Solution Structure of an RNA·DNA Hybrid Duplex Containing a 3′-Thioformacetal Linker and an RNA A-Tract†

Collin W. Cross, Jeffrey S. Rice,‡ and Xiaolian Gao*

Department of Chemistry, University of Houston, Houston, Texas 77204-5641

Received September 20, 1996; Revised Manuscript Received December 18, 1996⊗

ABSTRACT: Neutral and achiral backbone linkers are promising replacements for the phosphodiester linkages of antisense oligonucleotides that target mRNA sequences. Results are presented here for the solution structure elucidation by NMR of an RNA·DNA hybrid duplex, r(GCGCAAAACGCG)·d(CGCGTT-SCH_2O-TTGCGC) (designated RIII), containing a 3′-thioformacetal (3′-TFMA) backbone substitution in the DNA strand. The 3′-thioformacetal linker can be accommodated in the hybrid duplex in a conformation that is drastically different from its form in a DNA·DNA duplex but close to that of the canonical A-form helix, reflecting the sequence requirement for hybridization. While the global features of RIII are similar to what are described in the literature, the 3′-TFMA modification drives sugar puckers of the adjacent residues to more C3′-endo-like conformations and causes distortions in related twist angles and helical rises. The helical conformation analyses of each of the two strands and the hybrid duplex enable a clear account of the conformational variability of both the DNA and RNA strands. The A-tract in the RNA strand features an overall straight helix and a more prominent bend at the 3′-end CG step. The structure of RIII provides a structural basis for the improved thermal stability of RIII compared to the corresponding DNA·DNA duplex and insights into the factors that are important concerns for the design of new, effective antisense oligonucleotides.

In the last decade, synthetic oligodeoxynucleotides have been recognized as a new generation of compounds possessing great biomedical and therapeutic potential (Agrawal & Iyer, 1995; Cole-Strauss et al., 1996; Crooke, 1992; Milligan et al., 1993; Stein, 1996; Uhlmann & Peyman, 1990; Zamecnik & Stephenson, 1978). Specially designed oligodeoxynucleotides are expected to bind in a complementary fashion, to a specific site in a gene sequence, offering selectivity which has not been possible with other therapeutic agents. The mode of these interactions, between oligodeoxynucleotides and target nucleic acid sequences, has been proposed to be similar to that of natural RNA antisense sequences, which suppress gene expression by various mechanisms such as interfering with transcription and binding to translation sites of mRNA (Matthiue et al., 1992; Pestka, 1992). The antisense effect has also been correlated to the formation of hybrid duplexes between oligodeoxynucleotides and target RNA. This promotes RNA strand cleavage by RNase H enzymes. Because of the potential advantages of antisense oligonucleotides and their ready availability from chemical synthesis, the development of effective antisense molecules has been an active area of research (references cited in this paragraph; Baserga & Denhardt, 1992; Crook, 1992; Jain, 1996).

A major task of the research in antisense oligonucleotides is to overcome the problems inherent to natural phosphodiester backbone linkers (Gura, 1996; Jain, 1996; Stein & Cheng, 1993). Phosphodiester groups are substrates of nucleases and incompatible in polarity with the lipophilic cell membrane. These properties render phosphodiester-containing antisense oligonucleotides ineffective for *in vivo* applications. Various backbone substitution strategies have been extensively investigated. These include modification of phosphodiester moieties by sulfur, amidate, or alkyl groups or the *de novo* design of backbone linkers (Marshall & Caruthers, 1993; Nielsen et al., 1991; Richert et al., 1996; Sanghvi & Cook, 1993; Wiekstron, 1995). Backbone substitutions offer a means to improve the pharmacological properties of antisense oligonucleotides, while maintaining their specificity, conferred by Watson−Crick base pair formation, at the binding site. Although substituting all or part of the phosphodiester backbone with the appropriate chemical moieties can alleviate many of the problems mentioned above, these backbone surrogates present difficulties as well. For instance, the most extensively used antisense phosphorothioates bind nonselectively to proteins, resulting in cell toxicity; the peptide−nucleotide linkage, PNA, modified oligonucleotides can tightly bind to DNA or RNA, but they are insoluble in aqueous solutions and therefore are modified with a positively charged group at the 3′-end of the sequence. Despite the existing problems, these examples of antisense sequence development have provided valuable information for understanding the chemical structure−function relationship of antisense oligonucleotides. There have been increasing efforts to characterize and determine the structures of backbone-modified antisense duplexes, such as those containing phosphorothioate (Cruse et al., 1986; Williams et al., 1990; Gonzalez et al., 1995),

† This research is supported by the NIH (R29 GM49957-01) and the W. M. Keck Center for Computational Biology (a fellowship training grant to C.W.C.).

* To whom correspondence should be addressed. Tel: 713 743-2805. Fax: 713 743-2709. E-mail: gao@kitten.chem.uh.edu.

‡ Current address: Sealy Center for Structural Biology, Department of Human Biological Chemistry and Genetics, University of Texas Medical Branch, Galveston, TX 77555-1157.

⊗ Abstract published in *Advance ACS Abstracts,* March 15, 1997.

S0006-2960(96)02382-3 CCC: $14.00

phosphorodithioate (Piotto et al., 1991), 3′-methylene phosphonate (Heinemann et al., 1991), (N3′-P)-phosphoramidate (Ding et al., 1996), amide (Blommers et al., 1994; Fritsch et al., 1995), 3′-methylene-*N*-methylhydroxyamino or methylene methylimino (MMI, Mohan et al., 1995), sulfone (Egli, 1996), and *N*-(2-aminoethyl)peptide (PNA) (Brown et al., 1994; Eriksson & Nielsen, 1996) linkages.

Our studies of antisense sequences have focused on a class of nonionic, achiral backbone linker substitutions (Gao et al., 1992; Gao & Jeffs, 1994a,b; Veal et al., 1993; Veal & Brown, 1995), which include formacetal (FMA,[1] O3′−CH_2−O5′) (Matteucci, 1990) and 3′-thioformacetal (3′-TFMA or TFMA, S3′−CH_2−O5′) (Jones et al., 1993) groups as replacements for natural phosphodiester (O3′−PO_2−O5′) linkages. These neutral, achiral linkers are attractive because they can pass through the cellular membrane more easily than negatively charged phosphodiester linkers and are resistant to nucleases. Moreover, oligodeoxyribonucleotides containing FMA or 3′-TFMA modifications form stable duplexes with complementary DNA or RNA strands. A hybrid duplex containing multiple 3′-TFMA substitutions showed a higher T_m than an unmodified duplex of the same base sequence (Jones et al., 1993). To understand the molecular basis of the improved antisense properties of the 3′-TFMA-containing oligonucleotides, and to extend the range of the molecular design for this class of backbone modifications, structural and conformational information concerning these antisense oligonucleotides is indispensable. The TFMA linker contains a sulfur atom in place of the 3′-oxygen atom and a methylene group in place of the charged PO_2 moiety. Due to the longer C−S−C bond lengths compared to those of the C−O−P bonds (1.82 and 1.82 Å *versus* 1.44 and 1.62 Å) and larger van der Waals radius of the sulfur atom as compared to oxygen (1.09 *versus* 0.65 Å), the local conformation of the linker in the DNA strand must make adjustments in order to accommodate the substitution. In a previous NMR study, we demonstrated that the 3′-TFMA linker in a DNA·DNA duplex adopts a conformation that is distinctly different from a normal phosphodiester linkage (Gao *et al.,* 1994a). Recent NMR studies of the TFMA linker in an RNA·DNA hybrid duplex reveal yet another conformation for this linker (Rice & Gao, 1997).

In addition to the local conformation of the substituted backbone, another concern in antisense oligonucleotide applications is the overall structure of the hybrid duplex, because such features may hold the key to the ribonuclease cleavage of hybrid duplexes *in vivo*. Because the RNA and DNA strands have different properties, one or both of the strands in the hybrid duplex will change their conformation in order to form a stable complex. Most recent reports have suggested that the RNA·DNA duplex structure is heterogeneous, having the RNA strand in an A-like form and the DNA strand in an intermediate A- and B-form (Arnott *et al.*, 1986; Benevides & Thomas, 1988; Gao & Jeffs, 1994b; Gonzalez et al., 1995; Katahira *et al.,* 1990; Reid *et al.*, 1993; Rice & Gao, 1997).

The three-dimensional solution structure of RIII, a hybrid RNA·DNA duplex comprised of r(GCGCAAAACGCG) and d(CGCGTT-SCH_2O-TTGCGC) strands, has been determined in this laboratory by simulated annealing and NMR restrained molecular dynamics (rMD) (Brunger & Karplus, 1991; Nilsson et al., 1986) followed by relaxation matrix (RM) refinement based on volume restraints (Nilges, et al., 1991). By analyzing calculated structures and comparing the results with the existing information on related duplexes, we investigate the global features of the modified hybrid duplex, the specific details of the individual RNA and DNA strands and the RNA A-tract, the conformation of the linker backbone, and the local conformational perturbations caused by the accommodation of the linker.

MATERIALS AND METHODS

The RIII duplex and the related NMR data used in the calculation have been described in a separate article from this laboratory (Rice & Gao, 1997).

