Structural Effect of Complete [Rp]-Phosphorothioate and Phosphorodithioate Substitutions in the DNA Strand of a Model Antisense Inhibitor-Target RNA Complex

Patrick Furrer[1†], Todd M. Billeci[1], Alessandro Donati[1‡], Chojiro Kojima[1], Boleslaw Karwowski[2], Agnieszka Sierzchala[2¶], Wojciech Stec[2], and Thomas L. James[1*]

Keywords: Antisense; RNA•DNA hybrid structure; phosphorothioate; phosphorodithioate; RNase H; multidimensional NMR

[1]*Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143-0446 USA.*

[2]*Department of Bioorganic Chemistry, Centre of Molecular and Macromolecular Studies, Polish Academy of Sciences, 90-363 Lód´z Sienkiewicza 112, Poland.*

---

*Corresponding author.

†Present Address: Department of Mathematics, Swiss Federal Institute of Technology (EPFL), 1015 Lausanne, Switzerland.

‡Present Address: Department of Chemical Sciences and Technologies, Pian dei Mantellini, 44 53100, Siena, Italy.

¶Present Address: Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309 USA.

Abbreviations used: UV, ultraviolet; NMR, nuclear magnetic resonance; 2D, two-dimensional; NOESY, nuclear Overhauser effect spectroscopy; 2QF-COSY, double quantum filtered corelation spectroscopy; FID, free induction decay; TPPI, time proportional phase increment; RNase H, ribonuclease H.

ABSTRACT

Chemically modified DNA oligonucleotides have been crucial to the success of antisense therapeutics. Although such modifications are ubiquitous in the clinic, high-resolution structural studies of pharmaceutically relevant derivatives have been limited to only a few molecules. We have completed a high-resolution NMR structural study of three DNA•RNA hybrids with the sequence d(CCTATAATCC)•r(GGAUUAUAGG). All hybrids contain an unmodified RNA strand, whereas the DNA strand of each hybrid contains one of three different sugar-phosphate backbone linkages at each nucleotide: 1) phosphate, 2) [Rp]-phosphorothioate, or 3) phosphorodithioate. The UV and NMR melting profiles revealed that the normal hybrid was more stable than the [Rp]-phosphorothioate, which in turn was more stable than the phosphorodithioate. Homonuclear two-dimensional nuclear Overhauser effect spectroscopy and double quantum-filtered correlation spectroscopy afforded nearly complete nonlabile proton assignments. The three molecules show nearly equivalent chemical shifts, with the exception of H3' protons, which are shifted downfield in a manner that appears correlated with the degree of sulfur substitution at phosphate. All three hybrids exhibit unusually broad linewidths for deoxyribose protons H2' and H2".

Distance restraints were calculated from NOE cross-peak intensities *via* a complete relaxation matrix approach using the program RANDMARDI. Detailed comparison of interproton distances from each hybrid indicates that the three molecules share a common structure, with neither strand in canonical A or B form. Correlation of R factors, calculated using the program CORMA with DNA H2'-base and H3'-base distances, revealed a relative increase in the population of B-type sugar conformations for deoxyriboses in the AT-rich center of the hybrid sequence. As ribonuclease H activity varies sulfur substitution, these data suggest that factors other than global structure may mediate ribonuclease H recognition of these species.

INTRODUCTION

Antisense inhibitors are oligonucleotides with sequence complementarity to cellular mRNA or pre-mRNA targets. Binding of a DNA antisense inhibitor to its target RNA results in a DNA•RNA hybrid, affording the sequence-specific inhibition of gene expression *in vivo*. Antisense inhibitors developed as therapeutics are targeted to RNA sequences implicated in disease. However, oligonucleotides with unmodified phosphodiester backbone linkages do not fulfill the pharmacological requirements of effective membrane permeability, target affinity and stability to nucleases (Matteucci, 1996). As a result, all antisense agents currently in clinical trials have modifications of the sugar-phosphate backbone (Figure 1). Backbone modifications substituting one phosphate oxygen with sulfur ([Rp]-phosphorothioate; Figure 1, 1b) or both phosphate oxygens by sulfur (phosphorodithioate; Figure 1, 1c) confer sufficient hydrophobicity to permit delivery across the lipid bilayer membrane of cells. Phosphorothioates and phosphorodithioates also confer resistance to nucleases targeted to single-stranded DNA, thereby increasing the half-life of these oligonucleotides *in vivo* (Spitzer & Eckstein, 1988). Activity of the enzyme ribonuclease H (RNase H), which selectively hydrolyzes the RNA strand of DNA•RNA hybrids, is a key component of the dominant antisense mechanism. Recruitment of RNase H is enhanced if the DNA strand of DNA•RNA hybrids contains more than 50% phosphorothioate-substituted nucleotides. As a result, phosphorothioates direct increased enzymatic destruction of target mRNA, and may substantially increase the potency of phosphorothioate-containing antisense drugs. In contrast, hybrids formed from DNAs containing a sulfur atom substituted for the pro-S phosphoryl oxygen, such as [Sp]-phosphorothioate and phosphorodithioate analogs, show a relative decrease in both susceptibility to RNase H digestion and thermal stability, key components of effective antisense inhibition (Lesser et al., 1992, Cummins et al., 1996, Ghosh et al., 1993).

The pharmaceutical relevance of these modifications has provided some impetus to determine the role of structure in RNase H discrimination between phosphodiester, [Rp]-phosphorothioate, and phosphorodithioate linkages in the DNA backbone of DNA•RNA hybrid substrates. The role of structural differences in the relative stability of these analogs is also of interest. This paper is among the initial attempts to understand the impact of complete [Rp]-phosphorothioate and phosphorodithioate modifications upon the structure of a model antisense inhibitor-target RNA complex at high resolution. Recently, the solution structure of an eight-base-pair DNA•RNA hybrid containing an enzymatically synthesized phosphorothioate DNA strand, [Rp]-d(GCGTCAGG)•r(CCUGACGC), was solved and its structure compared with that of an unmodified analog (Bachelin et al., 1998). Both constructs were double-stranded helices consistent with the A family of helices, excepting DNA-strand sugar puckers and some helical parameters. Given the structural similarity of the [Rp]-phosphorothioate and its phosphodiester analog, Bachelin et al. suggested that phosphoryl group chirality alone could be key to differences in RNase H activity: phosphorodithioate and [Sp]-phosphorothioate-modified hybrids, with a sulfur atom in place of the pro-S phosphoryl oxygen, may substitute less stable S--H hydrogen bonding interactions for more stable O--H hydrogen bonds in the enzyme-hybrid complex (Fedoroff et al., 1993). However, the precise factors mitigating RNase H recognition remain an open question. Particularly, the role of global conformation in RNase H recognition and thermal stability of phosphorodithioate and [Sp]-phosphorothioate-modified hybrids, for which high-resolution structures are unavailable, remains unknown.

We have completed a high-resolution NMR structural study of three DNA•RNA hybrids with the sequence d(CCTATAATCC)•r(GGAUUAUAGG) containing the Pribnow box sequence flanked on the 3' and 5' termini by two G•C base-pair clamps (Figure 2). All hybrids contain an unmodified RNA strand, whereas the DNA strand of each hybrid contains either global phosphate, [Rp]-phosphorothioate, or phosphorodithioate linkages at each deoxynucleotide (Figure 1). Selection of the Pribnow box sequence for this study permitted comparison with a previously reported DNA•RNA hybrid solution structure containing a single [Rp]-phosphorothioate modification (González et al., 1995). Two-dimensional nuclear Overhauser effect spectroscopy (2D NOESY) cross-peak intensities were integrated and distance restraints calculated *via* a complete relaxation matrix approach using the program RANDMARDI (Liu et al., 1995), and these data were coupled with R factor analysis using the program CORMA (James, 1991) to obtain a detailed comparison of the interproton distances from each hybrid. These data, coupled with hybrid stability studies obtained *via* UV and NMR spectroscopy, were used to determine the impact of sulfur substitution on hybrid structure, with the goal of obtaining insights into how RNase H recognizes these therapeutically important molecules.

RESULTS AND DISCUSSION

Hybrid stability, chemical shift effects and dynamics.

Using UV spectrophotometry, melting profiles were examined at low hybrid concentration (~3 - 6 μM). Under these conditions, a melting transition was observed at a higher temperature for the normal hybrid (21 °C) than for the [Rp]-phosphorothioate (18 °C). The least stable molecule was the phosphorodithioate (T_m = 7 °C). To observe a complete melting transition for the phosphorodithioate, additional UV melting experiments were required at higher hybrid concentrations (~ 43 μM). All melting transitions were monophasic, and presumably corresponded to a transition between helix and random coil. An attempt to increase the stability of the phosphorodithioate-modified hybrid was made by including $MgCl_2$ in the sample buffer while keeping the concentrations of other buffer constituents constant. No increase in stability was observed up to a $MgCl_2$ concentration of 40 mM (data not shown).

