of flexibility necessary to fit the observed data. In this case, it was necessary to assume sugar flexibility in the PSO strands to explain certain experimental distance restraints, most notably intraresidue H2′-H6/H8 and H3′-H6/H8 (see Results). Such a representative ensemble is not unique; it can be selected in different ways; the ensemble of 1000 MDtar structures is likely to be redundant, in the sense that a smaller set of structures should be able to explain the experimental data. Apart from these considerations, 1000 structures are not very convenient to deal with when analyzing them.

Previously, we have used PDQPRO in combination with MDtar to select a small ensemble capable of fitting the NMR data (Schmitz et al., 1998; Aramini et al., 2000). Given a pool of potential conformations, PDQPRO finds a subset of structures and their probabilities with the best fit to the experimental data. Even though there is no intrinsic mechanism in PDQPRO to reduce the size of the ensemble, in our experience, finding the best fit does reduce the potential pool significantly, most probably due to the elimination of redundant structures. Out of 1000 MDtar structures, PDQPRO calculations selected 26 conformers for the [Sp]-hybrid and 15 for [Rp]-hybrid. During the selection process, PDQPRO optimizes a quadratic function of NOE-derived cross-relaxation rates, but not directly distance deviation or NOE R^x-factor. Because of that, it cannot be expected that ensemble figures of merit, $dev\langle r^{-3} \rangle$ and $ensR^x$ will further improve; however, they did improve a little for both hybrids compared to full MDtar trajectories (Table 3). Importantly, PDQPRO ensembles appear to be good representatives of the full MDtar trajectories: they have similar average and standard deviations of AMBER energy, almost identical average pairwise atomic RMSD (Tables 3 and 4). Even the distributions of sugar pseudorotation phase angles and helical parameters (Figs. 6 and 7) have many common features between MDtar and PDQPRO ensembles.

Sugar conformations

The most striking feature of the solution ensembles for both hybrids is rigid C3′-endo conformations of riboses in RNA strands and very flexible deoxyriboses in PSO strands (Fig. 6). This feature of PSO-RNA hybrid is common with regular DNA-RNA hybrids (Gonzàlez et al., 1995; Gyi et al., 1998). It has been known that deoxyriboses are also flexible in DNA-DNA duplexes (Tonelli and James, 1998; Schmitz et al., 1990, 1992b; Weisz et al., 1992; Mujeeb et al., 1992; Rinkel and Altona, 1987; Tonelli et al., 1998). However, the population of N-conformations is relatively small in DNA duplexes and it often can be ignored during refinement without significant deterioration of quality of structures. In

TABLE 4 Atomic RMSD for the ensembles of the [Sp]- and [Rp]-hybrids

	[Sp]-hybrid			[Rp]-hybrid		
Ensemble	All	DNA	RNA	All	DNA	RNA
rMD	2.4 (1.3)	1.8 (0.9)	1.9 (1)	0.9 (0.2)	0.7 (0.2)	0.7 (0.2)
MDtar	1.4 (0.4)	1.2 (0.3)	1.1 (0.4)	1.5 (0.5)	1.2 (0.4)	1.0 (0.4)
PDQPRO	1.3 (0.4)	1.2 (0.3)	1.1 (0.4)	1.5 (0.5)	1.2 (0.4)	1.0 (0.4)

Average and standard deviations (in parentheses) for pairwise RMSD values (Å) for each of the ensembles. RMSD values were calculated using heavy atoms excluding terminal residues.

contrast, populations of both N- and S-puckers are high for deoxyriboses in both DNA-RNA and PSO-RNA duplexes. Interestingly, distributions of sugar pucker are complex and not consistent with a simple two-state jump model (Fig. 6). Ignoring sugar flexibility in this case could either lead to completely unphysical structures (as in the case of rMD refinement of the [Sp]-hybrid), or lead to a suboptimal virtual structure (as in the case of the [Rp]-hybrid); this has been also noted by others (Fedoroff et al., 1997; Gyi et al., 1998). Flexibility of deoxyriboses appears to be more pronounced for the [Sp]-hybrid and to depend on sequence (Fig. 6 *A*). However, it is not clear to what degree this result depends on what types of distance restraints were measured for each residue; some of the NOE crosspeaks could not be integrated because of spectral overlap. For regular DNA-RNA hybrids, sequence dependence of sugar flexibility has been reported before; pyrimidine-rich DNA appears to be more flexible than purine-rich DNA strands (Gyi et al., 1998).