Distance and Dihedral Angle Restraints. The H2′,2″−H1′ coupling constants and the qualitative analysis of H2′,2″−H3′ and H3′−H4′ coupling cross peaks from the COSY-35 and DQF-COSY spectra were used in combination to derive pseudorotation angles (Saegner, 1984; Chary et al., 1987) for each of the sugar rings in RIII. The pseudorotation angles were then used to generate dihedral angle restraints for the υ_1 (torsion of C1′−C2′) and υ_3 (torsion of C3′−C4′) angles (Saenger, 1984). Error bounds used for the 48 torsion angle restraints were ±10°. Distance restraints were from three NOESY spectra for nonexchangeable protons recorded at 25 °C with mixing times of 70, 140, and 200 ms and NOESY spectra for exchangeable protons recorded at 15 °C with a mixing time of 150 ms. The 2D cross peaks from the NOESY spectra were integrated using FELIX 2.30 (Molecular Simulations, Inc.) and their assignments completed using a grid search of the FELIX cross peak database files (software developed in this laboratory). The integration yielded 35 volumes from the 150 ms exchangeable NOESY and 234, 288, and 284 volumes from the 70, 140 and 200 ms nonexchangeable NOESY spectra, respectively (volumes from fixed distances, such as those connecting H2′ and H2″ or base H5 and H6, are excluded). These volumes are summarized in Tables S1A and S1B, respectively, of Supporting Information. The TFMA linker has 19 distance restraints which are directly associated with the stereospecifically assigned HP1 (*pro-S*) and HP2 (*pro-R*) protons. These restraints are in addition to the normal T6 and T7 inter- and intraresidue restraints, allowing for a well-defined linker region.

The three sets of volumes from the nonexchangeable proton NOESY spectra were converted to distances (Table S1C) using the MARDIGRAS program (Borgias & James, 1989; Keepers & James, 1984). For each volume set, three distance sets were computed using isotropic correlation times of 1.5, 2.5, and 3.5 ns and an A-form model structure. This procedure yielded a total of nine sets of distances, which were then statistically averaged to yield a total of 335 distance restraints. The MARDIGRAS distances were divided into three subsets on the basis of the standard deviations (0.3 Å, 0.3−0.6 Å, and greater than 0.6 Å). The standard deviations of the distances were multiplied by a

[1] Abbreviations: 1D, one dimensional; 2D, two dimensional; FMA, formacetal; 3′-TFMA, 3′-thioformacetal; COSY, correlation spectroscopy; DQF, double quantum filtered; g, gauche; t, trans; ISPA, isolated spin pair approximation; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect and exchange spectroscopy; RM, relaxation matrix; RMSD, root-mean-square deviation; rMD, restrained molecular dynamics; T_m, melting temperature; TOCSY, total correlation spectroscopy; UV, ultraviolet; VDW, van der Waals.

Table 1: Summary of Structure Calculation Results[a]

		initial	post MARDIGRAS	post RM
coordinate convergence	RMSD (Å)	3.10 ± 0.50	0.50 ± 0.10	0.42 ± 0.09
NOE intensity restraints	$R^{1/6}$-factor	0.160 ± 0.020	0.078 ± 0.001	0.043 ± 0.001
	E(VDW) (kJ/mol)	bond length (Å)	bond angle (deg)	DIHE restraint (deg)[b]
ideal geometry	−78 ± 1	0.00041 ± 0.00001	2.53 ± 0.01	0.31 ± 0.02

[a] RMSD are pairwise comparisons of the eight structures. Values reported are the averages and the statistical deviations of the eight structures. [b] Averaged violation of the dihedral angle restraints derived from NMR analysis for the final eight structures.

scaling factor of 7.0 to give the initial error bounds. NOE force constants of 100, 50, and 25 kcal/(mol Å^2) were assigned to the three groups of distance restraints, respectively, to give the best data subset the most weight. Exchangeable NOE cross peak volumes were integrated using FELIX and converted to distances using the isolated spin pair approximation (ISPA). Error bounds were initially ±50% and were adjusted in later dynamic runs. The distance restraints used for A·T and C·G Watson−Crick hydrogen bonding include 35 restraints derived from NOEs and 32 restraints between the donor and acceptor atoms of nucleobases [three per C·G and two per A·T base pair; parameters were from Saenger (1984)] for maintaining base pairing during the calculations. These restraints were given a force constant of 200 kcal/(mol Å^2). Since no cross peaks for interstrand contacts between sugar protons were detected, repulsive distance restraints were applied at a later stage of calculations to separate these protons by greater than 5.5 Å to prevent the minor groove from collapsing.

Computational Details. All calculations were performed using the X-PLOR program (Brunger, 1992). Initial structures were obtained by free molecular dynamics (MD) simulations of ideal A- and B-form duplexes generated from the QUANTA program (Molecular Simulations, Inc.). The process was initiated by assigning each structure a randomly chosen seed number, which sets the velocity for MD calculations. The simulation involved Powell energy minimization of the input structure for 200 steps and then Verlet MD simulations with temperatures increased from −73 (200 K) to 523 °C (800 K) over 8.0 ps. The structure was then equilibrated at 523 °C for an additional 4.0 ps and minimized for 400 steps. Ten randomized initial structures were generated, five from A-form and five from B-form, and their coordinate RMSD was 3.10 ± 0.50 Å (Table 1).

Structure calculations were performed iteratively using distance and dihedral angle rMD methods and RM refinement (Brunger & Karplus, 1991; Nilges et al., 1991; Nilsson et al., 1986). The final restraint set contains 450 entries including those defining the interproton separations, Watson−Crick base pairs, and backbone and sugar dihedral angles. A CHARMM (Brooks et al., 1983) based force field modified by J. Veal (Veal et al., 1995) to optimize sugar pucker refinement was used throughout. All electrostatic interactions were turned off because these terms are not accurate in the absence of water and are not critical in NMR-based structure refinement. A 9.0 Å nonbond interaction cutoff radius was used. Square-well penalty functions were used for the NOE distance restraints and the dihedral angle restraints, while biharmonic penalty functions were used for the Watson−Crick base pair distance restraints. For the RM refinement, a square-well relaxation energy term was appended to the total potential energy for the calculation of the gradient. A 1.0 fs time step was used in all calculations.

Each of the ten starting duplexes was energy minimized without restraints for 200 steps. The NOE force constants were then ramped from 0.0 to their maximum levels [100, 50, and 20 kcal/(mol Å^2) based on statistical deviations, *vide infra*] over 50 fs. The structures were then heated from −173 (100 K) to 27 °C (300 K) over 4.0 ps, equilibrated at 27 °C for 2.0 ps, and then cooled to −173 °C over 4.0 ps. After the completion of the first simulated annealing cycle, two structures (one from A- and one from B-form structure) were excluded because they became trapped in local minima and had unreasonable geometries, resulting in a total of eight first generation structures.

After each annealing cycle, a program developed in this laboratory was used to calculate the distances between each proton pair for which there was a restraint. The average distances of the eight structures and their standard deviations were also computed. The program then compared the average distance, plus or minus the standard deviation, to the target distance and the existing error bounds. For those averaged distances which fall within the given restraints, the original distance bounds were narrowed to give a set of improved distance restraints. These comparisons also permitted identification of distances with large deviations. These distances were then examined against the original experimental data for potential errors in the initial data input. The rMD simulated annealing, distance comparison, data reexamination, and restraint revision procedures were repeated for a total of six iterations. During these iterations, the calculated distances were also compared with the MARDIGRAS distances to verify the accuracy of the assignments. These comparisons allowed us to correct errors and to add additional restraints.

The resulting structures were further refined using complete RM-rMD simulations and a new generation of NOE volume restraints. Potential errors in these volume restraints were eliminated by careful examination against MARDIGRAS distances using the latest generation of the averaged structure. These "cleaned up" volumes were used to simultaneously compute the gradient of the relaxation energy for each of the four (three nonexchangeable and one exchangeable) NOESY spectra. Watson−Crick base pair distance restraints and dihedral angle restraints for sugar ring puckers were also included in the calculation. A force constant of 500 kcal/mol was used for the relaxation energy, while the force constant for the Watson−Crick distance restraints was 200 kcal/(mol Å^2). A mass of 100 Da was assigned to all atoms in order to suppress high-frequency motions within the structure. An isotropic correlation time, τ_c, of 2.5 ns was determined by using X-PLOR to perform a grid search to minimize the RMSD of the average structure with respect to τ_c. Order parameters, which are indicative of local dynamics, were qualitatively assigned, according to the type of protons, to backbone, sugar, or base groups

(groups 1, 2, and 3, respectively). These values are vectors between atoms of group 1 and atoms of group 1 (1,1 interactions) = 0.64, groups 2,2 = 0.75, groups 3,3 = 0.85, groups 1,2 = 0.70, groups 1,3 = 0.75, and groups 2,3 = 0.80. The calculations were, in fact, found to be insensitive to the values of the order parameters within the chosen range. The initial RM refinement involves energy minimization, for 25 time steps, of the eight rMD simulated structures without restraints. The force constants for volume, distance, and dihedral restraints were then increased from 0.0 to their set values over 100 fs. The structures were annealed by heating from −173 to 27 °C (100−300 K) over a time period of 1.0 ps and then cooling to −173 °C over 9.0 ps. The eight converged structures from this calculation were energy minimized to give the reported results.

The calculated structures of RIII show satisfactory bonding geometry and a low nonbonding energy and are in good agreement with the experimental data as judged by low NOE and experimental dihedral angle violations (Table 1). Table 1 lists statistics for coordinate and ideal geometry convergence and NOE $R^{1/6}$-factors (Nilges et al., 1991) for the final eight structures at various points in the calculation. The RMSD results were calculated for all atoms with respect to an average structure. The $R^{1/6}$-factors were calculated for all NOE volumes and are averaged over the eight final structures. Spectral simulations using NMRCompass/QUANTA were used to visually inspect the overall quality of these structures. The overall structure of the RIII duplex was analyzed on the basis of a set of parameters describing the helical geometry and relative orientation of each residue using the CURVES program (Lavery & Sklenar, 1988). Helical parameters were also computed for canonical A- and B-form duplexes for purposes of comparison.