The T_m values for each hybrid were measured at concentrations consistent with those employed for NMR experiments by acquiring spectra from single-pulse (90° pulse) NMR experiments at varying temperatures and monitoring the signal amplitudes of well-isolated peaks (Figure 3). The NMR profiles showed the same trends as in the UV experiments:

$$T_m(\text{phosphodiester}) > T_m(\text{[Rp]-phosphorothioate}) > T_m(\text{phosphorodithioate})$$

The T_m values obtained under these conditions were approximately 33 °C, 29 °C, and 22 °C for the normal, [Rp]-phosphorothioate and phosphorodithioate, respectively. As with the UV data, all melting curves were monophasic.

The inverse relationship between increasing sulfur substitution and duplex stability has been observed previously *via* low-resolution NMR studies of a phosphorothioate DNA duplex consisting of a random mixture of diastereomers (Jaroszewski et al., 1996). In these studies, the base pair lifetimes were considerably shorter for phosphorothioates than for unmodified duplexes, providing a possible explanation for such stability differences. Other previous structure-stability studies also corroborate these findings (Freier & Altmann, 1997, Cummins et al., 1996, Koziolkiewicz et al., 1995).

Proton chemical shifts were measured as a function of temperature using well-isolated adenine H2 proton peaks observed in one-dimensional NMR spectra and monitored through the hybrid melting transition. The chemical shift variation with temperature was determined to be about 0.002 ppm/°K. This relatively low value characterizes the strand association-dissociation kinetics as a slow exchange process compared to the NMR timescale. This slow exchange signature is common to all three molecules, and therefore may be a feature of the Pribnow box sequence itself. It has been reported that TA base steps in DNA duplex sequences may produce slow dynamic processes occurring on a timescale up to the order of milliseconds (Lingbeck et al., 1996), and $T_{1\rho}$ (spin-lattice relaxation time in the rotating frame) measurements an a DNA duplex containing the Pribnow box sequence revealed exchange on the μs-ms time scale which was particularly prevalent in the AT-rich Pribnow box region (Schmitz et al., 1992). Attempts to characterize the dynamics of the phosphorodithioate-modified analog by exchange spectroscopy failed (data not shown), suggesting that the DNA•RNA hybrids studied here may exhibit motions on a timescale even slower than that for DNA duplexes.

Proton resonance assignments.

The nucleotides in the three molecules are numbered as in Figure 2. Sequential resonance assignments of the nonexchangeable protons were made independently for each hybrid using 2D NOESY and 2QF-COSY data. Standard sequential resonance assignment walks involving NOEs between ribose and base protons could be traced without interruption in the 2D NOESY spectra of each hybrid. The unbroken walks and NOE cross-peak intensities for all three molecules were consistent with those expected for right-handed nucleic acid duplexes, especially DNA•RNA hybrids (González et al., 1994, Lane et al., 1993). The chemical shifts of proton resonances in the DNA strand and RNA strand of all three molecules are given in Tables 1a and 1b, respectively. The aromatic-H1' region of the 100 ms phosphorodithioate NOESY spectrum is given in Figure 4.

Significantly, the chemical shift assignments of the central six base-pairs of the normal hybrid, corresponding to the Pribnow box sequence, corroborated those published for a DNA•RNA hybrid containing the Pribnow box sequence with a single [Rp]-phosphorothioate substitution (González et al., 1994). Although the unmodified and singly-substituted hybrid assignments were in good agreement, this was not the case for the hybrids containing complete [Rp]-phosphorothiote and phosphorodithioate substitutions. For these species a pattern was apparent: the H3' proton chemical shifts in the DNA strand (see Table 1a) were shifted downfield for the [Rp]-phosphorothioate, and even further downfield for the phosphorodithioate, such that H3'-base walks completely overlapped the H_2O resonance line to be located just below the first H1' resonances. Similar to the differences in hybrid melting temperature described above, this H3' chemical shift displacement may be correlated with the degree of sulfur substitution at the phosphoryl group. This effect likely results from locally different electronegativities, as the H3' proton is closest to the locus of sulfur substitution. The chemical shift effect provided an additional check of proton assignments, and may prove a useful tool to facilitate the assignment of NMR spectra of modified nucleic acids with novel sequences.

Linewidths.

The H2' and H2" resonances of all three hybrid DNA strands were unusually broad. This linewidth broadening afforded near-complete overlapping of H2'-H2" NOE cross-peaks such that these peaks were unusable for distance calibrations. Further, 2QF-COSY cross-peaks involving these protons were too broad to permit an unambiguous sugar conformation analysis. Linewidths derived from gaussian linefitting in the program SPARKY were collected and averaged 1) over all cross-peaks corresponding to a single proton, and 2) over the entire hybrid sequence (see Table 2). The differences between the two strands are striking for the H2'-H2" signals, where average linewidths for the DNA are approximately twice that of the H1' RNA linewidths. These unusually broad H2' and H2" linewidths may possibly be related to the strand association-dissociation reaction. Dynamic processes, such as those derived from molecular flexibility, may also be involved.

Relaxation time measurements and estimation of correlation time.

The measurements of spin-lattice relaxation time T_1 and spin-spin relaxation time T_2 utilized six well-isolated peaks corresponding to four adenine H2s and two thymine methyl groups. As expected, H2 protons have longer T_1 values than the other protons (see Table 3) due to the limited local proton density. A global correlation time τ_c was calculated from T_1 and T_2 values employing equation (3). An average τ_c value of 2.2 ± 0.3 ns, assuming isotropic motions, was measured for the phosphorodithioate hybrid, and served as input for subsequent MARDIGRAS calculations.

From NOE experiments to distance restraints.

We obtained distance restraints independently for the following 2D NOE data sets: normal hybrid, 150 ms; [Rp]-phosphorothioate, 150 ms and 300 ms; phosphorodithioate, 100 ms and 150 ms. One data set was selected for each molecule, on the basis of spectral signal-to-noise, overall spectral quality, and comparable mixing times. Distance comparisons were based upon 2D NOE spectra acquired at 150 ms, 150 ms and 100 ms for the unmodified (309 interproton distances), [Rp]-phosphorothioate (212 interproton distances) and phosphorodithioate (249 interproton distances), hybrids, respectively.

Interproton distances can be a powerful probe of three-dimensional structure. However, there can be cases where immeasurably small changes in interproton distances correspond to radically different molecular conformations. As a result, it is crucial to determine as many interproton distances as possible with the greatest accuracy. To this end, all interproton distances were computed *via* a complete relaxation matrix approach using the program MARDIGRAS (Liu et al., 1992). This method has the advantage of accounting for indirect magnetization transfer effects (spin diffusion), yielding accurate interproton distances from measured 2D NOE intensities. To account for complications in assessing distance bounds arising from signal-to-noise limitations and integration errors, the RANDMARDI option in MARDIGRAS was employed (Liu et al 1995).

Integration errors were estimated from the relative differences in the volumes of NOE cross-peaks on each side of the diagonal. For example, the normal hybrid afforded 309 integrated resonances, 230 of which could be integrated on both sides of the diagonal. The relative differences gave a skewed distribution of integration errors, with values ranging from 0 % to 90%. The mean value for these data was 16%, the median 9.5%, and the standard deviation 17%. Based on these calculations, we used an integration error of 20% in subsequent RANDMARDI calculations, corresponding to a value slightly higher than the mean integration error calculated for the normal hybrid data. Due to the sixth-root relationship between integration volumes (intensities) and distances, this 20% average error of integration is translated by RANDMARDI into fluctuations around the average distances of 3.2±1.9% (median at 2.6%). In other words, a 20% integration error used with RANDMARDI limits interproton distances to a precision better than 0.1 Å.

Comparison of interproton distances

Average interproton distances from the three hybrids were computed *via* either an interproton rmsd or an average relative difference (see equations 4 and 5 in the Material and Methods section). Table 4 summarizes the results of this comparison. Analysis of interproton distances based on the rmsd or relative average differences requires considerable caution, however: these indicators may not be sensitive to individual differences which computations based on averages will not reveal. As a result, relative distance differences were also calculated and compared for all three molecules.

The above distance comparison protocol was applied to the phosphorodithioate distances computed using RANDMARDI analysis of NOESY cross-peak intensities from spectra acquired with mixing times of 100 ms and 150 ms. This calculation provided a reference for rmsd and relative difference measurements. The two data sets have 179 interproton distances in common, with an rmsd of 0.3 Å, a relative difference of 7.7%. Fifty-seven distances differed by more than 10% and 13 distances differed by more than 15%. This defines the experimental accuracy limit, providing a threshold below which interproton distance differences between the three hybrids can be considered insignificant.

