DNA Duplexes Flanked by Hybrid Duplexes: The Solution Structure of Chimeric Junctions in [r(cgcg)d(TATACGCG)]$_2$[†,‡]

Leiming Zhu,[§] Miguel Salazar,[§,||] and Brian R. Reid*[,§,⊥]

Departments of Chemistry and Biochemistry, University of Washington, Seattle, Washington 98195

Received August 26, 1994; Revised Manuscript Received December 2, 1994[⊗]

ABSTRACT: Hybrid duplexes and chimeric duplexes containing hybrid segments linked to pure DNA (or pure RNA) segments are involved in transcription and replication, as well as reverse transcription. A complete understanding of the mechanism of these processes requires detailed information on such duplexes and the junctions between duplexes of differing structure. Using two-dimensional NMR, restrained molecular dynamics and mechanics, and back-calculation refinement against the nuclear Overhauser effect spectra at various mixing times, we have determined the solution structure of the chimeric duplex [r(cgcg)d(TATACGCG)]$_2$ containing a pure DNA segment in the center of a hybrid duplex. The solution structure differs from the previously determined X-ray structure of the analogous duplex [r(gcg)d(TATACGC)]$_2$, which was found to be A-form throughout [Wang, A. H.-J., et al. (1982) *Nature 299,* 601−604]. The basic features of the solution structure are (a) the RNA residues are all A-form with C3′-endo sugar conformations, (b) the central DNA segment is B-form, (c) the transition from A-form RNA sugar conformations to B-form DNA sugar conformations involves only the dT5 base step, and (d) although the sugar conformations of the DNA residues A6−G12 are closer to B-form, the basic helical properties of the peripheral RNA·DNA hybrid segments are closer to typical A-form than to B-form.

Hybrid duplexes are formed during transcription when DNA directs the polymerization of RNA and during reverse transcription when RNA is the template for DNA polymerization. In lagging strand synthesis during DNA replication, Okazaki fragments are initiated by a short RNA primer containing a few ribonucleotides [reviewed by Ogawa and Okazaki (1980)]. These RNA primers are hydrogen-bonded to the DNA template via Watson−Crick base pairing to form a DNA·RNA hybrid duplex. DNA polymerase then covalently extends the 3′ end of the hybrid's RNA strand to form an Okazaki fragment consisting of a short segment of hybrid duplex covalently linked to a longer segment of pure DNA duplex.

In order to fully understand the processes of transcription, reverse transcription, and DNA replication, it is important to obtain a clear picture of both the structure of hybrid duplexes and the nature of the junction between DNA·RNA hybrids and DNA·DNA duplexes. Several laboratories have investigated the structure of hybrid duplex−DNA duplex chimeras by a variety of methods. Using CD, NMR, and optical melting experiments, Selsing and Wells (1979a) concluded that if [dG_n]·[dC_n] duplexes are assumed to be in a B-form conformation and [dG_i]·[rC_i] hybrid duplexes are assumed to be in an A-form conformation, then the hybrid chimeric duplex d[G_n]·[rC_{13} dC_{17}] should be B-form at the pure DNA end of the duplex and A-form at the hybrid end of the duplex. In subsequent modeling studies, Selsing and co-workers (1979b) further proposed that the burden of the transition from the hybrid segment to the DNA segment should be borne by the DNA·DNA base pair at the junction and not by the junction RNA·DNA base pair; they proposed a bend of 26° at the junction between the two segments. Furthermore, their modeling studies indicated that the junction should not be a kink, and that the transition from A-form to B-form should not significantly disturb base stacking or hydrogen-bonding of the Watson−Crick base pairs.

The structure of the (gcgTATACGC)$_2$ chimeric duplex (using lower case to designate ribonucleotides) has been solved by X-ray crystallography (Wang et al., 1982) and was found to adopt an A-form conformation throughout the duplex, with no structural discontinuity between the hybrid duplex segment and the DNA duplex segment, in the crystalline state. More recently, Egli et al. (1993) reported the crystal structures of the symmetrical chimeras (gCGTATACGC)$_2$ and (GCGTaTACGC)$_2$ and the nonsymmetrical chimeric duplex (gcgTATACCC)·(GGGTATACGC). They found that, regardless of the RNA content or position, these duplexes all adopt an A-form conformation in crystals. Thus, the crystalline-state results of Wang et al. (1982) and of Egli et al. (1993) differ completely from the modeling studies of Selsing et al. (1979a,b).