Helical parameters

The overall geometries of [Sp]- and [Rp]-hybrids are shown in Fig. 8, *A* and *B*. The conformations of the two hybrids are similar, with the atomic RMSD between the most probable PDQPRO conformers of 1.3 Å. The sulfur atoms are pointing toward solution in the [Sp]-hybrid; they are directed more toward the major groove in the [Rp]-hybrid, but they are not making any specific interactions with atoms from the other strand or from the walls of the groove. The global conformational parameters of the hybrids are clearly intermediate between those typical for classical B- and A-conformations (Table 5). X-displacement is approximately −2 Å; it still creates a characteristic central hole in the duplex, but not as large as in the canonical A-conformation (a typical X-displacement for A-forms is approximately −4 Å). Inclination of basepairs relative to the global axis of the duplex is negative, which is typical of B-conformations; however, the variation of this parameter between individual conformers is very high for both hybrids, with standard deviations of 8–9°. Local helical parameters are sequence dependent and exhibit a high degree of variation between conformers; some of them appear to have bimodal distributions (Fig. 7). Of note is, e.g., the T5-A6 step, which has low-twist and high-twist conformers (see also Fig. 8, *E* and *F*). For the [Rp]-hybrid, twist varies between 22° and 40° for this step; for the [Sp]-hybrid, the range is 16–34°. This step precedes a stretch of two adenines; in DNA, TA steps preceding a stretch of adenines have been shown to undergo conformational averaging (Schmitz et al., 1992b; Kennedy et al., 1993; McAteer et al., 1995; McAteer and Kennedy, 2000), although the structural details of this dynamic process are not known.

Implications for RNase H recognition

DNA-RNA and PSO-RNA hybrid duplexes are recognized by RNase H with the subsequent cleavage of RNA residues. All these types of hybrids are extremely flexible in solution, with deoxyriboses of DNA or PSO strands undergoing conformational switches between C3′-endo and C2′-endo puckers. On the other hand, RNA-RNA duplexes with rigid C3′-endo riboses are not substrates for RNase H. Furthermore, introducing 2′-substituents that lock deoxyriboses in C3′-endo conformations and result in rigid DNA-RNA structures also result in the disappearance of RNase H recognition. We propose a hypothesis that flexibility of the hybrid duplex is required for RNase H recognition. This hypothesis is consistent with the recent crystal structure of a complex of HIV-1 reverse transcriptase with a DNA-RNA hybrid duplex (Sarafianos et al., 2001). As expected, all RNA residues in this hybrid have C3′-endo sugar puckers. On the other hand, the DNA residues clearly exhibit conformational flexibility: part of the hybrid duplex interacting with the polymerase domain has deoxyriboses in C3′-endo conformation, whereas DNA residues interacting with the RNase H domain have C2′-endo sugar puckers.

TABLE 5 Average and standard deviation values for selected helical parameters for the ensembles of the [Sp]- and [Rp]-hybrids

	Roll	Slide	Twist	Rise	Inclin	Propeller	X-displ	Minor groove width
[Sp]-hybrid								
rMD	0.8 (11.6)	−2.04 (0.48)	28.6 (4.7)	3.41 (0.38)	−13.1 (13.0)	−7.8 (15.6)	−1.83 (1.10)	8.34 (0.57)
MDtar	3.8 (9.8)	−1.86 (0.40)	29.6 (4.2)	3.34 (0.37)	−8.7 (9.1)	−12.0 (12.8)	−2.06 (0.91)	8.77 (0.55)
PDQPRO	4.9 (9.7)	−1.75 (0.43)	29.8 (4.6)	3.29 (0.38)	−5.1 (7.9)	−14.8 (12.1)	−2.35 (0.73)	9.04 (0.59)
[Rp]-hybrid								
rMD	4.0 (6.7)	−1.88 (0.31)	31.0 (3.4)	3.22 (0.33)	−2.5 (6.0)	−20.9 (9.8)	−2.94 (0.62)	9.36 (0.22)
MDtar	3.9 (8.5)	−1.86 (0.40)	30.5 (4.5)	3.27 (0.36)	−7.8 (8.9)	−16.5 (11.9)	−2.20 (0.94)	9.03 (0.49)
PDQPRO	3.8 (8.6)	−1.88 (0.41)	30.0 (4.7)	3.26 (0.35)	−11.8 (7.5)	−15.9 (12.1)	−1.75 (0.75)	8.81 (0.50)
HIV-1*	−6.3	−1.61	29.1	3.81	−20.4	1.6	−1.01	7.60
HIV-1†	0.7	−1.74	32.9	3.21	−3.8	1.6	−2.16	9.04

All parameters were calculated using CURVES (Lavery and Sklenar, 1996). Roll and Slide are local helical parameters and the rest are global helical parameters. Translations are in Å, and rotations are in degrees.