Another method for estimating the accuracy of this method could be based on an analysis of the starting-structure dependence of calculated distances. The starting structure simulates missing experimental intensities at the beginning of each MARDIGRAS calculation. As the number of experimental intensities increases, the influence of the starting model decreases. Also, increased accuracy is positively correlated with an increase in the number of experimental intensities. To measure the effect of starting structure on distance calculations, the distance comparison protocol described above was applied to the interproton distances outputs from RANDMARDI calculations using different starting models. Four starting model structures were utilized for RANDMARDI calculations. The Pure-A, Pure-B, Mixed and singly-substituted [Rp]-phosphorothioate starting structures (see Materials and Methods) minimally impacted interproton distance calculations. The average rmsd deviation for interproton distances of the phosphorodithioate data set (249 distances), using Pure-A and Pure-B as starting models (the A-form and B-form double helices) was 0.28 Å. This defines the accuracy limit to be closer to 0.3 Å, rather than the 0.1 Å limit calculated for mixing time effects. An rmsd between interproton distances with a value below this threshold would be insignificant. These results suggest that, above a 0.3 Å limit, it is possible to compare interproton distances from three macromolecules without interference from systematic errors derived from the choice of starting structure. However, as a general rule, it is probably wise to carefully examine the effect of starting structure when studies of this type are performed.

Among the three pairwise comparisons, the most accurate, in terms of the number of common distances available from RANDMARDI, was that between the unmodified and phosphorodithioate data, with 169 common interproton distances, among which only 7 differ by between 10 and 16%. Thus, 162 of the distances in common differ by less than 10%. The majority of the largest differences (>10%) may be assigned to overlapping cross-peaks in either one of the spectra which pose significant integration problems.

The pairwise comparisons with the [Rp]-phosphorothioate data show a little larger rmsd (0.22 and 0.25 Å), with significantly more candidates with larger distance differences (11 for the unmodified hybrid and 24 for the phosphorodithioate (see Table 4). However, the differences between the three molecules lie at the limit of the ~ 0.3 Å accuracy limit. Thus, the differences between these molecules, if any, are too subtle to be quantified given the accuracy limit of the distance comparison method. Using the alternate pairwise distance comparisons, all exhibit an average interproton distance rmsd smaller or no larger than the typical rmsd due to starting structure dependence. In summary, we were not able to identify any interproton distances differing by more than 10% between two molecules which could not be attributed to starting structure dependence or peak overlap when individual spectra were examined.

Analysis of the DNA•RNA hybrid structure.

Now that our three molecules have been shown to possess a common structure, the most accurate data set of the three, that of the normal hybrid (with 309 distance restraints) can be analyzed in detail. Although refinement of this structure is still in progress, useful insights can be gleaned when interproton distances are considered, notably those discriminating A and B-type helices.

On the DNA strand, the H3'-base intraresidue and H2'-base interresidue distances show a peculiar profile, with distances about 1 Å larger at the center of the sequence (A4-T5-A6-A7) than at either end of the molecule (Figure 5a). The same pattern is exhibited for H1'-base intraresidue distances, with larger values in the center of the DNA strand (from 3.4-3.8 Å) than at its ends (from 3.1 to 3.4 Å). The H1'-H4' intraresidue (2.9±0.2 Å) and H2"-base intraresidue (3.3±0.5 Å) distances are typical of canonical B-DNA; however, the H2'-base intraresidue (2.7±0.2 Å), H3'-base interresidue (3.5±0.3 Å) and H2"-base interresidue (2.9±0.2 Å) distances all show distances between those typical of canonical A and B forms. The single DNA H2-H2 sequential distance (A6-A7, 3.7 Å) clearly typifies B-DNA (3.6 Å), compared to A-DNA (4.8Å). However, the sequential H5-Me (C2-T3, 3.4 Å) or Me-H5 (T8-C9, 3.9 Å) connectivities usually absent in B-form helices, clearly suggest either a noncanonical conformation or some degree of conformational averaging in the DNA strand.

On the RNA strand, the H1'-H3' intraresidue distances are particularly short, ranging between 2.6 and 3.1Å (both A and B-forms should be 3.6 Å). Whereas H3'-base intraresidue and interresidue distances exhibit typical A-form values, other distances like H2'-base intraresidue (3.3±0.3 Å) and interresidue (2.6±0.1 Å) distances indicate that the RNA conformation is not purely A-form. Surprisingly in the H1'-base sequential distances, the characteristic canonical B-form distance of 3.4 Å is accommodated by both the DNA strand (3.3±0.3 Å) and the RNA strand (3.3±0.1 Å), substantially deviating from the 4.6 Å value typical of A-form.

Cross-strand distances are more difficult to analyze in terms of A- versus B-form characteristics. For DNA duplexes, typical cross-strand NOEs involving H2s on one strand and H1's on the other strand are a clear signature of A-type DNA. For DNA•RNA hybrids, these distances are usually observed (Gonzalez et al., 1994, Mujeeb et al., 1997) and provide valuable restraints for structure refinement. For these hybrids, these cross-strand distances were shorter when the H2 was on the DNA strand (about 3.2 Å), and slightly longer when the H2 was located on the RNA strand (A4H1'-A18H2=3.6 Å, A6H1'-A16H2=3.9 Å).

R^X factor analysis.

Sixth-root R factor calculations were carried out for the normal hybrid using four different starting models (see Materials and Methods). Global R^X factors measured for the four starting structures were 0.14, 0.11, 0.10 and 0.09 with the Pure-A, Pure-B, MIXED and CONTROL starting structures respectively, suggesting that our hybrid is better fitted with a mixed A-form RNA strand and B-form DNA strand than with pure A-type or B-type double helices. As these structures have not been refined, the global R^X values can only be compared qualitatively. However, these global R^X factors can be decomposed into local intraresidue (n=168) and interresidue (n=141) observables, as well as into per-residue factors. From Figure 4a, where intraresidue R^X values are plotted for every nucleotide, the Pure-A starting model fits the RNA strand better than the Pure-B, except for residue U15. This behavior reverses on the DNA strand, where the Pure-B model agrees well with the data for A4 and T5. The Pure-B model is in significantly better agreement with the structure than Pure-A for all residues except for C1. As the MIXED model takes advantage of being closer to B and A in the DNA and RNA strands, respectively, it manifests better R^X factors both locally and globally. For the interresidue distances, the poor fit between experimental data and the Pure-A model along the RNA strand is largely improved by the MIXED model. Here again, the Pure-B is strongly preferred on the DNA strand, as well as for the U14-U15 on the RNA strand, which confirms the intraresidue U15 preference for B-type conformation.

In summary, analysis of interproton distances and R^X factors suggests that the DNA strand is more B-like and the RNA strand more A-like. This heterogeneity of the two strands has been observed previously in backbone-modified and unmodified DNA•RNA hybrids (Lane et al., 1993, Mujeeb et al., 1997). However, in contrast to a similar DNA•RNA hybrid (González et al., 1995) with all distance restraints compatible with riboses in the N-type (C3'-endo) domain, the RNA strands here exhibit some average interproton distances clearly inconsistent with pure N-type puckering modes. These include 1) H1'-base interresidue distances with values typical of B-form, and 2) H1'-H4' intraresidue and H2'-base inter- and intra-residue distances for which RANDMARDI accurately excludes both the A- and B-type conformations. The H2'-base interresidue, H1'-base intraresidue and H3'-base intraresidue DNA distances lend support to the hypothesis that the central TATAAT DNA bases are closer to B-type nucleotides than the terminal bases, observed previously in MD-tar simulations in water for a similar sequence (Gonzalez et al., 1995). Significantly, these effects have been correlated with the decreased minor groove width of DNA•RNA hybrids relative to DNA duplexes, as well as extensive deoxyribose repuckering, hypothesized to play a role in recognition of hybrids by the enzyme RNase H.

It is somewhat apparent how both [Rp]-phosphorothioate-containing hybrids and unmodified hybrids could serve as RNase H substrates, given the substantial structural similarity of all three molecules as determined by these methods. However, it is much less clear why phosphorodithioate-containing hybrids fare much more poorly as RNase H substrates, given these species share a structure similar to that of phosphorothioate-containing and normal hybrids. It is possible that dynamic processes, binding thermodynamics, and stabilization of the transition state mediated by sulfur could play some role. Additionally, the possibility of subtle, yet functionally significant structural differences beyond the detection limit of this method must not be ruled out.