In solution studies using NMR, Mellema et al. (1983) investigated the sugar conformations of the same symmetrical (gcgTATACGC)$_2$ chimeric duplex studied by Wang et al. (1982) in the crystalline state. They found that, in solution, there is a change in sugar conformation at the junction between the hybrid duplex and the DNA duplex, in agreement with the modeling predictions of Selsing and co-workers (1979b). However, in contrast to the results of Selsing et al. (1979a,b), Mellema and co-workers (1983)

† This work was supported by NIH Grants GM42896 and GM32681 to B.R.R.

‡ The coordinates have been deposited with the Protein Data Bank (104D).

* Author to whom correspondence should be addressed.

§ Department of Chemistry.

|| Present address: Drug Dynamics Institute, The University of Texas at Austin, Austin, TX 78712.

⊥ Department of Biochemistry.

⊗ Abstract published in *Advance ACS Abstracts,* January 15, 1995.

0006-2960/95/0434-2372$09.00/0

showed that the sugars of the DNA strand of the hybrid segments never achieve an A-form conformation and are closer to a B-form conformation.

Benevides et al. (1986) examined the hybrid chimeric duplex $(gcgCGC)_2$ as a single crystal and in solution. They reported that, in the crystalline state, the duplex assumes an A-form conformation, but in solution the Raman data are consistent with at least two of the DNA residues being in the C2′-endo or B-form conformation. These results, and the results of Mellema et al. (1983), strongly suggest that the crystal structure of chimeric hybrid duplexes may be influenced by dehydration effects (Saenger et al., 1986) and/or crystal packing forces (Dickerson et al., 1994). Furthermore, Fedoroff et al. (1993) recently showed that hybrid duplexes in solution do not have a perfectly A-form structure but have a DNA strand that is neither A-form nor B-form, creating a unique minor groove that appears to explain the specificity of RNase H. Taken together, these studies underscore the need to reexamine the structure of hybrid-containing chimeras in solution.

In an effort to better understand the solution structure of chimeric duplexes with hybrid segments and the nature of the junction between RNA·DNA hybrids and DNA·DNA duplexes in Okazaki fragments, we have determined the three-dimensional structure of the double chimera duplex $(cgcgTATACGCG)_2$ by two-dimensional NMR, restrained molecular dynamics and mechanics, and back-calculation refinement against the NOESY spectra at various mixing times.

MATERIALS AND METHODS

Sample Preparation. The dodecanucleotide cgcgTATACGCG was synthesized on a 10-μmol scale on an automated DNA synthesizer (Applied Biosystems Model 392) by introducing the ribonucleoside phosphoramidites in the last four cycles, as described previously (Chou et al., 1989, 1991). The sample was deprotected and purified as described previously (Salazar et al., 1993a,b), lyophilized to dryness, and redissolved in 0.4 mL of buffer containing 20 mM sodium phosphate, 200 mM NaCl, and 0.8 mM EDTA at pH 7.0. The sample was then repeatedly lyophilized to dryness, first from the aqueous buffer and then from D_2O solution. Finally, 0.4 mL of 99.996% D_2O was added and the solution was transferred to a 5 mm NMR tube.

NMR Spectroscopy. All NMR data were collected on a home-built 500 MHz spectrometer at a sample temperature of 30 °C with a spectral width of 4386 Hz and the carrier frequency placed at the HDO resonance. NOESY spectra with mixing times of 60, 120, and 180 ms were acquired in the phase-sensitive hypercomplex mode (States et al., 1982) with a relaxation delay of 3 s. For each NOESY spectrum the mixing time was randomly varied within 10% of the designated mixing time to eliminate zero-quantum coherence transfer. For all NOESY experiments 2048 complex points in the t_2 dimension and 400 complex points in the t_1 dimension were collected, and 32 scans were averaged for each t_1 increment. The E.COSY spectrum (Griesinger et al., 1986, 1987) was collected in the TPPI mode (Drobny et al., 1979; Marion et al., 1983) with 2048 complex points in the t_2 dimension and 800 real points in the t_1 dimension. A total of 48 scans were averaged for each t_1 increment with a relaxation delay of 3 s between scans.