RESULTS

Using NMR we have studied a set of antisense duplexes which differ in a single backbone linker in the center of the DNA strand. The linkers include phosphodiester, FMA, 3′-TFMA, and MMI (Vasseur et al., 1992) groups. These studies have permitted a systematic examination of the conformation and the conformational preferences of backbone linker modifications. Here we report the high-resolution structure of the RIII duplex (Figure 1), which is the first structure determined in this series of antisense duplexes.

Computational Results. The spectra simulated from a representative structure of RIII and the experimental spectra are displayed in panels A and B of Figure 2, respectively. The low nonbonding energy and the agreement of molecular geometry between the calculated structures and the expected experimental values are given in Table 1. The final eight structures converge well, as reflected by the 0.42 ± 0.09 Å RMSD based on comparison of the coordinates of these structures with respect to those of an average structure (Table 1). The $R^{1/6}$-factor (Table 1), which measures the deviation of the calculated spin intensity matrix from experimental restraints is 0.043 ± 0.001, indicating an overall good agreement between the two data sets. The quality of the calculated structures is judged not only by the empirical energy terms and numerical statistics but also by spectral back-calculation (Figure 2). An excellent agreement between the calculated and experimental spectra has been achieved. Spectral comparisons in important regions containing H1′ to base proton correlations and linker to sugar proton

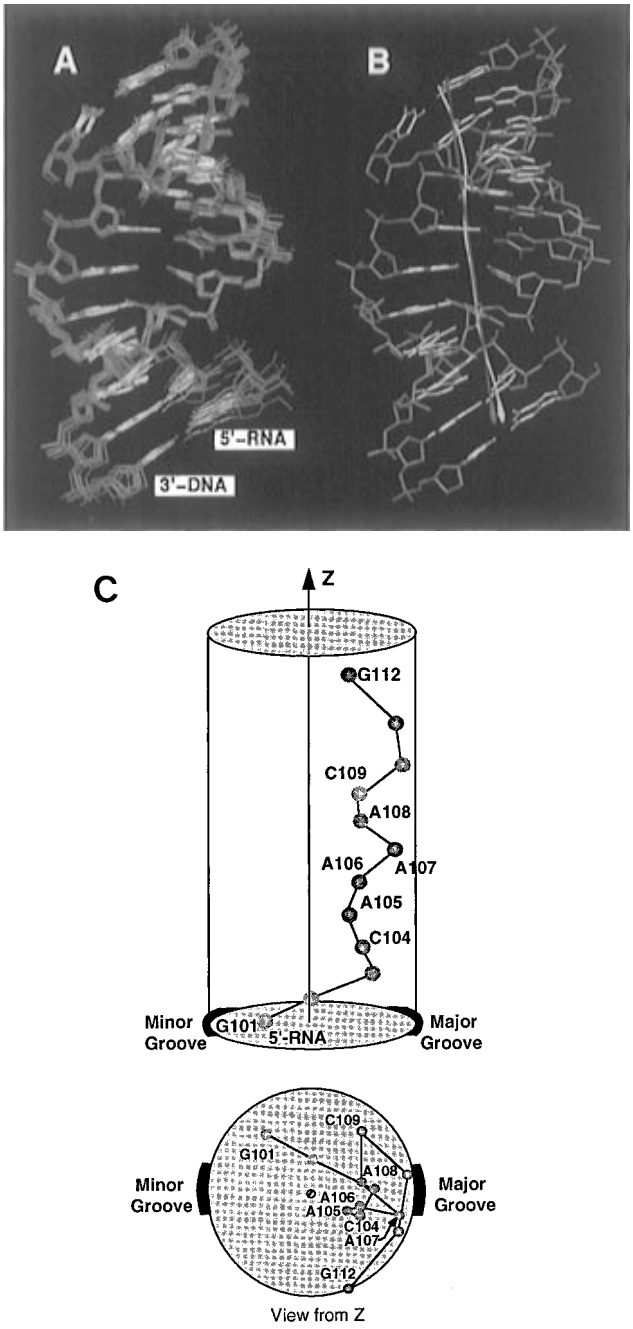

FIGURE 1: Structural drawings of the RIII duplex (a view into the minor groove), 5′-d(CGCGTT-*-TTGCGC)·r(GCGCAAAACGCG)-3′ (* is S3′-CH_2-O5′; the DNA strand is numbered from the 5′-end, from 1 through 12 and the RNA strand from 101 to 112). Backbones are drawn in magenta and bases in white. T residues of the DNA strand are colored in blue, while the TFMA linker is in green. For clarity of display, hydrogen atoms are not shown. (A) Overlay of the entire family of eight structures. (B) Average of the eight calculated structures showing best fit (red) and local (yellow) helical axes. (C) Schematic drawing of the projection of the helical axes generated using the ORIGIN program and translated coordinates (*x*, *y*, 0). Each point specifies the position of a base pair, taken from the relative position of C1′, with respect to the axis drawn. The light balls represent the residues into the paper plane, and the darker balls represent the residues out of the paper plane. The drawing shows views projected into *yz* or *xy* planes. The relative positioning of the sequentially connected residues indicates the direction and amplitude of the local helical axis bending (drawings are not proportional in three dimensions; rather they demonstrate relative movement as a function of sequence). The relative orientation typically is described by bending toward the minor groove (if a residue is located relatively on the left of adjacent residues) or the major groove (the *vice versa*). The local helical axis also indicates bending modes, such as that observed at the C109 step, which do not cause compression of either groove.

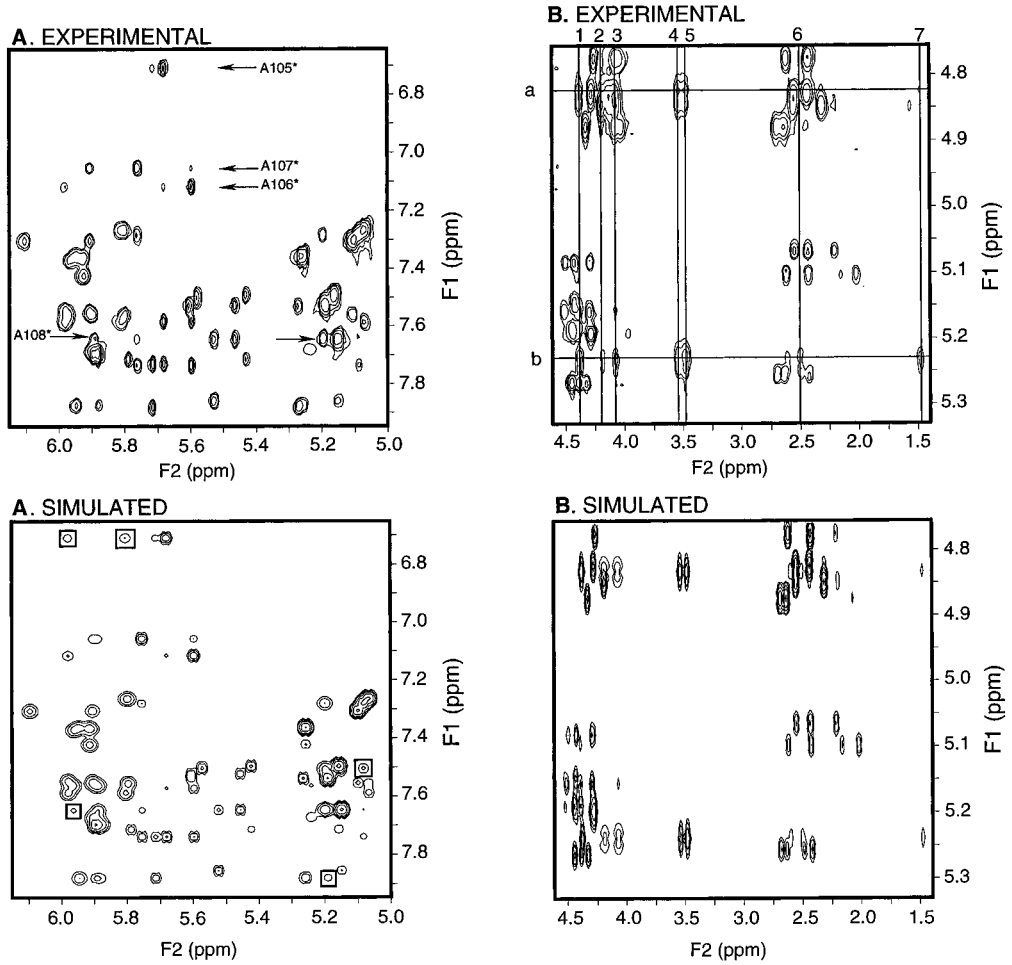

FIGURE 2: Experimental and simulated 200 ms NOESY spectra. (A) H1′−base proton (F_2−F_1) spectral region. Spectral simulation uses a representative structure of the eight final structures as input for the calculation using QUANTA/NMRcompass (Molecular Simulations, Inc.). Arrows indicate A H2 resonances. Cross peaks enclosed with a box originate from A H2 peaks or C H5 resonances and do not appear in the experimental spectra. These discrepancies may be associated with longer than average T_1 values which potentially can cause signal attenuation in NMR spectra or imperfection of simulated spectra. (B) Linker to sugar interproton NOE connectivities. Important proton assignments are indicated in the figure by solid lines, which are those in parallel with the F_2 axis (a, HP1; b, HP2) or those in parallel with the F_1 axis (1, T7 H5″; 2, T7 H4′; 3, T6 H4′; 4, T7 H5′; 5, T6 H3′; 6, T6 H2′; 7, T7 methyl).

correlations are shown in panels A and B of Figure 2, respectively. The characterization of the double helix of RIII and its minor groove configurations is summarized in Tables 2 and 3 and Figure 3.