Conclusions

We have examined how global [Rp]-phosphorothioate and phosphorodithioate substitutions affect the stability and structure of a model antisense DNA oligonucleotide hybridized with its RNA target. The three DNA•RNA hybrids studied showed differences in thermal stability *via* both UV and NMR spectroscopy. The stability decreases with the degree of sulfur substitution in the DNA backbone. The H3' chemical shift of the deoxyribonucleotides in each hybrid was shifted downfield in a manner that correlated with the degree of sulfur substitution at the phosphoryl moiety. The hybrids also showed unusually broad DNA H2' and H2'' linewidths, suggesting some exchange.

A detailed comparison of accurate interproton distances calculated with the program RANDMARDI allowed us to conclude that, within the accuracy of our experimental method (0.3 Å), our three molecules share a common structure. The salient features of this structure have been studied by qualitative comparison of local sixth-root R factors with model structures, as well as a quantitative analysis of interproton distances for the unmodified hybrid. Neither the RNA nor DNA strand is in canonical A or B form. The NOE-derived interproton distances, 2QF-COSY data and local R^x values clearly suggest that the DNA strand manifests conformational flexibility, with a larger percentage of B-form characteristics at the AT-rich center of the DNA strand. The 2QF-COSY data are mostly consistent with RNA sugars in C3'-endo conformations (except G11), as already observed for high-resolution solution structures of other hybrids with partial modification of the DNA strand (Gonzalez, 1994, Mujeeb, 1997). Several interproton distances located in the RNA strand deviate from canonical A form. Importantly, these effects have been correlated with the decreased minor groove width of DNA•RNA hybrids relative to DNA duplexes, as well as extensive deoxyribose repuckering, hypothesized to play a role in recognition of hybrids by the enzyme RNase H. The noncanonical hybrid structure may also play a role in recognition by this enzyme. Although the substantial structural similarity of all three molecules suggests how both [Rp]-phosphorothioate-containing hybrids and unmodified hybrids could serve as RNase H substrates, it remains unclear precisely why phosphorodithioate-containing hybrids should fare much more poorly as RNase H substrates than analogs quite similar in conformation. More detailed studies will be required to gain further insights into RNase H-substrate recognition.

MATERIALS AND METHODS

Sample preparation.

The [Rp]-phosphorothioate and phosphorodithioate DNA oligomers were synthesized as described (Ohrustec et al., 1995, Stec et al., 1995). The unmodified DNA strand was synthesized *via* standard phosphoramidate chemistry, and the RNA strand synthesized using a method reported previously (Scaringe et al., 1990).

Crude oligonucleotides were purified by HPLC in three steps. First was a reversed-phase step using a C18 column with elution *via* gradient of 5-40% acetonitrile in 100 mM triethylammonium acetate increased linearly over 30 minutes. The second purification step utilized a Mono Q anion exchange column (Pharmacia LKB Biotechnology). Finally, a gel filtration step with a G3000 (Tosoh) column (21.5 mm i.d. X 600 mm) was performed using a flow rate of 6 ml/min. Samples were reconcentrated by lyophilization and/or using MACROSEP 1K membranes with centrifugation overnight. The buffer used for all experiments was 0.1 mM EDTA, 100 mM NaCl, 20 mM sodium phosphate, pH 6.5. To form the hybrids for NMR experiments, the DNA and RNA strands were mixed in a 1:1 stoichiometry, heated to 330 °K, and slowly cooled to room temperature for annealing in an insulated bottle (total cooling time ~3 hours). Final hybrid duplex concentrations were about 3 mM. Two to three steps of further lyophilization were then followed in order to exchange water for D_2O. The phosphorodithioate sample was filtered through a z-spin membrane before placement in a 5 mm Shigemi NMR tube.

UV melting profiles.

Melting profiles were acquired on a CARY 118 spectrophotometer by measuring the absorbance at a wavelength of 260 nm as a function of temperature. Samples were heated at a rate of 0.2 °K/minute. Absorbance data were collected at 0.2 degree intervals with a signal averaging time of 1-2 seconds. Hybrid concentrations for the UV measurements were in the range of 0.5 - 1 O.D. measured in quartz cuvettes with a path length of 1 cm. The sample buffer employed in UV melting studies was identical to that used in NMR experiments (*vida supra).* Melting transitions for the phosphorodithioate sample were also acquired using a shorter path length cuvette (1 mm instead of 1 cm) to increase the melting temperature sufficiently to observe a complete transition curve. We also attempted to increase the stability of the phosphorodithioate-modified hybrid by increasing the concentration of $MgCl_2$ in the sample buffer, up to a $MgCl_2$ concentration of 40 mM, while the concentrations of other buffer constituents were held constant.

NMR experiments.

The NMR experiments were performed at 500 MHz for the phosphorodithioate hybrid and 600 MHz for the unmodified and [Rp]-phosphorothioate hybrids, using either a General Electric GN-500 or a Unityplus 600 NMR spectrometer. The NMR data were processed using the program NMRPipe (Delaglio et al., 1995) on Silicon Graphics workstations, and the files imported into our locally developed NMR data analysis program SPARKY for assignment and integration. All 2D NMR experiments were performed in D_2O containing 0.5 mM sodium 3-(trimethylsilyl) propionate-2,2,3,3-d_4 for chemical shift reference.

Pure absorption-phase 2D NOE spectra were acquired in hypercomplex mode with a spectral width of 5,000 Hz in both dimensions and 2048 X 512 data points along the ω_2 and ω_1 dimensions, respectively. The phosphorodithioate spectra were acquired with a spectral width of 4,000 Hz. Sixteen free induction decay signals (FIDs) were acquired for each t_1 value using a repetition delay of 10 s. The data were apodized with a parametric sine-bell function which was applied sharply for assignment purposes and somewhat more mildly for integration. The solvent signal was suppressed by polynomial subtraction, and zero-filling was applied to achieve a final file size of 2K X 2K. Baselines were corrected by fitting with a 5^{th} order polynomial equation.

Pure absorption-double-quantum-filtered COSY (2QF-COSY) spectra were acquired in TPPI mode, with 2048 X 512 points, and a spectral width of 5,000 Hz in both dimensions. The number of scans was 64, the repetition delay was 3 s, and sharp apodization parameters were used during data processing.

Spin-lattice relaxation times (T_1) were measured *via* the inversion-recovery method using 14 delays between 200 ms and 10 sec. Each FID was repeated 16 times, with a repetition delay of 10 sec. Spin-spin relaxation times (T_2) were measured *via* the spin-echo method, using 9 delays between 10 and 200 ms. Each FID was repeated 16 times, with a repetition delay of 10 s. Six isolated cross peaks were monitored for measuring T_1 and T_2, and their line heights M plotted and fitted using the program Kaleidagraph with the following exponential decay laws:

$$M=M_0*(1-2*\exp(-t/T_1)) \qquad (1)$$

$$M=M_0*\exp(-t/T_2) \qquad (2)$$

where t represents time and T_1, T_2 and M_0 are the parameters to fit.

A single correlation time τ_c can be estimated from measured T_1 and T_2 according to the following formula (Suzuki et al., 1986):

$$\tau_c = \frac{2}{\omega}\sqrt{\frac{T_1}{3T_2}} \qquad (3)$$

where $\omega = 2\pi$ x represents the spectrometer frequency.

Starting model structures.

Four starting models were used for MARDIGRAS calculations. Two of these, designated hereafter as Pure-A and Pure-B, were constructed as canonical A-form and B-form double helices, respectively. A third starting model, designated hereafter as the MIXED model, was constructed using an A-form RNA strand and a B-form DNA strand. These three models were made using the NUCGEN module of the AMBER suite of programs (Pearlman et al., 1991), and energy-minimized with the SANDER module of AMBER *via* 400 conjugate gradient steps followed by 100 steps using the steepest descent method. A distance-dependent dielectric constant was used to mimic bulk solvent effects and solvating cations were added as hexahydrated sodium. The fourth starting model for our experiments was generated from the published solution structure of the DNA•RNA hybrid d(GCTATAA$_{ps}$TGG)•r(CCAUUAUAGC), which contained a single [Rp]-phosphorothioate modification and a sequence which differed slightly from ours at the termini (González et al., 1995, PDB accession number 219d). To generate a starting model from this structure, we replaced three base-pairs to give the correct sequence and energy minimized in the neighborhood of the substitutions.

From 2D NOE intensities to distance restraints.