*Polypurine tract of a RNA·DNA hybrid from the crystal structure of the complex with the HIV-1 reverse transcriptase (PDB 1HYS; Sarafianos et al., 2001) downstream of the out-of-register basepairs (see text).

†Portion of the same RNA·DNA hybrid upstream of the out-of-register basepairs.

Furthermore, the irregularities in the basepairing in the polypurine tract (PPT) region of the hybrid DNA-RNA duplex in the vicinity of the RNase H active site (Sarafianos et al., 2001) suggest that the helical conformation of the hybrid is also very flexible, similar or even higher than observed for the PSO-RNA hybrids in this work. Two basepairs within the PPT are out-of-register mismatches, which may be a feature of the PPT sequence possessing an unusual resistance to RNase H cleavage. X-displacement of basepairs for the portion of the hybrid interacting with RNase (downstream of the out-of-register basepairs) is -1 Å on average and the average inclination has a very unusual value of $-20°$ (Table 5). The minor groove width for this portion of the hybrid is greatly diminished and the depth is increased compared to typical A-conformations, which facilitates contacts between RNase H residues and the hybrid. The authors argue that the RNase cleavage specificity is controlled by the width of the minor groove and by the overall very unusual trajectory of DNA-RNA duplex (Sarafianos et al., 2001).

The structural ensembles calculated here for both PSO-RNA hybrid duplexes are consistent with the notion that such molecules have enough structural plasticity to accommodate the unusual requirements for the RNase H recognition. The minor groove width of [Sp]- and [Rp]-hybrids is also greatly reduced compared to A-conformations (Fig. 9). Relatively small X-displacement and negative inclination (Table 5) are very unusual features, clearly distinguishing these structures from classical B- and A-conformations, not unlike the structure of the DNA-RNA hybrid complexed with the HIV-1 reverse transcriptase. Most importantly, both helical geometry and sugar conformations of deoxyriboses are extremely flexible (Figs. 5–7), suggesting that these structures could easily fit with the RNase H. In contrast, introducing 2′-substituents into deoxyriboses of the antisense strand locks sugars in the C3′-endo conformations and makes the hybrid structures rigid (K.-S. Ryu, unpublished data). It would be impossible for such rigid structures to adopt very unusual conformations required for the RNase H recognition, which is consistent with them not being a RNase H substrate.

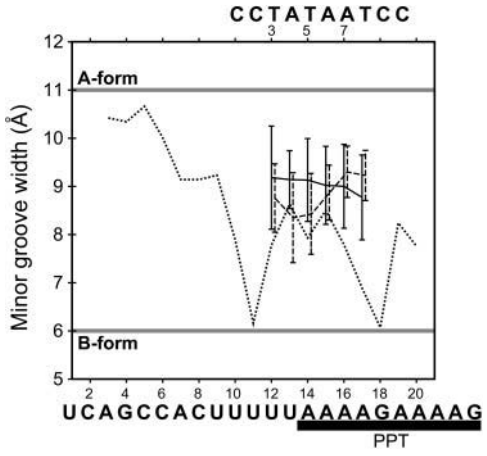


FIGURE 9 Minor groove width (Å) calculated with the CURVES program (Lavery and Sklenar, 1996). (*Dotted line*) RNA·DNA hybrid from the crystal structure of the complex with the HIV-1 reverse transcriptase (PDB 1HYS (Sarafianos et al., 2001)). The lower *x* axis shows the numbering and sequence of the RNA strand. The downstream portion of the sequence with the narrow minor groove corresponds to the region interacting with the RNase H site. The solid and dashed lines with the error bars show the average and standard deviation values for the PDQPRO ensembles of [Sp]- and [Rp]-hybrids, respectively; the upper *x* axis shows the numbering and the sequence of the PSO strands. The minor groove width values for the canonical A- and B-conformations are 11 Å and 6 Å, respectively.