The NMR data were transferred to an IRIS-4D workstation and processed with FTNMR and/or FELIX software (Hare Research, Inc.). The NOESY data were zero-filled to 2048 points and apodized with a 90° phase-shifted sine-squared window function in both dimensions. The E.COSY data were zero filled to 2048 points in each dimension and processed with a Lorentz−Gauss transformation in the t_2 dimension with 6 Hz of exponential line narrowing and 6 Hz of Gaussian line broadening and a skewed-sinebell window function in the t_1 dimension. In both NOESY and E.COSY spectra, t_1 ridges were attenuated by multiplying the first t_1 slice by 0.5 (Otting et al., 1985).

Structure Determination. The structure determination protocol involves two basic steps. First, an interim structure is obtained via molecular dynamics and mechanics using initial NMR restraints. The interim structure is then iteratively refined until its theoretical NOESY spectra match the experimental NOESY spectra at each mixing time. These procedures are briefly outlined below.

Step 1. Determination of the Structure Using Molecular Dynamics with Initial NMR Restraints. Interproton distances were calculated for each resolved cross peak by measuring the initial build-up rates compared to standard cross peaks corresponding to known fixed H5−H6 distance of cytosine residues. A total of 458 distance restraints were obtained. The sugar conformations were determined by H1′−H4′ (24 constraints) and H1′−H2′/H2″ distances (40 constraints). *J*-coupling information was used only in a very qualitative way, and only to supplement the NOEs, since it should be pointed out that the use of *J*-coupling data to determine exact dihedral angles is questionable for slow-tumbling biopolymers, i.e., correlation times longer than 2−3 ns (Harbison, 1993; Zhu et al., 1994). The χ angles were determined by H6/H8−H1′ distances (46 constraints), H6/H8−H2′/H2″ distances (76 constraints), H6/H8-H3′ distances (38 constraints), H5−H1′/H2′/H2″ distances (12 constraints), and intraresidue H2−H1′ distances (4 constraints). Base−base stacking was constrained by (*n*)H6/H8−(*n*+1)H6/H8 (22 constraints) and (*n*)H6/H8−(*n*+1)H5 (six constraints) distance bounds. The relative orientation of the two strands was constrained using four cross-strand H2−H1′ distance constraints and 24 generous distance bounds (1.7−2.1 Å) for the hydrogen bond lengths. A total of 22 (*n*)H3′−(*n*+1)P distance constraints (2.8−3.2 Å) were obtained by line width analysis of the H3′ resonances, as described by Kim et al. (1992). Lower bounds for (*n*)H6/H8-(*n*)H5′/H5″ (48 constraints) and (*n*)H1′−(*n*,*n*+1)H5′/H5″ (92 constraints) distances were also included to further restrain the backbone conformation.

These distances were then used as initial distance restraints in determining the structure. The DISCOVER program (Biosym Technologies, Inc.) with AMBER potential functions was used to produce the structure from an approximate starting model by means of molecular mechanics and molecular dynamics. A distance-dependent dielectric constant of 4*r* was used, and no counter-ions were included in the calculations. Computations were carried out with a nonbonded cutoff distance of 12 Å. Two different canonical model structures were used as starting structures, one with a classical A-form conformation throughout (denoted "AAA") and the other consisting of a classical A-form conformation for the peripheral hybrid segments and a classical B-form conformation for the pure DNA segment (denoted "ABA").

Chart 1

```
      1   3   5   7   9   11
5'- c g c g T A T A C G C G
    G C G C A T A T g c g c - 5'
      23  21  19  17  15  13
```

Prior to the molecular dynamics simulations, the structures were first minimized using 2000 cycles of conjugate gradient minimization. The minimized structures were then refined by molecular dynamics at 300 K for 15 ps with a step size of 1 fs. Structures were archived every 1 ps, and the last five structures were minimized for 2000 cycles. The structure with the lowest energy out of these five was then chosen for further refinement by back-calculation of its NOESY spectra.