Once assigned, NOE cross-peaks were integrated with gaussian fitting using the program SPARKY. Some peaks were integrated by fitting with an ellipse in cases where the baseline was not amenable to gaussian fitting and/or the peaks were of relatively low intensity. Cross-peaks on both sides of the diagonal were integrated wherever possible, and the average of the two intensities was used in subsequent MARDIGRAS calculations. The differences between the volumes across the diagonal served as an estimate of integration error. For intensity normalization, we compared two different protocols available within the MARDIGRAS program: 1) normalization using fixed intensities, and 2) normalization using all intensities. Using normalization against fixed intensities, the ten Me-H6 and H5-H6 distances in the molecule were set as fixed reference distances with values of 2.46 Å and 2.98 Å, respectively. Alternatively, normalization against all intensities utilized the sum of all distances as a reference. The results of our comparisons between the three molecules were negligibly affected by choice of normalization protocol. To account for the effects of spectral noise and integration errors, we employed the RANDMARDI option in MARDIGRAS (Liu et al., 1995). Fifty MARDIGRAS cycles were run for every NOE dataset. A three-state jump model was adopted for methyl proton internal motions (Liu et al., 1992).

RMS-mardi calculations. To compare MARDIGRAS (or RANDMARDI) lists of interproton distances from the different hybrids, we built a program which evaluates the lists of *n* distances present in two output files and calculates the root-mean-square deviation between them (rmsd/Å), as well as their average relative difference (ardiff /%), according to the following formulas:

$$rms = \sqrt{\frac{\sum_{j=1,n} (d_a(j) - d_b(j))^2}{n}} \qquad (4)$$

$$ardiff = \frac{100}{n} \cdot \sum_{j=1,n} 2 \cdot \frac{|d_a(j) - d_b(j)|}{d_a(j) + d_b(j)} \qquad (5)$$

where d_a and d_b refer to the values for the same interproton distance in each of two lists a and b. The program also produces a list of the distances which differ by more than a user-defined value when the two lists are compared. This algorithm was employed iteratively. At each round, outlying data were reassessed, and either reassigned or re-integrated. This method proved most useful for the identification of misassignments, and permitted the critical reevaluation of the integration of low intensity peaks, as MARDIGRAS can be sensitive to this source of error.

R^x Factor Calculations. The complete relaxation matrix program CORMA permits comparison of experimental 2D NOE intensities with values calculated for a

given structure. The fit of experimental and calculated intensities is reflected in the term R^X, a residual index analogous to the crystallographic R factor weighted according to the sixth-root dependence of 2D NOE intensities on distance (James, 1991):

$$R^X = \frac{\sum_i \left| \sqrt[6]{a_0(i)} - \sqrt[6]{a_c(i)} \right|}{\sum_i \sqrt[6]{a_0(i)}} \qquad (6)$$

where a_o and a_c are the observed and calculated intensities, respectively. The R^X values were calculated using all four starting models. The R^X values for each individual residue, separating intraresidue or interresidue observables, were computed to assess the conformational preferences of individual nucleotides.

ACKNOWLEDGMENTS

The authors wish to thank He Liu, Nick Ulyanov and Marco Tonelli for their valuable assistance with MARDIGRAS and RANDMARDI calculations. David Konerding generously assisted us in generating starting model structures. Tom Goddard provided very useful modifications of the program SPARKY. Finally, we wish to thank Anwer Mujeeb and Peter Lukavsky for sharing their insight and expertise. We acknowledge use of the UCSF Computer Graphics Laboratory, which is supported by a grant from the National Institute of Health (RR01081). This work was supported by National Institute of Health grant No. GM39247 (to T.L.J.), the Polish-American Maria Curie-Sklodowska Fund II grant No. MZ/HHS-95-228 (to W.J.S.), a CNR (National Research Council)-NATO grant, notice of competition 215.28/16 (to A.D.), and a Swiss National Science Foundation (Fonds National Suisse de la Recherche Scientifique) fellowship (to P. F.).

REFERENCES

Bachelin, M., Hessler, G., Kurz, G., Hacia, J. G., Dervan, P. B., Kessler, H., (1998) *Nature Struct. Biol. 5*, 4429-4443.

Cummins, L., Graff, D., Beaton, G., Marshall, W. S., & Caruthers, M. H. (1996) *Biochemistry 35*, 8734-8741.

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., Bax, A. (1995) *J. Biomol. NMR 6* 227-293.

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

Freier, S. M., Altman, K.-H. (1997) *Nucl. Acids Res. 25* 4429-4443.

Ghosh, M. K., Ghosh, K., Dahl, O., & Cohen, J. S. (1993) *Nucleic Acids Res. 21*, 5761-5766.

González, C., Stec, W., Kobylanska, A., Hogrefe, R. I., Reynolds, M., & James, T. L. (1994) *Biochemistry 33*, 11062-11072.

González, C., Stec, W., Reynolds, M., & James, T. L. (1995) *Biochemistry 34*, 4969-82.

James, T. L. (1991) *Curr. Opin. Struct. Biol. 1*, 1042-1053.

Jaroszewski, J. W., Clausen, V., Cohen, J. S., & Dahl, O. (1996) *Nucleic Acids Res. 24*, 829-834.

Koziolkiewicz, M., Krakowiak, A., Kwinkowski, M., Boczkowska, M., & Stec, W. J. (1995) *Nucl. Acids Res. 23*, 5000-5.

Lane, A, Ebel, S., & Brown, T. (1993) *Eur. J. Biochem. 215*, 297-306.

Lesser, D. R., Grajkowski, A., Kurpiewski, M. R., Koziolkiewicz, M., Stec, W. J., & Jen-Jacobson, L. (1992) *J. Biol. Chem. 267* 24810-24818.

Lingbeck, J., Kubinec, M. G., Miller, J., Reid, B. R., Drobny, G. P., & Kennedy, M. A. (1996) *Biochemistry 35(3)*, 719-34.

Liu, H., Thomas, P. D., & James, T. L. (1992) *J. Magn. Reson. 98*, 163-175.

Liu, H., Spielmann, H. P., Ulyanov, N. B., Wemmer, D. E., & James, T. L. (1995) *Journal of Biomolecular NMR 6*, 390-402.

Matteuci, M. (1996) *Perspectives in Drug Discovery and Design 4*, 1-16.

Mujeeb, A., Reynolds, M. A., & James, T. L. (1997) *Biochemistry 36(9)*, 2371-2379.

Okruszek, A., Sierzchala, A., Fearon, K. L., Stec, W. J. (1995) *J. Org. Chem. 60*, 6998-7005.

Pearlman, D. A., Case, D. A., Caldwell, J. C., Seibel, G. L., Singh, U. C. Weiner, P., & Kollman, P. A. (1991) AMBER 4.0, University of California, San Fancisco.

Scaringe, S.A., Francklyn, C., & Usman, N. (1990) *Nucl. Acids Res. 18*, 5433-5441.

Schmitz U., Sethson, I., Egan, W., & James, T. L. (1992) *J. Mol. Biol. 227*, 510-531.

Spitzer, S., & Eckstein, F. (1988) *Nucl.Acids Res.16(24)*,11691-11704.

Stec, W. J., Grajkowski, A., Karwowski, B., Kobylanska, A., Koziolkiewicz, M., Misiura, K., Okruszek, A., Wilk, A., Guga, P., & Boczkowska, M. (1995) *J. Am. Chem. Soc., 117*, 12019-12029.

Suzuki, E., Pattabiraman, N., Yon, G., & James, T. L. (1986) *Biochemistry 25*, 6854-6865.

Table 1 a: Chemical shifts (ppm) for assigned protons in the normal hybrid DNA strand. If the chemical shift of the same proton differs by more than 0.1 ppm in the [Rp]-phosphorothioate or phosphorodithioate analogs, the values for each are reported in italic and bold italic, respectively.

Base	H1'	H2'	H2"	H3'	H4'	H5/M7/H2	H6-H8
C1	5.84 *5.83* ***5.73***	2.56	2.62	4.62 *4.89* ***5.06***	4.11	5.99	7.98
C2	6.03 *5.93* ***5.90***	2.42 *2.55* ***2.55***	2.69	4.77 *4.99* ***5.18***	4.24 *4.32* ***4.34***	5.64	7.78 *7.98* ***8.07***
T3	6.03 *5.91* ***5.87***	2.48	2.71	4.93 *5.05* ***5.19***	4.30	1.62 *1.73* ***1.71***	7.65
A4	6.23	2.57	2.87	4.92 *5.06* ***5.25***	4.43	7.04	8.13 *8.19* ***8.31***
T5	5.85	2.22	2.58	4.88 *5.02* ***5.18***	4.23 *4.27* ***4.34***	1.33	7.25
A6	6.00	2.64	2.77	5.00 *5.06* ***5.21***	4.41 ***4.53***	6.51	8.02
A7	6.16	2.52	2.80	4.82 *4.96* ***5.17***	4.41	7.65	7.88
T8	5.91	2.22 *2.32* ***2.34***	2.54	4.79 *4.91* ***5.05***	4.21	1.20	7.25
C9	6.05 *5.94* ***5.81***	2.25	2.56	4.73 *4.83* ***5.00***	4.25	5.44	7.53 *7.63* ***7.67***
C10	6.14	2.10	2.27	4.48	4.05 *4.14* ***4.17***	5.39	7.47 *7.65* ***7.75***

Table 1 b: Chemical shifts (ppm) for assigned protons in the normal hybrid RNA strand. If the chemical shift of the same proton differs by more than 0.1 ppm in the [Rp]-phosphorothioate or phosphorodithioate analogs, the values are for each reported in italic and bold italic, respectively. A dash denotes an unassigned peak. A blank space denotes an inapplicable assignment.