CONCLUSIONS

We have used homonuclear 2D NMR to solve the structure of [Sp]-phophorothioated DNA-RNA hybrid duplex and the corresponding [Rp]-isomer with the same sequence. Our results clearly indicate that the DNA strands in these hybrids are more flexible than in a corresponding DNA duplex. Indeed, the [Sp]-hybrid required the use of time-averaged restraints during MD simulations to get an ensemble of structures capable of describing the experimental data. The [Rp]-hybrid appears to be less flexible, so we were able to produce stable trajectories even with regular rMD simulations. It did, however, show a significant amount of flexibility when MDtar was used, yielding an ensemble of structures that fit the NMR data better (see Results). This flexibility is not just localized for the sugars of the DNA strands; the overall helical structure is also flexible, as indicated by the high variability of helical parameters in our final ensembles of structures. Some helical parameters even exhibited a clear bimodal distribution (see Results and Discussion).

In principle, it is possible that this conformational variability is not real, but rather is an indirect result of the paucity of our NMR restraints that are relatively few in number and have quite large flat-well widths (the average widths are 1.51 Å and 1.66 Å in the [Sp]- and [Rp]-hybrids, respectively). However, we argue that sugar rings in the RNA strand of both hybrids are locked in C3′-endo conformations and do not show any sign of repuckering, even though they are restrained by fewer distance bounds that the DNA deoxyribose sugars (35% and 18% less distance restraints for the RNA strand in the [Sp]- and [Rp]-hybrid, respectively). Moreover, the crystal structure of a DNA-RNA hybrid bound to HIV-1 reverse transcriptase (Sarafianos et al., 2001) shows a great deal of conformational heterogeneity, not only limited to the pucker of the DNA sugars but also involving the helical geometry, that resembles the structures of our hydrids (see Discussion).

We conclude that our PSO-RNA hybrids must indeed be flexible, both at the ribose level and at the level of global helical geometry. Taking into account that RNA-RNA duplexes with rigid sugar rings are not substrates for RNase

H, we hypothesize that this flexibility must be critical for the RNase H recognition of duplexes.

SUPPLEMENTARY MATERIAL

An online supplement to this article can be found by visiting BJ Online at http://www.biophysj.org.

We gratefully acknowledge Dr. Vladimir Basus for help in running the NMR spectrometers, Eric Pettersen for the help with molecular graphics, Davide Tonelli for the help with running MD on a linux cluster, and the University of California, San Francisco, Computer Graphics Lab for the use of their facilities.

This work was supported by National Institutes of Health grants GM39247 and CA25644.

REFERENCES

Agrawal, S., and R. P. Iyer. 1997. Perspectives in antisense therapeutics. *Pharmacol. Ther.* 76:151–160.

Altmann, K.-H., D. Fabbrot, N. M. Dean, T. Geigert, B. P. Monia, M. Müller, and P. Nicklin. 1997. Second-generation antisense oligonucleotides: structure-activity relationships and the design of improved signal-transduction inhibitors. *Biochem. Soc. Trans.* 24:630–637.

Aramini, J. M., A. Mujeeb, N. B. Ulyanov, and M. W. Germann. 2000. Conformational dynamics in mixed alpha/beta-oligonucleotides containing polarity reversals: a molecular dynamics study using time-averaged restraints. *J. Biomol. NMR.* 18:287–302.

Berendsen, H. J. C., J. P. M. Postma, W. F. van Gunsteren, A. Di Nola, and J. R. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684–3690.

Bonvin, A. M. J. J., and A. T. Brünger. 1995. Conformational variability of solution nuclear magnetic resonance structures. *J. Mol. Biol.* 250:80–93.

Borgias, B. A., and T. L. James. 1990. MARDIGRAS—procedure for matrix analysis of relaxation for discerning geometry of an aqueous structure. *J. Magn. Reson.* 87:475–487.

Broido, M. S., G. Zon, and T. L. James. 1984. Complete assignment of the non-exchangeable proton NMR resonances of $[d\text{-}(GGAATTCC)]^2$ using two-dimensional nuclear Overhauser effect spectra. *Biochem. Biophys. Res. Commun.* 119:663–670.

Case, D. A., D. A. Pearlman, J. W. Caldwell, I. T. E. Cheatham, W. S. Ross, C. Simmerling, T. Darden, K. M. Merz, R. V. Stanton, A. Chen, J. J. Vincent, M. Crowley, V. Tsui, R. Radmer, Y. Duan, J. Pitera, I. Massova, and P. A. Kollman. 2000. Amber 6.0. University of California, San Francisco.