Step 2. Further Structure Refinement by Back-Calculation of the NOESY Spectra. The theoretical NOESY cross peak intensities were calculated from the interim structures using the NOESY simulation program BIRDER (Zhu & Reid, 1995). BIRDER is an improved NOESY simulation program that permits the use of faster pulse repetition rates by taking into account incomplete recovery of Z-magnetization and also allows for the possibility of differential external relaxation rates for different protons. The method to experimentally determine these external relaxation rates has been described previously (Zhu & Reid, 1995). The calculated NOESY cross peak intensities were compared with the experimental NOESY data, and the differences were used to generate a refined distance bounds file. The new set of distance bounds was then used in the next round of molecular mechanics and molecular dynamics, as described in step 1, to generate an improved structure with a lower *R* factor. The procedure was repeated several times until the *R* factor did not improve significantly. The earlier stages of refinement utilized only resolved cross peaks and backbone constraints. During the later stages of refinement more constraints were added to include overlapped cross peaks. Some lower bound distance restraints were also added for those cross peaks that were either undetectable or very weak. At this stage, the GNOE program (Hare Research, Inc.) was used for visual comparison of the back-calculated and experimental NOESY spectra.

To make sure conformational space was adequately sampled, the final structures obtained above were further heated to 1000 K and then cooled down to 700, 400, and 200 K for 5 ps at each temperature stage with a step size of 1 fs. The structures were finally minimized for 2000 cycles.

Table 1: Chemical Shifts (ppm) of Nonlabile Protons of [r(cgcg)d(TATACGCG)]$_2$ at 30 °C

residue	H6/H8	H5/H2/M5	H1′	H2′	H2″	H3′	H4′
c1	7.65	5.58	5.11	4.34		4.28	4.06
g2	7.59		5.57	4.44		4.52	4.31
c3	7.43	5.16	5.30	4.56		4.30	4.27
g4	7.36		5.53	4.31		4.30	4.27
T5	7.49	1.14	5.74	2.31	2.48	4.67	4.12
A6	8.07	7.09	6.13	2.43	2.76	4.78	4.26
T7	7.17	1.26	5.71	2.05	2.45	4.76	4.10
A8	8.00	7.07	6.03	2.47	2.69	4.85	4.26
C9	7.13	5.08	5.60	2.05	2.31	4.61	4.06
G10	7.42		5.77	2.34	2.50	4.72	4.19
C11	7.24	5.03	5.91	2.10	2.40	4.69	4.14
G12	7.52		5.98	2.27	2.30	4.48	4.03

RESULTS

Resonance Assignments. The numbering system for the (cgcgTATACGCG)$_2$ duplex is shown in Chart 1. All protons, except for the geminal H5′/H5″ pair, were assigned from the NOESY connectivities combined with E.COSY data, and the assignments are listed in Table 1. Figure 1 shows the sequential assignment of the expanded H6/H8 to H1′ and H6/H8 to H2′/H2″ regions of the NOESY spectrum at 120 ms.

External Relaxation Rates and Tumbling Correlation Times. The correlation times for parallel and perpendicular tumbling of this particular duplex were determined as described previously (Zhu & Reid, 1995): they are $\tau_1 = 1/6D_1 = 8.0$ ns and $\tau_2 = 1/6D_2 = 16.0$ ns, where D_1 and D_2