Base	H1'	H2'	H2"	H3'	H4'	H5/M7/H2	H6-H8
G11	5.40 ***5.31***	4.73		4.43	- *4.14*		7.78
G12	5.77	4.69		4.61	4.48		7.40
A13	5.93	4.59		4.52	- *4.67*	7.79	7.78
U14	5.42	4.38		4.39	4.43	5.04	7.59
U15	5.65	4.67		4.62	4.47	5.60	8.07
A16	5.92	4.45		4.61	4.55	6.91	8.18
U17	5.34	4.39		4.49	4.39	5.08	7.63
A18	5.98	4.65		4.72	4.47	6.88	8.08
G19	5.45	4.39		4.35	- *4.04* ***4.21***		7.13
G20	5.81	4.03		4.19	- *4.01*		7.28

Table 2: Linewidth measurements (Hz) for cross-resonances involving non-labile protons from the phosphorodithioate 2D NOESY dataset at 100 ms mixing time and T=20 °C. Each value is an average of all cross-peaks assigned for the corresponding proton. Average values for the DNA and RNA strands are reported in boldface.

Base	H1'	H2'	H2"	H3'	H4'	H6/H8	H5/M7/H2
C1	21.1	32.9	54.2	33.7	20.5	17.7	14.6
C2	27.0	52.0	40.1	34.8	-	22.4	14.2
T3	18.2	44.5	27.5	28.0	28.2	18.1	13.4
A4	15.5	43.8	37.5	30.6	13.8	12.8	10.7
T5	19.3	41.0	38.8	30.2	16.9	17.5	11.8
A6	18.1	36.6	43.0	31.3	14.7	16.1	9.9
A7	16.1	52.8	25.9	25.5	19.6	18.9	12.3
T8	17.0	52.9	40.8	23.7	22.7	16.5	12.9
C9	19.7	43.1	13.4	25.2	36.5	26.2	16.9
C10	16.2	37.8	36.5	17.5	24.5	25.4	15.7
DNA	**18.8**	**43.7**	**35.8**	**28.1**	**21.9**	**19.2**	**13.2**
Avg.	±3.4	±7.0	±11.1	±5.2	±7.2	±4.3	±2.2
G11	14.5	19.5		26.2	20.5	14.2	
G12	11.0	18.5		24.5	25.6	16.1	
A13	12.4	22.8		20.2	-	15.6	11.2
U14	14.9	35.0		29.6	-	22.1	16.1
U15	12.4	24.0		31.2	18.7	22.6	14.5
A16	12.6	31.4		32.5	30.3	15.7	9.3
U17	12.4	19.1		30.8	-	22.3	16.4
A18	12.6	16.1		24.7	23.3	15.5	10.4
G19	13.5	20.0		26.1	-	13.6	
G20	9.7	17.4		15.0	-	12.2	
RNA	**12.6**	**22.4**		**26.1**	**23.7**	**17.0**	**13.0**
Avg.	±1.5	±6.2		±5.4	±4.5	±3.9	±3.0

Table 3: Relaxation and correlation times for the phosphorodithioate. Relaxation times were obtained using inversion-recovery experiments for T_1 and spin-echo experiments for T_2. The six peaks used were well-isolated in 1D spectra, permitting accurate relaxation time measurements and a good estimate for the correlation time using equation (3).

Base Proton	T_1/ms	T_2/ms	τ_c/ns
A4H2	3.6	84	2.4
A6H2	3.9	77	2.6
A16H2	3.6	107	2.1
A18H2	4.3	94	2.5
T3CH3	1.8	59	2.0
T5CH3	1.7	76	1.7
			$\langle\tau_c\rangle=2.2\pm0.3$

Table 4: Comparison of normal, [Rp]-phosphorothioate, and phosphorodithioate interproton distances. Lists of interproton distances produced by RANDMARDI are compared by computing the rmsd (Å) between protons in two lists, their average relative difference, and the number of interproton distances differing by more than 10%. The reference corresponds to a comparison of 100 ms and 150 ms spectra of the same molecule (phosphorodithioate; see text).

Statistical Parameter	Reference Comparison	Normal vs. [Rp]	Normal vs. Dithioate	[Rp] vs. Dithioate
No. Common Distances	175	148	169	132
RMSD (Å)	0.31	0.22	0.18	0.25
Avg. Rel. Diff. (%)	7.7	4.94	4.13	5.87
>10%	57	11	7	24

Table 5: Experimental interproton distances (Å) for the unmodified hybrid. Average values for the DNA and RNA strand are given separately for each type of interproton distance, with canonical A- and B-form values tabulated for comparison. Intra-residue and inter-residue distances are labeled intra and inter, respectively. Standard deviations are added when statistically relevant. The RNA and DNA strands contain some distances typical of canonical A-form and B-form, respectively, such as the DNA H2"-base intra-residue and the RNA H3'-base distances. However, overall the average distances deviate distance from both A and B form references, suggesting a conformation that may be characterized as neither B-form, nor A-form.

Distance	DNA Avg.	RNA Avg.	A-DNA	B-DNA
H1'-H4' intra	2.9±0.2	3.1	3.3	2.8
H2'-base inter	2.8±0.3	2.6±0.1	2.0	3.9
H2'-base intra	2.7±0.2	3.3±0.3	3.8	2.2
H3'-base inter	3.5±0.3	2.9±0.2	3.0	4.6
H3'-base intra	3.1±0.3	2.8±0.1	2.8	4.5
H2"-base inter	2.9±0.2	-	3.8	2.1
H2"-base intra	3.3±0.5	-	4.5	3.5
H1'-base inter	3.3±0.3	3.3±0.1	4.6	3.5
H2-H2 inter	3.7	-	4.8	3.6
H5-H5 inter	-	3.5	3.7	4.7
H5-Me inter	3.4	-	3.3	4.7
Me-H5 inter	3.9	-	4.3	-

FIGURE CAPTIONS

Figure 1 - Sugar-phosphate backbone modifications examined in this study. The three molecules share the same unmodified phosphodiester RNA strand. Modifications (1b, 1c) occur on the DNA strand.

Figure 2 - The RNA•DNA hybrid sequence. The central Pribnow box sequence is flanked on the 5' and 3' ends by two G•C base-pair clamps.

Figure 3 - Hybrid melting profiles. Normalized line heights in 1D NMR spectra were measured using well-isolated A6H2 proton signals. Line heights are plotted a function of temperature for the unmodified (square), phosphorothioate (triangle) and phosphorodithioate (diamond) analogs. As the temperature exceeded the melting point (T_m), the A6H2 peak intensity decreased. Reducing the temperature below T_m, the intensity reached a plateau. Curves shown are polynomial fits to the data. Due to the variation in nucleic acid concentration, T_ms obtained by UV differ but, show the same overall trend in hybrid stability.

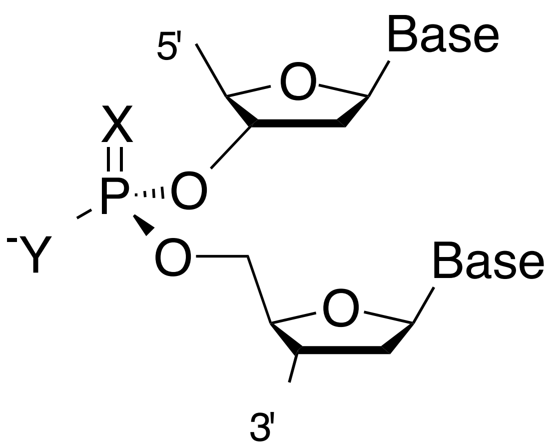
Figure 4 - The 100 ms NOESY spectrum for the phosphorodithioate-modified hybrid. The base-H1' region is displayed, with walks for the DNA strand outlined by a solid lines. The RNA strand walk is indicated by a dashed line. Seven strong H5-H6 cross-peaks also reside in this region. Since the DNA H3' protons for this particular phosphorodithioate molecule were strongly downfield-shifted towards the H1's, another DNA walk along the base-H3' protons can be observed. Isolated cross-peaks upfield from the base protons correspond to sequential or cross-strand H2'-H1' connectivities. Intra-residue ribose-base NOEs of terminal residues are labeled.