Our calculation has taken the advantage of using distances derived from complete relaxation matrix implemented in the MARDIGRAS program (Borgias & James, 1989; Keepers & James, 1984). In addition to offering a more theoretically sound approach to process NOE information, these calculations have practical advantages as well. The resonance assignments, cross peak integration, and other experimental or human errors are potential sources of poor quality data. These errors should be screened using MARDIGRAS or similar programs before the rMD simulation is started. The data output from MARDIGRAS provides important statistical information on the quality of the cross peaks in various data sets, thus prompting verification of resonance and cross peak assignments for those which show large statistical deviations among the data sets of different mixing times. The program permits the combined use of all available data and can be used to weigh distances according to their statistical significance. These features are particularly useful for examination of initial distances because a few ill-defined distances at the beginning stage can be detrimental to the entire structure generation process. The dependence of MARDIGRAS upon the model structure does present limitations, however, which may be alleviated by applying the program after generation of improved structures.

Global Helical Structure. The helical conformation of RIII is characterized by a set of global helical parameters generated using the CURVES program (Figures 1C and 3 and Table 2) (Saenger, 1984; Lavery & Sklenar, 1988). In Figure 3, the major structural parameters are displayed for each strand and for the duplex, from which the conformations of the two strands separately and the structural characteristics of the duplex can be assessed. Although limited by the available NMR information due to lack of protons or spectral resolution for nucleobases and backbone regions, the collective pattern and trend of the structural parameters shown in Figure 3 and Table 2 provide a qualitative description that permits a better understanding of the structural information concerning RIII.

RIII is characterized by two complementary strands in heteromerous helical conformations (Table 2 and Figures 3 and 4). The RNA strand shows several typical A-form features, such as the base displacement along the *X*-axis of the helix, the inclination of base planes with respect to the helical axis, and the interbase rise (Figure 3a,b,g). But atypical characteristics of A-form are also detected in the RNA strand, such as large negative tip angles (Figures 3c and 4A,B). The conformation of the DNA strand is clearly heterogeneous, being present in a range of noncanonical

Table 2: Sugar Pucker and Backbone Conformation[a]

residue	phase[b]	glycosidic	α	β	γ	δ	ε	ζ
C1	80 ± 1	−165 ± 3			63 ± 1	95 ± 1	−132 ± 3	−60 ± 2
G2	111 ± 1	−119 ± 2	−93 ± 1	170 ± 1	53 ± 1	106 ± 1	−179 ± 2	−92 ± 1
C3	97 ± 1	−141 ± 1	−72 ± 2	171 ± 1	57 ± 2	97 ± 1	−166 ± 1	−72 ± 1
G4	94 ± 1	−143 ± 2	−88 ± 1	177 ± 1	56 ± 1	97 ± 1	−174 ± 1	−80 ± 2
T5	95 ± 3	−139 ± 3	−73 ± 1	173 ± 1	59 ± 1	97 ± 1	−173 ± 2	−86 ± 1
T6	**82 ± 1**	−145 ± 2	−71 ± 1	162 ± 1	73 ± 1	**87 ± 1**	**−154± 1**	**−79± 1**
T7	88 ± 1	−138 ± 2	**−69 ± 1**	**159 ± 1**	**76 ± 1**	96 ± 1	−179 ± 1	−75 ± 1
T8	97 ± 1	−125 ± 2	−86 ± 1	−174 ± 1	63 ± 1	107 ± 1	169 ± 1	−70 ± 3
G9	102 ± 1	−139 ± 5	−82 ± 2	174 ± 2	56 ± 3	100 ± 1	−175 ± 2	−83 ± 2
C10	105 ± 1	−136 ± 3	−73 ± 2	171 ± 1	60 ± 1	106 ± 1	−167 ± 2	−80 ± 2
G11	103 ± 1	134 ± 1	−81 ± 3	171 ± 1	60 ± 1	101 ± 1	−177 ± 1	−85 ± 1
C12	98 ± 1	−135 ± 3	−69 ± 1	168 ± 1	60 ± 1	102 ± 9		
G101	54 ± 3	−166 ± 2			56 ± 3	81 ± 2	−172 ± 2	−76 ± 3
C102	9 ± 1	−160 ± 2	−82 ± 1	177 ± 1	63 ± 1	78 ± 2	−173 ± 3	−61 ± 3
G103	15 ± 7	−155 ± 3	−81 ± 3	173 ± 1	65 ± 2	86 ± 4	−175 ± 1	−54 ± 1
C104	56 ± 1	−156 ± 2	−107 ± 2	−167 ± 3	66 ± 1	93 ± 2	−174 ± 1	−72 ± 3
A105	38 ± 10	−148 ± 3	−88 ± 2	−173 ± 1	72 ± 5	84 ± 3	−166 ± 3	−59 ± 2
A106	31 ± 5	−168 ± 2	−90 ± 1	167 ± 1	73 ± 2	65 ± 1	−163 ± 3	−58 ± 3
A107	22 ± 8	−140 ± 2	−87 ± 5	−167 ± 1	73 ± 6	85 ± 3	−176 ± 1	−64 ± 3
A108	15 ± 1	−163 ± 3	−84 ± 2	178 ± 1	49 ± 2	71 ± 1	−174 ± 1	−57 ± 1
C109	20 ± 1	−156 ± 1	−86 ± 1	−177 ± 1	68 ± 1	74 ± 1	−162 ± 1	−63 ± 1
G110	20 ± 1	−178 ± 2	−79 ± 3	169 ± 1	72 ± 2	76 ± 1	−161 ± 3	−61 ± 5
C111	28 ± 5	−171 ± 2	−87 ± 3	179 ± 4	70 ± 3	75 ± 2	−168 ± 4	−50 ± 4
G112	64 ± 7	−158 ± 3	−83 ± 2	−179 ± 1	71 ± 3	80 ± 6		
A-form	14	−153	−88	−149	47	83	171	−44
B-form	192	−95	−39	−151	31	156	159	−99

[a] Mean and deviation values of the structural parameters were derived from the final eight structures of RIII using CURVES. Phase and glycosidic angles are as defined in the literature (Saegner, 1984). Backbone dihedral angles are as follows: α, O3′−P−O5′−C5′; β, P−O5′−C5′−C4′; γ, O5′−C5′−C4′−C3′; δ, C5′−C4′−C3′−O3′; ε, C4′−C3′−O3′−P; ζ, C3′−O3′−P−O5′. Ideal values for the A- and B-forms are from QUANTA structures. [b] Sugar pucker−phase angle correlation: phase angle 0−36°, C3′-endo; 37−72°, C4′-exo; 73−108°, O4′-endo; 109−144°, C1′-exo; 145−180°, C2′-endo. Bold lettering indicates linker positions.

Table 3: Distances Involving the Minor Groove A H2[a]

	A105	A106	A107	A108	A-form	B-form
1. A H2 to adjacent A H2	$4.56^{A} \pm 0.08$	$4.42^{A/B} \pm 0.09$	$4.45^{A/B} \pm 0.13$	N/A	4.87	3.78
2. A H2 to cross strand 5′-imino H	$6.16^{A} \pm 0.06$	$6.11^{A} \pm 0.11$	$5.84^{A} \pm 0.17$	$4.94^{B} \pm 0.25$	5.89	4.99
3. A H2 to cross strand 3′-imino H	$4.66^{AB} \pm 0.06$	$3.26^{X} \pm 0.06$	$3.25^{X} \pm 0.05$	$3.58^{X} \pm 0.14$	5.32	4.29
4. A H2 to its own H1′	$4.63^{A/B} \pm 0.01$	$4.72^{A} \pm 0.01$	$4.55^{A/B} \pm 0.03$	$4.67^{A/B} \pm 0.02$	4.71	4.49
5. A H2 to 3′-H1′	$3.08^{X} \pm 0.05$	$3.19^{X} \pm 0.12$	$3.61^{A} \pm 0.18$	$3.60^{A} \pm 0.11$	3.64	4.55
6. A H2 to cross strand 3′-H1′	$4.25^{A} \pm 0.09$	$4.21^{A} \pm 0.07$	$4.42^{A} \pm 0.17$	$3.44^{X} \pm 0.08$	4.47	5.01

[a] Entries in lines 1, 4, and 5 are NOEs within the RNA strands and in lines 2, 3, and 6 are NOEs between RNA and DNA strands. Superscript letters: A, closer to A-form than to B-form; B, closer to B-form than to A-form; AB, in between A- and B-form; A/B, values for A- and B-forms are similar and distances cannot differentiate one form from the other; X, non-A- and non-B-form. N/A, not available.

forms. The deviations of the conformations of either of the strands from ideal A- or B-form values suggest that both DNA and RNA strands have undergone structural changes in order to hybridize. A notable difference between the bases of DNA and RNA strands is shown in their relative orientation along the base plane long axis with respect to the helical axis (*i.e.*, inclination angles, Figure 3b). The bases of the DNA strand show more positive inclination angles compared to those of the RNA strand. The magnitude of the fluctuations of the values for several helical parameters also differs considerably in the DNA and RNA strands. This can be seen in the tip angles for individual bases (Figure 3c) and the corresponding roll angles between sequential bases (Figure 3m). For DNA or RNA strands, the pattern of parameter variation as a function of sequence is also different. The conformational parameters of the consecutive residues in the DNA strand are often opposite to each other, resulting in zigzag patterns as shown in Figure 3a−j. In comparison, some of the conformational variations in the RNA strand are isolated, and the general trend is as shown in Figure 3c,g,i. This indicates that the DNA strand in the hybrid duplex is more flexible and adaptive to structural variations, whereas the RNA strand is conformationally more dominant. These DNA and RNA comparisons show that although the structural heterogeneity of the hybrid duplex has been described in general terms in several unmodified as well as in chiral phosphorothioate hybrid duplexes (Fedoroff et al., 1993; Gonzalez et al., 1995; Salazar et al., 1993b), the separate analysis of DNA and RNA strands in RIII has presented a much more detailed view of the microscopic properties of the hybrid duplex.