Conte, M. R., G. L. Conn, T. Brown, and A. N. Lane. 1997. Conformational properties and thermodynamics of the RNA duplex r(CGCAAAUUUGCG)2: comparison with the DNA analogue $d(CGCAAATTTGCG)_2$. *Nucleic Acids Res.* 25:2627–2634.

Crooke, S. T., and C. F. Bennett. 1996. Progress in antisense oligonucleotide therapeutics. *Annu. Rev. Pharmacol. Toxicol.* 36:107–129.

Delaglio, F., S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, and A. Bax. 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR.* 6:277–293.

DeLong, R. K., A. Nolting, M. Fisher, Q. Chen, E. Wickstrom, M. Kligshteyn, S. Demirdji, M. Caruthers, and R. L. Juliano. 1997. Comparative pharmacokinetics, tissue distribution, and tumor accumulation of phosphorothioate, phosphorodithioate, and methylphosphonate oligonucleotides in nude mice. *Antisense Nucleic Acid Drug Dev.* 7:71–77.

Essmann, U., L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.* 103:8577–8593.

Fedoroff, O. Y., Y. Ge, and B. R. Reid. 1997. Solution structure of r(gaggacug):d(CAGTCCTC) hybrid: implications for the initiation of HIV-1 (+)-strand synthesis. *J. Mol. Biol.* 269:225–239.

Fennen, J., A. E. Torda, and W. F. v. Gunsteren.1995. Structure refinement with molecular dynamics and a Boltzmann-weighted ensemble. *J. Biomol. NMR.* 6:163–170.

Ferrin, T. E., C. C. Huang, L. E. Jarvis, and R. Langridge. 1988. The MIDAS display system. *J. Mol. Graph.* 6:13–27.

Freier, S. M., and K. H. Altmann. 1997. The ups and downs of nucleic acid duplex stability: structure-stability studies on chemically-modified DNA:RNA duplexes. *Nucleic Acids Res.* 25:4429–4443.

Furrer, P., T. M. Billeci, A. Donati, C. Kojima, B. Karwowski, A. Sierzchala, W. Stec, and T. L. James. 1999. Structural effect of complete R_p phosphorothioate and phosphorodithioate substitutions in the DNA strand of a model antisense inhibitor-target RNA complex. *J. Mol. Biol.* 285:1609–1622.

Goddard, T. D., and D. G. Kneller. 1998. SPARKY 3.0. University of California, San Francisco.

Gonzàlez, C., W. Stec, A. Kobylanska, R. Hogrefe, M. Reynolds, and T. L. James. 1994. Structural study of a DNA-RNA hybrid duplex with a chiral phosphorothioate moiety by NMR: extraction of distance and torsion angle constraints and imino proton exchange rates. *Biochemistry.* 33:11062–11072.

Gonzàlez, C., W. Stec, M. Reynolds, and T. L. James. 1995. Structure and dynamics of a DNA-RNA hybrid duplex with a chiral phosphorothioate moiety. NMR and molecular dynamics with conventional and time-averaged restraints. *Biochemistry.* 34:4969–4982.

Görler, A., N. B. Ulyanov, and T. L. James. 2000. Determination of the populations and structures of multiple conformers in an ensemble from NMR data: multiple-copy refinement of nucleic acid structures using floating weights. *J. Biomol. NMR.* 16:147–164.

Gyi, J. I., A. N. Lane, G. L. Conn, and T. Brown. 1998. Solution structures of DNA-RNA hybrids with purine-rich and pyrimidine-rich strands: comparison with the homologous DNA and RNA duplexes. *Biochemistry.* 37:73–80.

Huang, C. C., G. S. Couch, E. F. Pettersen, and T. E. Ferrin. 1996. Chimera: an extensible molecular modeling application constructed using standard components. *Pac. Symp. Biocomp.* 1:724.

Jucker, F. M., and A. Pardi. 1995. Solution structure of the CUUG hairpin loop: a novel RNA tetraloop motif. *Biochemistry.* 34:14416–14427.

Keepers, J. W., and T. L. James. 1984. A theoretical study of distance determinations from NMR. Two-dimensional nuclear Overhauser effect spectra. *J. Magn. Reson.* 57:404–426.