Figure 5 - Comparison of interproton distances as a function of base sequence. Interproton distances for the normal hybrid (squares), the [Rp]-phosphorothioate (triangles) and the phosphorodithioate (diamonds) are plotted. The MIXED model distance is included for comparison (X's). **A.** (top panel) The H1'-base interresidue distances are homogeneous along the whole molecule (DNA=3.3±0.3 Å, RNA=3.3±0.1 Å), suggesting B-form geometry for the hybrid duplex (the canonical B-form H1'-base distance is 3.5Å). The note that the MIXED A-form/B-form model

shows canonical A-type geometry in the RNA. **B.** (center panel) The H2'-base inter-residual interproton distances show a clear profile along the DNA strand. This profile may be correlated to a larger percentage of S-type deoxyriboses in the AT-rich center of the DNA strand relative to the terminal bases (see text). The MIXED model shows H2'-base sequential distances typical of A-type geometry. For the RNA strand, this distance significantly deviates from the canonical A-form (2.1Å). **C.** (bottom panel) The H3'-base intra-residue distances show a similar, though less pronounced, profile as for the H2'-base intra-residues, confirming the B-form preference for the DNA central bases.

Figure 6 - Residual R^X-factors. Conformational preferences of individual nucleotides are shown for the Pure-A (black bins), Pure-B (white bins) , and MIXED (gray bins) model conformations. On the x-axis, the number(s) of the residue(s) involved are followed by the number of experimentally observed distances used in the R^X estimation, given in parentheses. A good match has a low R^X value. **A.** (top panel) Intra-residual R^X factors, better-suited to assess sugar conformation, indicate clear preferences both for B-type DNA deoxyriboses and A-type RNA riboses. **B.** (bottom panel) Inter-residual R^X factors confirm the B-type preference for the DNA strand, but rule out the Pure-A conformation for the RNA strand, as expected from the preliminary analysis of interproton distances (see text).

Figure 1



#	Backbone Linkage	X	Y
1a	phosphate	O	O
1b	[Rp]-phosphorothioate	O	S
1c	phosphorodithioate	S	S

			1	2	3	4	5	6	7	8	9	10		
DNA	5'	-	**C**	**C**	**T**	**A**	**T**	**A**	**A**	**T**	**C**	**C**	-	3'
RNA	3'	-	**G**	**G**	**A**	**U**	**A**	**U**	**U**	**A**	**G**	**G**	-	5'
			20	19	18	17	16	15	14	13	12	11		

Figure 2

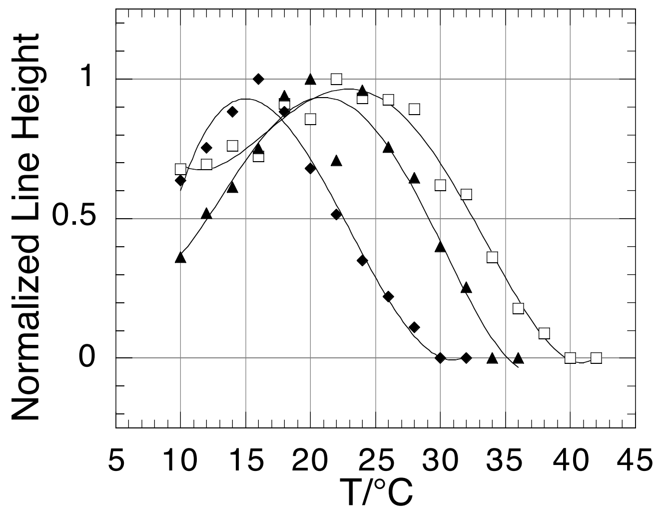
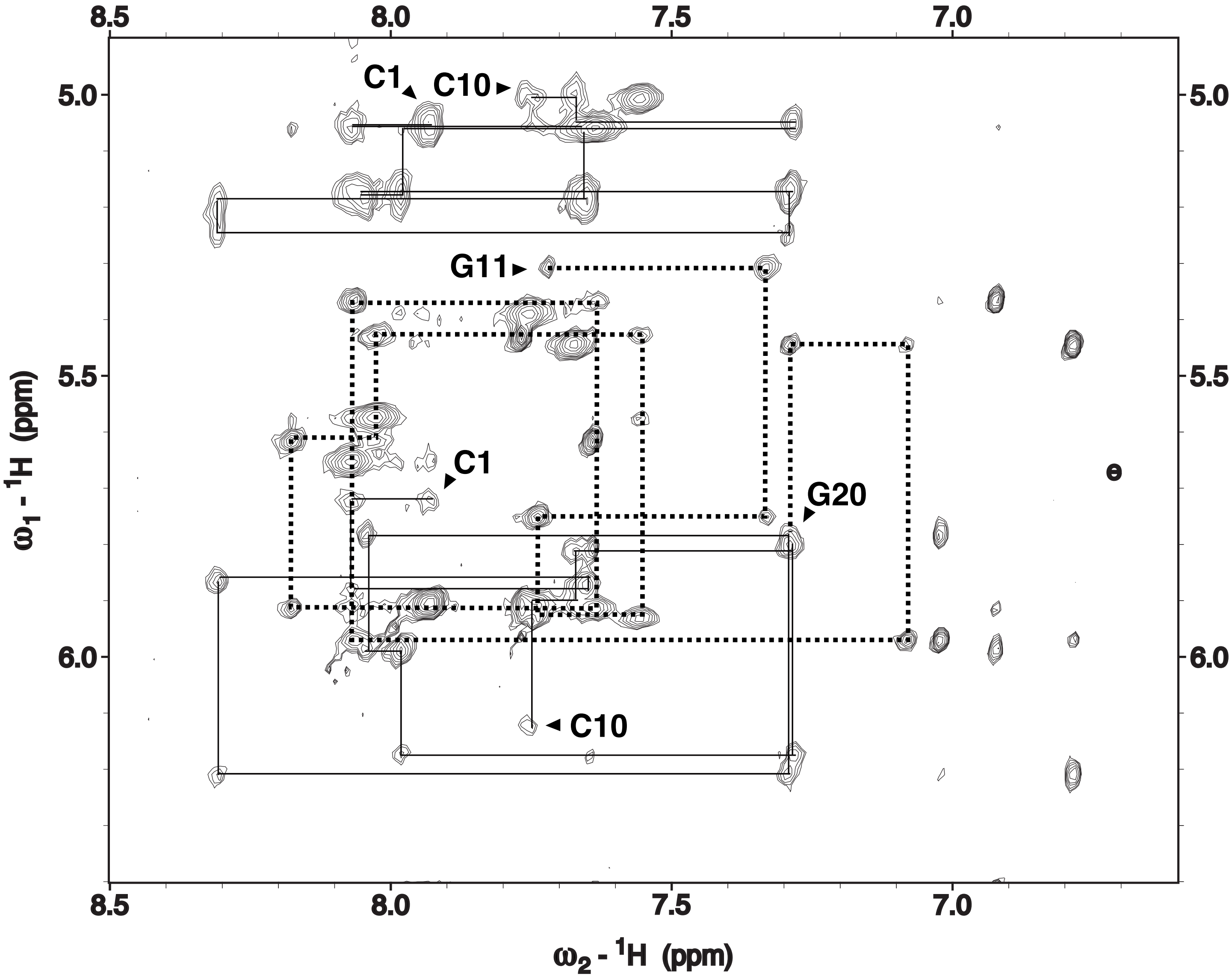
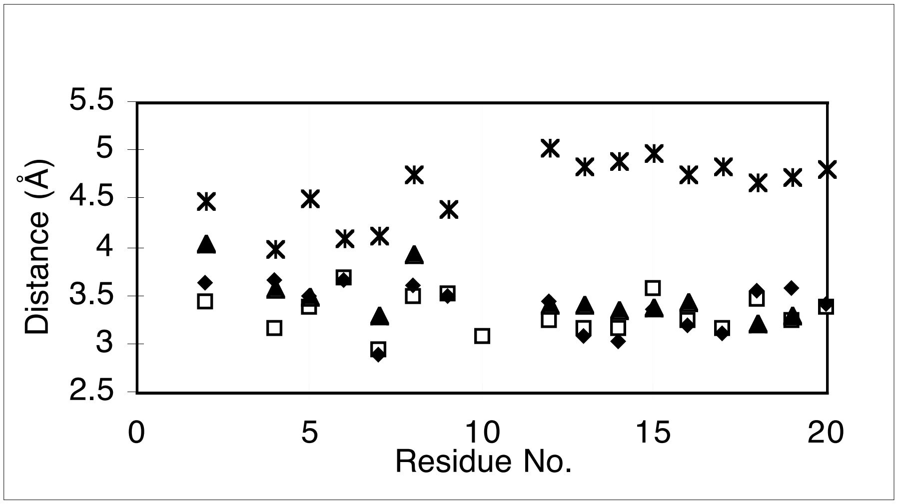
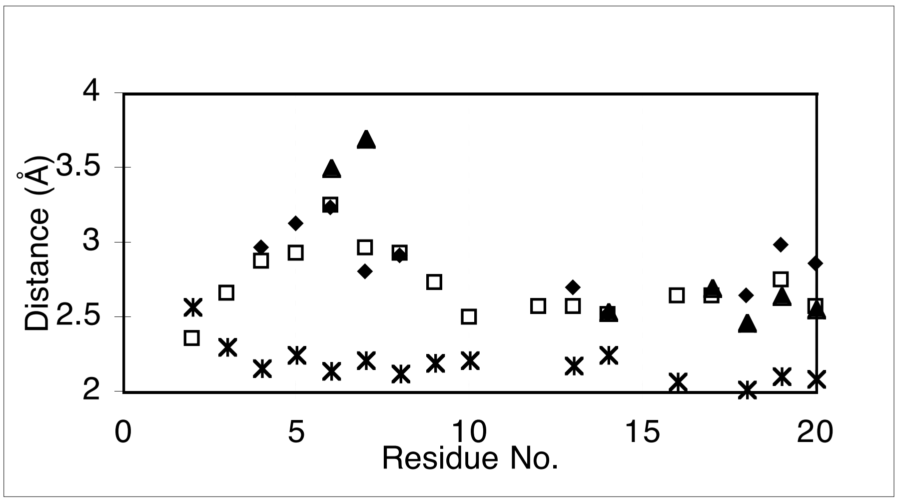