Figure 3d−f and Figure 3k−n are displays of the structural parameter measurements based on the axis defined by the elements from the residues of both strands in the duplex. Consequently, these parameters tend to reflect averaged properties of the two single strands. For instance, as shown in Figure 3b,e, the DNA residues have mostly larger and RNA residues have mostly smaller than 22° inclination angles, but the duplex inclination angles are about 22° and are very close to that of the A-form. Other structural parameters of the duplex also fall into the range of an A-form helix. However, locally, the relative orientations of the bases, such as buckle and propeller twist angles (Figure 3o,p), in the same Watson−Crick base pair are not restricted to either

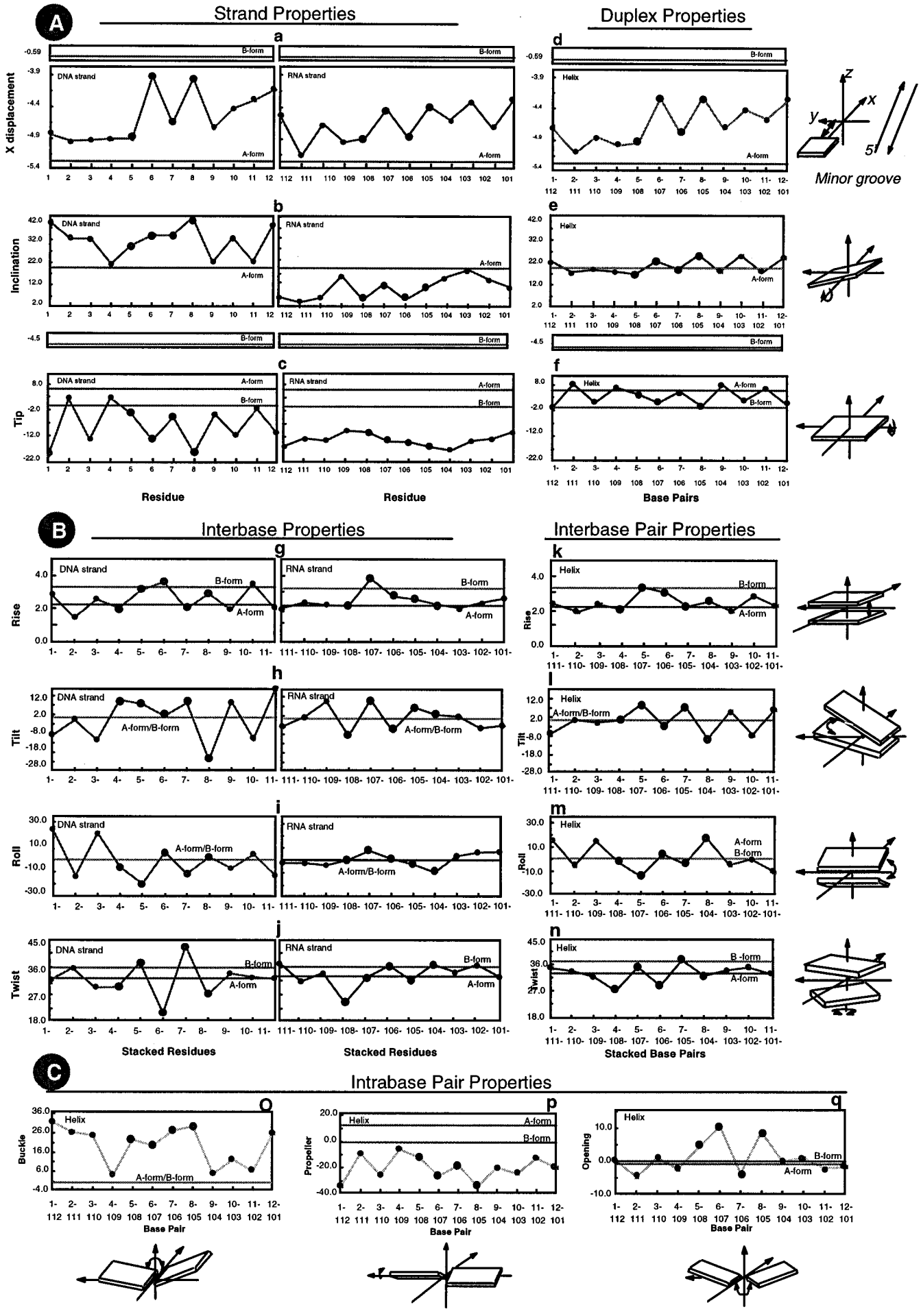
A
Strand Properties
Duplex Properties
a
B-form
DNA strand
RNA strand
A-form
X displacement
d
Helix
z
x
y
5'
Minor groove
b
Inclination
e
c
Tip
f
Residue
Residue
Base Pairs
B
Interbase Properties
Interbase Pair Properties
g
Rise
k
h
Tilt
A-form/B-form
l
i
Roll
m
j
Twist
n
Stacked Residues
Stacked Residues
Stacked Base Pairs
C
Intrabase Pair Properties
o
Buckle
p
Propeller
q
Opening
Base Pair
Base Pair
Base Pair

FIGURE 3: Graphical representation of helical parameters of RIII computed by the CURVES program (all based on the global axis) for the eight final structures, summarized in average values, and RMSD (RMSD not shown). Values expected from ideal A- or B-form helices are also indicated in the figures. The helical parameters are shown in three groups: (A) parameters in relation to helical axis for either of the two single strands (a−c) or the duplex (d−f); (B) parameters defining the relative orientation of the two bases which are hydrogen bonded (o−q). Charts a−c and g−j characterize separately the conformation of the DNA and the RNA strands, while charts d−f and k−n reflect the structural features of the duplex as a whole. The figures for the DNA strand are plotted from the 5′- to 3′-end, and those for the RNA strand are plotted from the 3′- to 5′-end. This arrangement gives a directional correlation of the plots for single strands with those for the duplex.

family of canonical helical parameters. There is no major helical disruption at the TFMA-modified T6 and T7 residues, although local variations are discernible (*vide infra*). Overall, the RIII duplex is A-form like as judged by the distribution of individual helical parameters, the majority of which are closer to A-form than to B-form. However, deviations from both types of canonical helical structures, in some portions of the helix, are evident. An additional indication of the hybrid nature of the RIII duplex is the width of the minor groove (7−9 Å), which lies between the values calculated for ideal A- (10−11 Å) and B-form (5−6 Å) duplexes (Saegner, 1984).

Sugar Pucker and Backbone Conformation. Sugar pucker conformations are characterized by torsion angles, δ, and pseudorotation phase angles (Saenger, 1984), which were calculated for all eight structures (Table 2). It should be noted these values were generated using the single restraint per torsion angle approach of rMD (conventional rMD) and, thus, have their limitations. As shown by Gonzalez et al. (1995) using molecular dynamic simulations based on time-averaged restraints, the sugar conformations thus derived are close to their conformational equilibrium when sugars are in typical C2′- or C3′-endo conformations. An intermediate sugar pucker (pseudorotational angle $\sim$ 65°) derived from conventional rMD may correlate to an increased population of C3′-endo for a DNA sugar originally in a C2′-endo pucker. In this study, the phase angles for the RNA strand sugar rings are less than 60°, consistent with the C3′-endo and C4′-exo families of conformations. The phase angles of DNA sugar rings (80−111°, Table 2), on the other hand, are intermediate to those of typical C3′- and C2′-endo forms. While these values indicate more C2′-endo form influence in the DNA strand than in the RNA strand, they may imply the presence of a family of sugar conformations with considerable contribution from C3′-endo pucker. This would be consistent with our observation of weak H1′−H4′ NOEs (70 ms mixing time) for most of the DNA sugar moieties except for the linker residues and the comparable H1′−H2′ and H1′−H2″ coupling constants ($J_{1'-2'}$ and $J_{1'-2''}$) and other *J* constant analysis (Rice & Gao, 1997). The H1′−H4′ NOEs were not observed in the RNA or T6 and T7 residues and are diagnostic of an O4′-endo type of conformation (Wuthrich, 1986). The exact nature of the DNA sugar conformational equilibrium cannot be determined at this time due to the limitations of the rMD method used. Except for the T6 and T7 linker sugars, the overall characteristics of RIII sugar moieties derived from rMD are similar to what have been reported for other hybrid duplexes (Arnott et al., 1986; Salazar et al., 1993a,b), while the detailed interpretation of these results remains controversial.

The statistical distribution of the backbone torsion angles of RIII is summarized in Table 2. The observed phosphodiester backbone conformation of RIII is the result of nonrestrained dynamics simulations at most stages of the calculation, during which only C3′−C4′ dihedral angle restraints derived from NMR data were applied. In the final refinement, additional backbone dihedral angle restraints were applied. These restraints include those based on some qualitative NOE information, such as the absence of base proton to H5′ and H5″ NOEs, as discussed by Reid and coworkers (Kim et al., 1992), and those based on the statistical distributions of backbone angles among the cluster of structures. The final refinement using additional backbone restraints helped the convergence of the calculation results and had marginal effects on local conformations. Due to the uncertainty in the quantitative aspect of the backbone parameters determined by NMR, here we treat these results on a general basis and emphasize the importance of the relative values rather than the accuracy of the absolute values. The results shown in Table 2 indicate that the phosphodiester backbone conformations of both strands are close to those adopted by canonical A- or B-form helices. The glycosidic angle measurements show that RNA residues are in the relatively high-anti region (−140 to −170°, average = −160°) compared to −153° for A-form RNA. The majority of the DNA residues are in the −128 to −150° range (average = −134°). This averaged value is neither A- nor B-form-like (glycosidic angle for A-form = −153° and for B-form = −95°).