Kemmink, J., and R. M. Scheek. 1995. Dynamic modelling of a helical peptide in solution using NMR data: multiple conformations and multi-spin effects. *J. Biomol. NMR.* 5:33–40.

Kennedy, M. A., S. T. Nuutero, J. T. Davis, G. Drobny, and B. R. Reid. 1993. Mobility at the TpA cleavage site in the T_3A_3-containing *Aha*III and *Pme* restriction sequences. *Biochemistry.* 32:8022–8035.

Koradi, R., M. Billeter, and K. Wüthrich. 1996. MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* 14: 29–32.

Krakowiak, A., A. Owczarek, M. Koziolkiewicz, and W. J. Stec. 2002. Stereochemical course of *Escherichia coli* RNase H. *Chem. Biochem.* 3:1242–1250.

Lavery, R., and H. Sklenar. 1996. CURVES 5.1. Helical Analysis of Irregular Nucleic Acids. Laboratoire de Biochimie Theoretique, CNRS, Paris.

Lima, W. F., and S. T. Crooke. 1997. Binding affinity and specificity of *Escherichia coli* RNase H1: impact on the kinetics of catalysis of antisense oligonucleotide-RNA hybrids. *Biochemistry.* 36:390–398.

Liu, H., H. P. Spielmann, N. B. Ulyanov, D. E. Wemmer, and T. L. James. 1995. Interproton distance bounds from 2D-NOE intensities: effect of

experimental noise and peak integration errors. *J. Biomol. NMR.* 6:390–402.

Marquis, J. K., and J. M. Grindel. 2000. Toxicological evaluation of oligonucleotide therapeutics. *Curr. Opin. Mol. Ther.* 2:258–263.

McAteer, K., P. D. Ellis, and M. A. Kennedy. 1995. The effects of sequence context on base dynamics at TpA steps in DNA studied by NMR. *Nucleic Acids Res.* 23:3962–3966.

McAteer, K., and M. A. Kennedy. 2000. NMR evidence for base dynamics at all TpA steps in DNA. *J. Biomol. Struct. Dyn.* 17:1001–1009.

McKay, R. A., L. J. Miraglia, L. L. Cummins, S. R. Owens, H. Sasmor, and N. M. Dean. 1999. Characterization of a potent and specific class of antisense oligonucleotide inhibitor of human protein kinase C-alpha expression. *J. Biol. Chem.* 274:1715–1722.

Monia, B. P. 1997. First- and second-generation antisense inhibitors targeted to human c-*raf* kinase: *in vitro* and *in vivo* studies. *Anticancer Drug Des.* 12:327–339.

Mujeeb, A., S. M. Kerwin, W. Egan, G. L. Kenyon, and T. L. James. 1992. A potential gene target in HIV-1: rationale, selection of a conserved sequence, and determination of NMR distance and torsion angle constraints. *Biochemistry.* 31:9325–9338.

Mujeeb, A., S. M. Kerwin, G. L. Kenyon, and T. L. James. 1993. Solution structure of a conserved DNA sequence from the HIV-1 genome: restrained molecular dynamics simulation with distance and torsion angle restraints derived from 2D NMR spectra. *Biochemistry.* 32:13419–13431.

Pearlman, D. A. 1996. FINGAR: a new genetic algorithm-based method for fitting NMR data. *J. Biomol. NMR.* 8:49–66.

Pearlman, D. A., and P. A. Kollman. 1991. Are time-averaged restraints necessary for NMR refinement? A model study for DNA. *J. Mol. Biol.* 220:457–479.

Rinkel, L. J., and C. Altona. 1987. Conformational analysis of the deoxyribofuranose ring in DNA by means of sums of proton-proton coupling constants: a graphical analysis. *J. Biomol. Struct. Dyn.* 4:621–649.

Ryckaert, J. P., G. Cicotti, and H. J. C. Berendsen. 1977. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* 23:327–341.

Sarafianos, S. G., K. Das, C. Tantillo, A. D. Clark, Jr., J. Ding, J. M. Whitcomb, P. L. Boyer, S. H. Hughes, and E. Arnold. 2001. Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. *EMBO J.* 20:1449–1461.

Schmitz, U., A. Donati, T. L. James, N. B. Ulyanov, and L. Yao. 1998. Small structural ensembles for a 17-nucleotide mimic of the tRNA T psi C-loop via fitting dipolar relaxation rates with the quadratic programming algorithm. *Biopolymers.* 46:329–342.