Figure 3

Figure 4




5.5
5
4.5
4
3.5
3
2.5
Distance (Å)
0
5
10
15
20
Residue No.


4
3.5
3
2.5
2
Distance (Å)
0
5
10
15
20
Residue No.

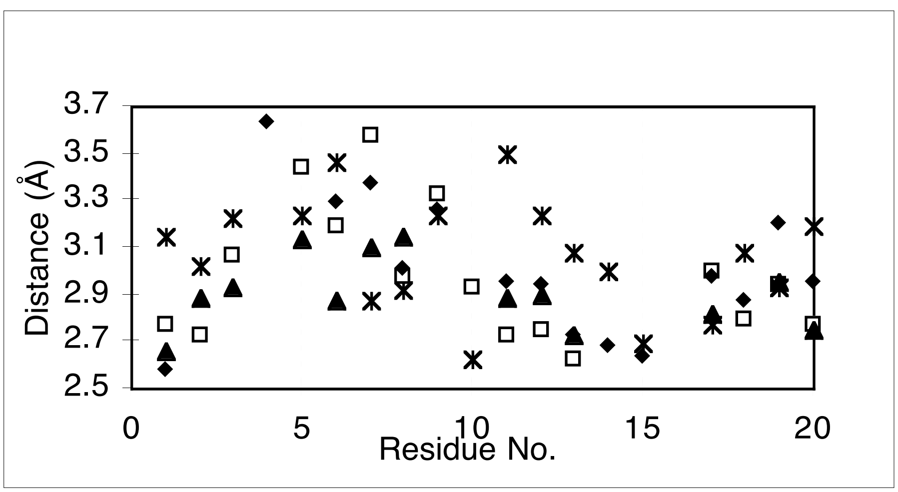
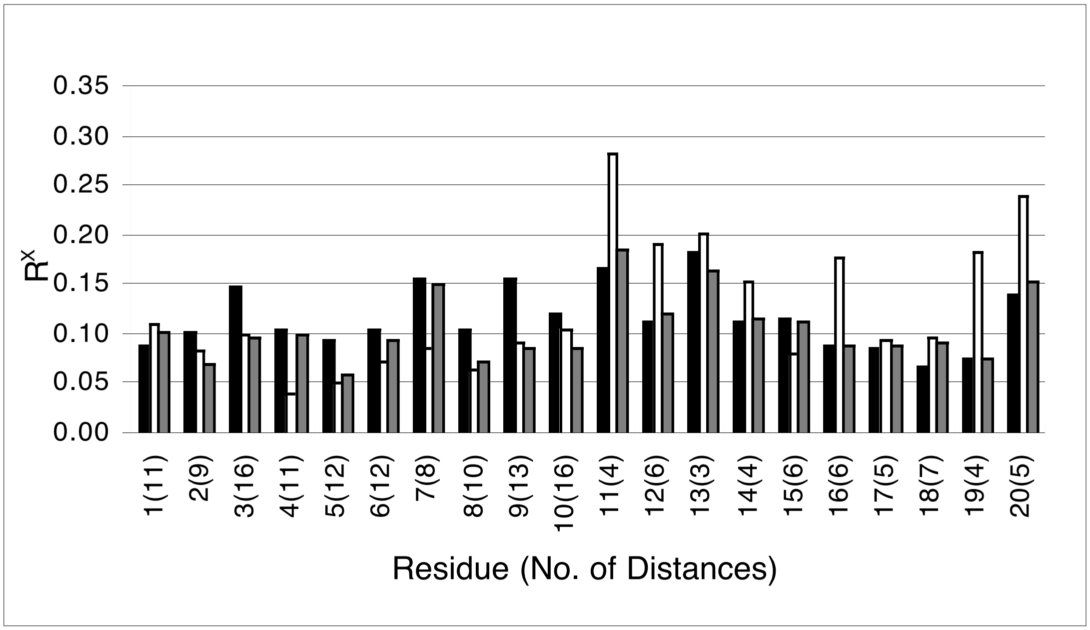

3.7
3.5
3.3
3.1
2.9
2.7
2.5
Distance (Å)
0
5
10
15
20
Residue No.

Figure 5


0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
R^X
1(11)
2(9)
3(16)
4(11)
5(12)
6(12)
7(8)
8(10)
9(13)
10(16)
11(4)
12(6)
13(3)
14(4)
15(6)
16(6)
17(5)
18(7)
19(4)
20(5)
Residue (No. of Distances)

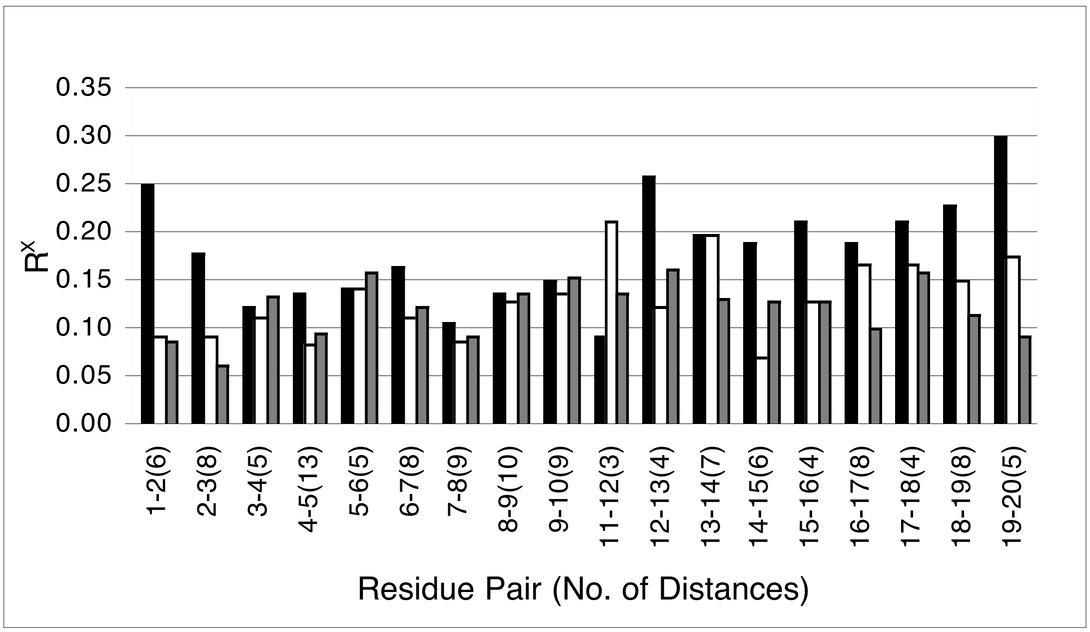

0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
R^X
1-2(6)
2-3(8)
3-4(5)
4-5(13)
5-6(5)
6-7(8)
7-8(9)
8-9(10)
9-10(9)
11-12(3)
12-13(4)
13-14(7)
14-15(6)
15-16(4)
16-17(8)
17-18(4)
18-19(8)
19-20(5)
Residue Pair (No. of Distances)

Figure 6