Local Conformation of the A4-Tract. Helical bending is an important global feature and is discernible from the perturbations in the roll, tilt, and propeller twist angles of base pairs. Bending has been found within the DNA A-tract or at the A-tract junctions (Alexeev et al., 1987; Edmondson & Johnson, 1985; Koo et al., 1986; Nadeau & Crothers, 1989; Nelson et al., 1987; Young et al., 1995). Figure 1B shows least squares fitted, averaged axis and the local axes of RIII, and the enlarged projections of these axes into the *xy* plane are shown in Figure 1C. The position of each residue in the projected plane contains information on the relative orientation of each residue with respect to the global helical axis. These displays indicate that the overall positioning of the helix is shifted toward the major groove, but the two terminal residues are more subject to computational artifacts. In addition, there is a local bending toward the minor groove, starting from C104 and continuing to A106. The A107 residue bends in an opposite direction toward the major groove, but A108 reverses the direction orientation to align with the A105 and A106 residues. The observed bend winds around the helix with the RNA 5′-end (C102-G103) being approximately 9−14° and the 3′-end (C109-G110) approximately 8−23° (depending on the direction of view). Both bends are observed to point into the major groove. The tilt, roll, and propeller twist angles within the A-tract are displayed in Figures 3l,m,p and 4C. The tilt angles of the A105 to A108 residues gradually change from −10° to positive +9° with the exception of A107, which is +7° (Figure 3l). A similar trend is also observed for roll and propeller parameters (Figure 3m,p), which show a consistent pattern for the A-tract residues with the exception of A107. If A107 is excluded, roll angles within the A-tract consistently vary from +17 to −15° and propeller twistangles vary from −34 to −13°.

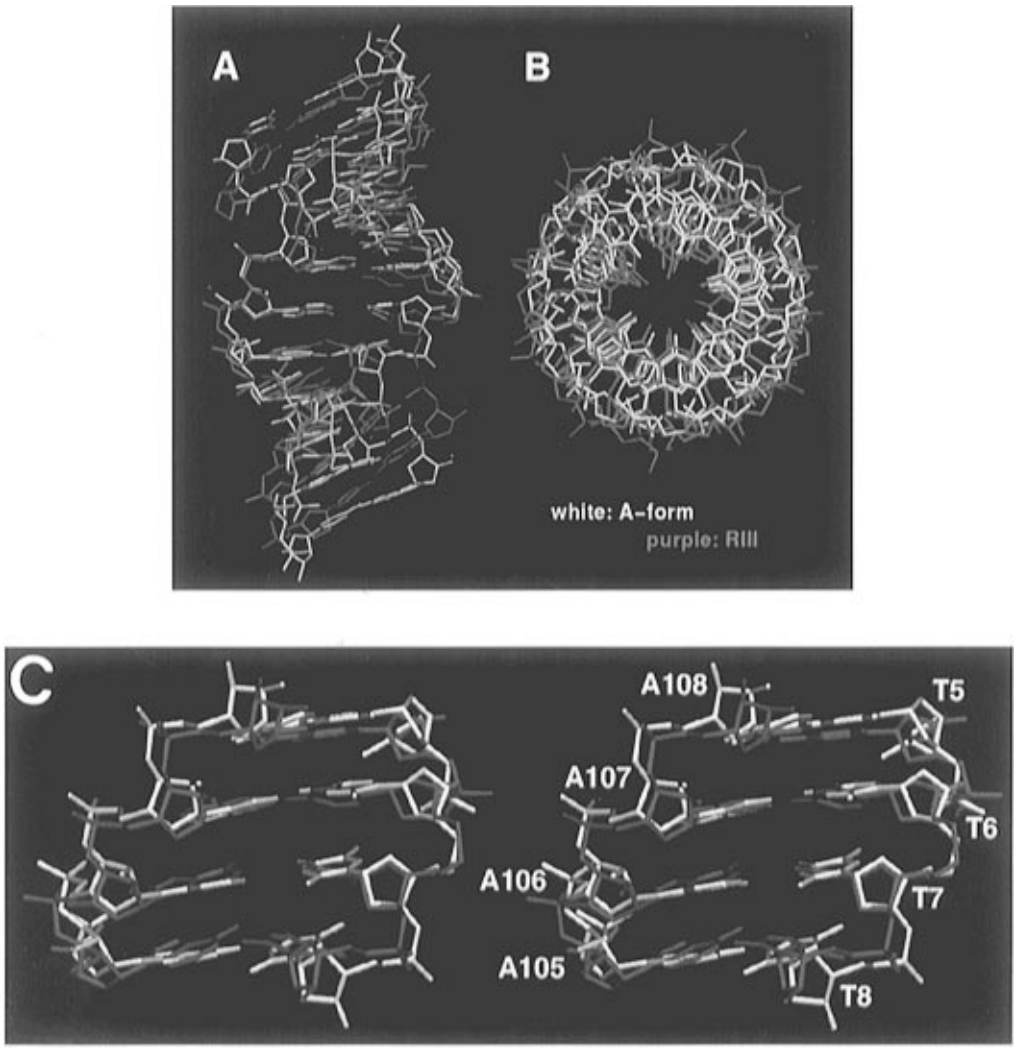

FIGURE 4: Structural comparison of the RIII (purple) and an ideal A-form (white, QUANTA coordinates) hybrid duplex. The two structures are positioned by best fitting the atomic coordinates. (A) A view looking into the minor groove. (B) A view looking down the helical axis. (C) Superpositioning of the A_4·T_4 segment in RIII and a canonical A-form sequence (white). The structure of RIII at the 3′-TFMA linker site shows heterogeneous features and major variations. The view is projected into the minor groove, demonstrating increased base plane perturbation along the 5′- to 3′-direction of the TT-SCH_2O-TT segment.

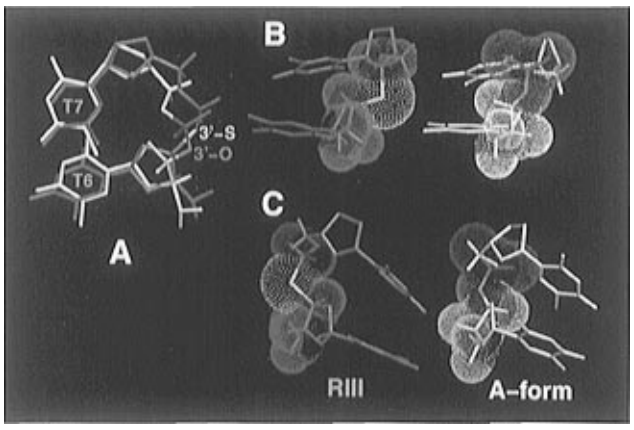

FIGURE 5: Structure of the T6-T7 3′-TFMA backbone linker (magenta) and its comparison with an ideal A-form T-T dimer (white). In the linker region, the color code is yellow for sulfur, green for oxygen, orange for the acetal linker methylene CH_2 group, and magenta for sugar protons. (A) A view looking down the helical axis. (B) The linker and A-form dimer shown with the major groove in front. Full van der Waals radii are shown for the backbone atoms only. (C) The linker and A-form dimer shown with the minor groove in front.

To further analyze the conformation of the A-tract in RIII, the calculated distances related to A H2 protons are compiled in Table 3. These protons are located in the minor groove and make contacts with adjacent imino and A H2 protons in the groove and with H1′ protons on the sugar−phosphate backbone. The interbase A H2 distances and A H2 to imino H distances, in general, decrease from 5′ to 3′of the A-tract and manifest mixed features of canonical A-, B-, or nonstandard forms (Table 3). Among these distances, only the inter-strand interproton distances between A H2 and the 3′-imino H are detectable in the NOESY spectrum of RIII. A H2 protons are also linked to the H1′ protons of its 3′-adjacent residue (intrastrand distances) and to the H1′ protons of the residues 3′ to the residue base paired to that A (interstrand distances). In RIII, such intrastrand distancesfor A105, A106, and A107 are shorter by ∼1 Å than interstrand distances, while A108 shows similar intra- and interstrand A H2 to H1′ separations. This pattern was also observed previously for the unmodified RNA·DNA duplexes (Gao & Jeffs, 1994b). In the DNA·DNA A-track, these distances are comparable and are in the range of 4 Å, as observed by NMR and crystallography (Gao et al., 1992; Nadeau & Crothers, 1989).

TFMA Linker Region. The conformation of the 3′-TFMA linker is illustrated in Figures 4C and 5, which provide a detailed view of how the longer C−S bond and the larger S

atom are accommodated in the backbone. Figure 4C displays an increased base plane perturbation from T5 to T8 (along the 5′- to 3′-direction of the DNA strand) when compared to their counterparts in an ideal A-form duplex. Panels B and C of Figure 5 demonstrate the three-dimensional relationship of sulfur to the surrounding chemical moieties. The pertinent backbone, glycosidic, and phase angles are highlighted in Table 2, and the corresponding helical parameters can be found in Figure 3. The resonances critical for defining the linker conformation, HP1, HP2, H5′, and H5″ of T7, are reasonably well resolved and stereospecifically assigned (Rice & Gao, 1997). This results in 19 distance restraints in the linker region, thereby lending credence to the calculated conformation of the TFMA linker. Figure 2B shows that the calculated and the experimental NOESY spectra for the linker region compare closely.