Schmitz, U., A. Kumar, and T. L. James. 1992a. Dynamic interpretation of NMR structural data: molecular dynamics with weighted time-averaged restraints and ensemble R-factor. *J. Am. Chem. Soc.* 114:10654–10656.

Schmitz, U., I. Sethson, W. Egan, and T. L. James. 1992b. Solution structure of a DNA octamer containing the Pribnow box via restrained molecular dynamics simulation with distance and torsion angle constraints derived from 2D NMR spectral fitting. *J. Mol. Biol.* 227:510–531.

Schmitz, U., N. B. Ulyanov, A. Kumar, and T. L. James. 1993. Molecular dynamics with weighted time-averaged restraints of a DNA octamer: dynamic interpretation of NMR data. *J. Mol. Biol.* 234:373–389.

Schmitz, U., G. Zon, and T. L. James. 1990. Deoxyribose conformation in $[d(GTATATAC)]_2$: evaluation of sugar pucker by simulation of double-quantum-filtered COSY cross-peaks. *Biochemistry.* 29:2357–2368.

Simmerling, C., R. Elber, and J. Zhang. 1995. MOIL-View—a program for visualization of structure and dynamics of biomolecules and STO—a program for computing stochastic paths. *In* Modeling of Biomolecular Structures and Mechanisms. A. Pullman, editor. Academic Press, Kluwer, Netherlands. 241–65.

Stec, W. J., B. Karwowski, M. Boczkowska, P. Guga, M. Koziokiewicz, M. Sochacki, M. W. Wieczorek, and J. Baszczyk. 1998. Deoxyribonucleoside 3′-O-(2-Thio- and 2-Oxo-''spiro''-4,4-pentamethylene-1,3,2-oxathiaphospholane)s: monomers for stereocontrolled synthesis of oligo(deoxyribonucleoside phosphorothioate)s and chimeric PS/PO oligonucleotides. *J. Am. Chem. Soc.* 120:7156–7167.

Temsamani, J., and P. Guinot. 1997. Antisense oligonucleotides: a new therapeutic approach. *Biotechnol. Appl. Biochem.* 26:65–71.

Thiviyanathan, V., B. A. Luxon, N. B. Leontis, N. Illangasekare, D. G. Donne, and D. G. Gorenstein. 1999. Hybrid-hybrid matrix structural refinement of a DNA three-way junction from 3D NOESY-NOESY. *J. Biomol. NMR.* 14:209–221.

Tonelli, M., and T. L. James. 1998. Insights into the dynamic nature of DNA duplex structure via analysis of nuclear Overhauser effect intensities. *Biochemistry.* 37:11478–11487.

Tonelli, M., E. Ragg, A. M. Bianucci, K. Lesiak, and T. L. James. 1998. Nuclear magnetic resonance structure of d(GCATATGATAAG): a consensus sequence for promoters recognized by σ^K RNA polymerase. *Biochemistry.* 37:11745–11761.

Torda, A. E., R. M. Scheek, and W. F. van Gunsteren. 1990. Time-average nuclear Overhauser effect distance restraints applied to tendamistat. *J. Mol. Biol.* 214:223–235.

Ulyanov, N. B., U. Schmitz, A. Kumar, and T. L. James. 1995. Probability assessment of conformational ensembles: sugar repuckering in a DNA duplex in solution. *Biophys. J.* 68:13–24.

Weisz, K., R. H. Shafer, W. Egan, and T. L. James. 1992. The octamer motif in immunoglobulin genes: extraction of structural constraints from two-dimensional NMR studies. *Biochemistry.* 31:7477–7487.

Weisz, K., R. H. Shafer, W. Egan, and T. L. James. 1994. Solution structure of the octamer motif in immunoglobulin genes via restrained molecular dynamics calculations. *Biochemistry.* 33:354–366.

Wüthrich, K. 1986. NMR of Proteins and Nucleic Acids. Wiley, New York.

Yao, L. J., T. L. James, J. T. Kealey, D. V. Santi, and U. Schmitz. 1997. The dynamic NMR structure of the TΨC-loop: implications for the specificity of tRNA methylation. *J. Biomol. NMR.* 9:229–244.

Zhurkin, V. B., N. B. Ulyanov, A. A. Gorin, and R. L. Jernigan. 1991. Static and statistical bending of DNA evaluated by Monte Carlo simulations. *Proc. Natl. Acad. Sci. USA.* 88:7046–7050.