Figure 3a−c and Figure 3g−j show that there are apparent local conformational variations in the structure of RIII. For instance, starting from T6, the linker residues show a zigzag pattern in the X-displacement; their values fluctuate between −4.0 and −4.7 Å (Figure 3a). The twist angles indicate that the T6-T7 step is unwound by ∼10°, but the following T7-T8 step is overwound by ∼10° (Figure 3j). Correspondingly, the T6-T7 step has a much larger interbase separation distance (3.7 Å) compared to that of the T7-T8 step (2.1 Å) (Figure 3g). The 3′-TFMA-modified T6 residue exhibits the lowest pesudophase and δ angles when compared to the other residues of the DNA strand with the exception of the terminal C1 residue (Table 2). The T6 sugar, therefore, adopts a distinguished conformation. This observation agrees with predictions based on quantum mechanical calculations (Veal & Brown, 1995). These calculations show that the C1′−C2′−C3′−S and O4′−C4′−C3′−S torsion angles, as a result of sulfur substitution, prefer trans conformations, which are better accommodated in a C3′-endo sugar conformation than in a C2′-endo form. Normal B-form DNA·DNA duplexes have their sugar puckers in a predominantly C2′-endo form. In an RNA·DNA duplex the DNA sugars tend to shift to a more C3′-endo-like pucker, moving their average phase angles into the range of O4′-endo puckers. This structural requirement, for hybridization with an RNA sequence, is in line with the energy preference of a 3′-S-substituted sugar. Thus, the presence of a 3′-S linkage is observed to enhance the C3′-endo population for the T6 sugar. The adoption of a more C3′-endo type conformation for the T6 sugar, in combination with the longer C−S bond lengths and larger van der Waals radius of the sulfur (Figure 5B,C), also affects the backbone angles and the adjacent sugar ring (T7) of the linker. The phase angle of the T7 sugar is calculated to be the second lowest value among those of nonterminal DNA sugar rings, indicating an average sugar conformation that is weighted more toward C3′-endo than the other sugars, except T6, in the DNA strand. In addition, there are minor adjustments of the backbone torsion angles within the T6-T7 step compared to the backbone conformation in other portions of the DNA strand of RIII (Table 2, boldfaced values). The overall pattern of the linker torsion angles are +g (87°), t, −g for T6 (δ, ϵ, and ζ) and −g, t, +g for T7(α, β, and γ), respectively (g = gauche; t = trans), which significantly differs from that of the DNA·DNA duplex containing the same linker modification [pattern: +g (79°), +g, +g, +g, t, t] (Gao & Jeffs, 1994a; Veal & Brown, 1995).

DISCUSSION

The high-resolution structure elucidation of the 3′-TFMA-modified RIII hybrid duplex (Figure 1) and the subsequent analysis reveal a number of interesting features important for understanding the fundamental properties of backbone modified hybrid duplexes. The primary goals of this study are to understand, through a detailed analysis, how the 3′-TFMA linker can be accommodated in the DNA strand of the antisense RNA·DNA duplex, RIII, and how this knowledge could contribute to the design of new internucleotide linkers. The following discussion will focus on three areas: (a) the strand properties of the RNA·DNA RIII hybrid duplex, (b) the comparison of the RNA·DNA A·T-tract of RIII with DNA·DNA A·T-tracts, and (c) the structure of the TFMA linker. As we have stated previously, even though our discussions are based on the calculation results of the eight final structures, limited by the inherent resolution of NMR spectroscopy of nucleic acids and restraints used, these data should not be treated as the static, specific definition of conformations. Rather, it is the relative values or comparisons that support the following analysis. A more comprehensive sampling of conformational isomers, especially the sugar pucker equilibrium, may be achieved by time-averaged rMD calculations (Gonzalez et al., 1995), which would require extensive, additional computing.

Our detailed analysis of the structural parameters for each strand of RIII and the comparison of these parameters with the corresponding duplex parameters, as shown in Figures 3 and 4, provide new insights into the properties of RNA and DNA hybridization. Inspection of the parameter plots for each of the two strands and the duplex in Figure 3 indicates that the duplex parameters, which are commonly reported in literature, are an average of the properties of the two strands. In RIII, even though the duplex appears to be an A-form type, neither of the two strands is completely in A-form. A striking example is the inclination (Figure 3b,e), which measures the degree of base plane tilt from a direction looking into the minor groove. The DNA strand residues show large, positive inclination angles ranging from 19 to 38°, compared to a 20° inclination for a canonical A-form duplex. The RNA strand residues show smaller inclination (4−19°) than the ideal A-form value. The hybridization of the two strands results in typical A-form inclination angles in the range of 16−25°. These results reflect the cooperativity of the two strands to achieve a low-energy duplex structure. It has been well accepted that in a hybrid duplex the RNA strand is rigid and the DNA strand is flexible and changes its conformation to meet the needs for binding to RNA. Here our analyses reveal the delicate balance of the conformational variations of the two strands; either DNA or RNA strands can be adaptive to conformational changes upon hybridization.

The local bending of the DNA A-tract sequences has been a subject of intensive investigation due to the potential roles of this structural motif in modulation of transcription and recombination (Nadeau & Crothers, 1989; Nelson et al., 1987; Thompson et al., 1995; Young et al., 1995). RIII contains an RNA A-tract of four consecutive residues. An interesting question is how does the structure of this RNA A-tract in a hybrid duplex compare with that of a DNA A-tract. To examine the related properties of the RIII A-tract, we have analyzed the distances involving interbase A H2 protons and A H2 to imino protons and to H1′ protons

(Table 3), the orientation of the helical axis (Figure 1C), and three helical parameters (tilt, roll, and propeller twist shown in Figure 3l,m,p and Figure 4C). These analyses indicate that the groove features of the A-tract are not affected in a significant way by the backbone modification (RIII compared to RI), but they are distinctly different from the DNA·DNA counter part (Chen et al., 1990; Gao et al., 1992; Nadeau, & Crothers, 1989; Young et al., 1995). This is evident in that the aforementioned helical parameters measured from the A-tract in RIII are always opposite to what are observed in a DNA A-Tract. Therefore, the molecular basis for RNA A-tract bending should be defined by a different set of rules, which remain to be explored when a larger set of structures containing A-tracts become available. The local helical bending analyses have led us to believe that, except for A107, the CAA-A (starting from C104, - is A107) segment is relatively straight. One bend can be found adjacent to the A-tract at each end of the duplex, involving G103/C104 and C109/G110. Both bends orient the A-tract slightly toward the minor groove. These features are consistent with the junction bending model (Koo et al., 1986; Nadeau et al., 1989) and the crystal structure of $d(CGCAAAAAAGCG)_2$ (Nelson et al., 1987). In fact, the structure of RIII matches a proposed model given in 1986 for junction bending caused by A-tract regions flanked by GC regions (Edmondson & Johnson, 1985). This model also predicts a more prominent bend at the 5′-end of the T-tract (or the 3′-end of the A-tract) than at the 3′-end and increasing tilt angles of the bases of the T residues upon going from the 5′- to the 3′-end. Our observations of a larger change in the direction of the helical axis at A108-C109 junction (corresponding to the 5′-end of the T-tract) than that observable at the C104-A105 step (Figure 1C) are consistent with the prediction (Edmondson & Johnson, 1985).

The conformation of the 3′-TFMA linker is defined by five backbone torsion angles and the conformations of the two T residues connected by the linker. These analyses indicate that although there is no major adjustment in the linker region compared to the global helical structure, the TFMA moiety does enforce a local conformational preference. This is characterized by the coordinated adjustment of the ϵ, ζ, α, and β bonds linking the T6 and T7 residues and an increased contribution of a C3′-endo sugar pucker to the two residues. Additionally, the linker step, T6-T7, is measured to have low twist angle and elongated interbase separations. The conformation of the TFMA linker in RIII is clearly different from that observed in the DNA·DNA duplex (Gao & Jeffs, 1994a), indicating the influence of the complementary sequence and the conformational flexibility to the conformation of the linker moiety.

The overall conformation of the linker region is heterogeneous with respect to the rest of the RIII duplex, indicating that the linker torsion angles and sugar conformations are more or less different from the other portions of the modified DNA strand. It is also possible that the observed linker conformation is not occupying the low-energy conformation that is preferred by the 3′-TFMA group, because of restrictions from neighboring groups. These factors may be the molecular origin for the observed lower stability of RIII compared to unmodified RI (Rice & Gao, 1997). Molecular forces at the transition points of different conformations may act against each other and lower the cooperativity in the event of duplex association or dissociation. We speculate that this may explain the observed increase in T_m, the duplex melting temperature, in a duplex containing multiple 3′-TFMA substitutions compared to that of the unmodified sequence (Jones et al., 1993), while RIII containing a single modification has a slightly lower T_m compared to the reference duplex RI (Rice & Gao, 1997). We have initiated the investigation of a disubstituted 3′-TFMA-containing hybrid duplex to further address this question.

The 3′-TFMA linker in a DNA strand is less destabilized when hybridized with an RNA strand (Rice & Gao, 1997), demonstrating the preference for a C3′-endo type of sugar pucker and the capability to adopt a regular backbone conformation by a 3′-S-containing linker. These properties form the basis for the linker to be a promising synthetic template that is effective for the design of neutral, achiral antisense oligonucleotides. These antisense molecules will likely have the advantage of being cell membrane permeable, stable, and highly specific to target RNA sequences. Furthermore, the structural information presented in this work, in combination with NMR analysis, clearly indicates other factors also play important roles. These include the overall structural requirements of the resulting duplex, the sequence-dependent conformational preferences determined by the target RNA and the intrinsic properties of the antisense oligonucleotide, and the preference for maintaining structural homogeneity.

ACKNOWLEDGMENT

The computer facilities at the Institute for Molecular Design at the University of Houston are in part supported by the W. M. Keck Center for Computational Biology. The 600 MHz NMR spectrometer at the University of Houston is funded by the W. M. Keck Foundation. The authors are grateful for the generous chemistry support of Gilead Sciences and Glaxo Pharmaceuticals, Inc. We are grateful to Jennifer Look, who is a 1996 W. M. Keck Summer Scholar Student, for her significant contribution in the preparation of Figure 3 and to G. Kenneth Smith for his careful reading of the manuscript and for valuable suggestions.

SUPPORTING INFORMATION AVAILABLE

One table giving statistics of NOE cross peak distribution (3 pages). Ordering information is given on any current masthead page.

REFERENCES

Agrawal, S., & Iyer, R. P. (1995) *Curr. Opin. Biotechnol. 6,* 12−19.

Alexeev, D. G., Lipanov, A. A., & Skuratovskii, I. Y. (1987) *Nature 325,* 821−823.

Arnott, S., Chandrasekaran, R., Millane, R. P., & Park, H.-S. (1986) *J. Mol. Biol. 188,* 631−640.

Baserga, R., & Denhardt, D. T. (1992) *Ann. N.Y. Acad. Sci.,* 600.

Benevides, J. M., &Thomas, G. J. (1988) *Biochemistry 27,* 3868−3873.

Blommers, M. J., Peiles, U., & De Mesmaeker, A. (1994) *Nucleic Acids Res. 22,* 4187−4194.

Borgias, B. A., & James, T. L. (1989) *Methods Enzymol. 176,* 169−199.

Borgias, B. A., Gochin, M., Kerwood, D. J., & James, T. L. (1990) *Prog. NMR Spectrosc. 22,* 83−100.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4,* 187−217.

Brown, S. C., Thomson, S. A., Veal, J. M., & Davis, D. G. (1994) *Science 265,* 777−780.

Brunger, A. T. (1992) *X-PLOR Manual Version 3.0,* Yale University, New Haven, CT.
Brunger, A. T., & Karplus, M. (1991) *Acc. Chem. Res. 24,* 54−61.
Burkhoff, A. M., & Tullius, T. D. (1988) *Nature 331,* 455−456.
Chary, K. V. R., Hosur, R. V., Govil, G., Zu-Kun, T., & Miles, H. T. (1987) *Biochemistry 26,* 1315−1322.
Chen, S.-M., Leupin, W., & Chazin, W. J. (1990) *Int. J. Macromol. 14,* 57−63.
Chou, S.-H., Flynn, P., & Reid, B. R. (1989) *Biochemistry 28,* 2435−2443.
Cole-Strauss, A., Yoon, K., Xiang, Y., Byrne, B. C., Rice, M. C., Gryn, J., Holloman, W. K., & Kmiec, E. B. (1996) *Science 273,* 1386−1389.
Crooke, S. T. (1992) *Annu. Rev. Pharmacol. Toxicol. 32,* 329−376.
Cruse, W. B. T., Salisbury, S. A., Brown, T., Cosstick, R., Eckstein, F., & Kennard, O. (1986) *J. Mol. Biol. 192,* 891−905.
Ding, D., Gryaznov, S. M., Lioyd, D. H., Chandrasekaran, S., Yao, S., Ratmeyer, L., Pan, Y., & Wilson, W. D. (1996) *Nucleic Acids Res. 24,* 354−360.
Edmondson, S. P., & Johnson, W. C. (1985) *Biopolymers 24,* 825−841.
Egli, M. (1996) *Angew. Chem., Int. Ed. Engl. 35,* 1894−1909.
Eriksson, M., & Nielsen, P. E. (1996) *Nat. Struct. Biol. 3,* 410−411.
Fedoroff, O. Y. U., Salazar, M., & Reid, B. R. (1993) *J. Mol. Biol. 233,* 509−523.
Fritsch, V., Mesmaeker, A. De., Waldner, A., Lebreton, J., Blommers, M. J., & Wolf, R. M. (1995) *Bioorg. Med. Chem. 3,* 321−335.
Gao, X., & Jeffs, P. W. (1994a) *J. Biomol. NMR 4,* 17−34.
Gao, X., & Jeffs, P. W. (1994b) *J. Biomol. NMR 4,* 367−384.
Gao, X., Brown, F. K., Jeffs, P., Bischofberger, N., Lin, K.-Y., Pipe, A. J., & Noble, S. A. (1992) *Biochemistry 31,* 6228−6236.
Gonzalez, C., Stec, W., Reynolds, M. A., & James, T. L. (1995) *Biochemistry 34,* 4969−4982.
Gura, T. (1996) *Science 270,* 575−577.
Heinemann, U., Rudolph, L.-N., Claudis, A., Morr, M., Heikens, R. F., & Blocker, H. (1991) *Nucleic Acids Res. 19,* 427−433.
Jain, R. (1996) *Science 271,* 1079−1080.
James, T. L. (1994) *Curr. Opin. Struct. Biol. 4,* 275−284.
Jie, J., Hogan, M., & Gao, X. (1996) *Structure 4,* 425−435.
Jones, R. J., Lin, K.-Y., Milligan, J. F., Wadwani, S., & Matteucci, M. D. (1993) *J. Org. Chem. 58,* 2983−2991.
Katahira, M., Lee, S. J., Kobayashi, Y., Sugeta, H., Kyogoku, Y., Iwai, S., Ohtsuka, E., Benevides, J. M., & Thomas, G. J. (1990) *J. Am. Chem. Soc. 112,* 4508−4512.
Keepers, J. W., & James, T. L. (1984) *J. Magn. Reson. 57,* 404−426.
Kim, S.-G., Lin, L.-J., & Reid, B. R. (1992) *Biochemistry 31,* 3564−3574.
Koo, H.-S., Wu, H.-M., & Crothers, D. M. (1986) *Nature 320,* 501−506.
Lavery, R., & Sklenar, H. (1988) *J. Biomol. Struct. Dyn. 6,* 63−91.
Lipanov, A. A., & Chuprina, V. P. (1987) *Nucleic Acids Res. 15,* 5833−5844.
Marshall, W. S., & Caruthers, M. H. (1993) *Science 259,* 1564−1570.
Matteucci, M. (1990) *Tetrahedron Lett. 31,* 2385−2388.
Matthiue, J.-M., Tosic, M., & Coach, A. (1992) *Ann. N.Y. Acad. Sci. 660,* 188−192.
Milligan, J. F., Matteucci, M. D., & Martin, J. C. (1993) *J. Med. Chem. 36,* 1923−1937.
Mohan, V., Griffey, R. H., & Davis, D. R. (1995) *Tetrahedron 51,* 8655−8668.
Nadeau, J. G., & Crothers, D. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86,* 2622−2626.
Nelson, H. C. M., Finch, J. T., Luisi, B. F., & Klug, A. (1987) *Nature 330,* 221−226.
Nielsen, P. E., Egholm, M., Berg, R. H., & Buchardt, O. (1991) *Science 254,* 1497−1500.
Nilges, M., Habazettl, J., Brunger, A. T., & Holak, T. A. (1991) *J. Mol. Biol. 219,* 499−510.
Nilsson, L., Clore, G. M., Gronenborn, A. M., Brunger, A. T., & Karplus, M. (1986) *J. Mol. Biol. 188,* 455−475.
Pestka, S. (1992) *Ann. N.Y. Acad. Sci. 660,* 251−262.
Piotto, M. E., Granger, J. G., Cho, Y., Farschtschi, N., & Gorenstein, D. G. (1991) *Tetrahedron 47,* 2449−2461.
Rice, J. S., & Gao, X. (1997) *Biochemistry 36,* 399−411.
Richert, C., Roughton, A. L., & Benner, S. A. (1996) *J. Am. Chem. Soc. 118,* 4518−4531.
Saenger, W. (1984) *Principles of Nucleic Acid Structure,* pp 14−20, Springer-Verlag Press, New York.
Salazar, M., Champoux, J. J., & Reid, B. R. (1993a) *Biochemistry 32,* 739−744.
Salazar, M., Fedoroff, O. Y., Miller, J. M., Ribeiro, N. S., & Reid, B. R. (1993b) *Biochemistry 32,* 4207−4215.
Sanghvi, Y. S., & Cook, P. D. (1993) in *Nucleosides and nucleotides as antitumor and antiviral agents* (Chu, C. K., & Baker, D. C., Eds.) pp 311−324, Plenum Press, New York.
Schmitz, U., & James, T. (1995) *Methods Enzymol. 261,* 3−44.
Stein, C. A. (1996) *Chem. Biol. 3,* 319−323.
Stein, C. A., & Cheng, Y.-C. (1993) *Science 261,* 1004−1012.
Thompson, A. S., Sun, D., & Hurley, L. H. (1995) *J. Am. Chem. Soc. 117,* 2371−2372.
Uhlmann, E., & Peyman, A. (1990) *Chem. Rev. 90,* 544−584.
Vasseur, J. J., Debart, F., Sanghvi, Y. S., & Cook, P. D. (1992) *J. Am. Chem. Soc. 114,* 4006−4007.
Veal, J. M., & Brown, F. K. (1995) *J. Am. Chem. Soc. 117,* 1873−1880.
Veal, J. M., Gao, X., & Brown, F. K. (1993) *J. Am. Chem. Soc. 115,* 7139−7145.
Wickstron, E. (1995) in *Delivery strategies for antisense oligonucleotide therapeutics* (Akhtar, S., Ed.) pp 85−105, CRC Press, New York.
Williams, L. D., Egli, M., Ughetto, G., van der Marel, G. A., van Boom, J. H., Quigley, G. J., Wang, A. H.-J., Rich, A., & Frederick, C. A. (1990) *J. Mol. Biol. 215,* 313−320.
Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids,* John Wiley & Sons, New York.
Young, M. A., Srinivasan, J., Goljer, I., Kumar, S., Beverige, D. L., & Bolton, P. H. (1995) *Methods Enzymol. 261,* 121−144.
Zamecnik, P. C., & Stephenson, M. L. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75,* 280−284.

BI962382K