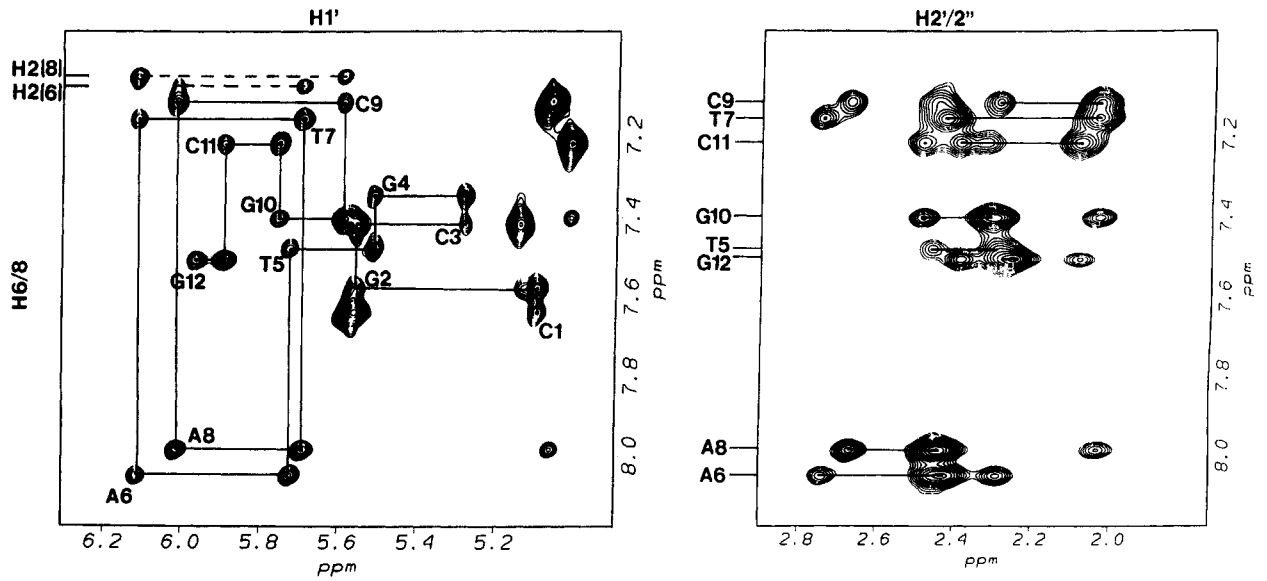


FIGURE 1: Expansion of the base to H1′ region (left) and the base to H2′/H2″ region (right) of the NOESY spectrum of (cgcgTATACGCG)$_2$ at a mixing time of 120 ms. In the base to H1′ region, the intraresidue H8/H6−H1′ cross peaks are labeled with their residue number. The horizontal dashed lines at the top connect the cross-strand (*n*)H2−(*m*+1)H1′ to the intrastrand (*n*)H2−(*n*+1)H1′. In the base to H2′/H2″ region, the horizontal lines connect the intraresidue NOEs between the H8/H6 protons and their own H2′ and H2″ protons.

Table 2: External Relaxation Rates for the Protons in the Sequence [r(cgcg)d(TATACGCG)]$_2$

sugar ring		base	
H1′ (RNA/DNA)	0.00/0.10	H6/8	0.00
H2′/2″ (RNA/DNA)	0.30/0.70	H2	0.10
H3′/4′ (RNA/DNA)	0.10/0.45	H5 (RNA/DNA)	0.50/0.70
H5′/5″	0.45	M5	0.00

represent the rotational diffusion coefficients parallel and perpendicular to the global helical axis, respectively. The empirically determined external relaxation rates for individual proton types are shown in Table 2. A brief explanation of why different protons have different external relaxation rates has been given previously (Zhu & Reid, 1995).

Structure Refinement and Back-Calculation of NOESY Spectra. The structure of the chimeric (cgcgTATACGCG)$_2$ duplex was obtained by iterative back-calculation refinement against the experimental NOESY spectra as described above. Table 3 lists the RMSD values, the NOE *R* factors and the energy terms for the final refined structures generated from both "AAA" and "ABA" starting models. Figures S1, S2, and S3 (Supplementary Material) show stack plots of the experimental and the back-calculated NOESY spectra at 60, 120, and 180 ms mixing times, respectively, for (a) the base to H1′, (b) the base to DNA H3′/H4′/H5′/H5″, and RNA H2′/H3′/H4′/H5′/H5″, and (c) the base to DNA H2′/2″ regions. Table 3 and Figures S1, S2, and S3 show that the final structures are well-converged and that the simulated NOESY spectra match the experimental ones very well.

Structural Features of the (cgcgTATACGCG)$_2$ Duplex in Solution. Figure 2 shows superimposed stereo views of the highly converged final four structures obtained from the "AAA" and "ABA" starting models, two of which were further annnealed at 1000 K and then gradually cooled down as described above. It is immediately obvious that there are bends in the helix axis at the junctions between the central pure DNA duplex segment and the flanking hybrid duplex segments. These bends at the two junctions lead to problems in defining a local helical axis from which to calculate helical parameters. Since there are three obvious segments to the overall helix, we chose local helical axes as follows: axis one includes base pairs rC1•dG24 through rG4•dC21, axis two includes base pairs dT5•dA20 through dA8•dT17, and axis three includes base pairs dC9•rG16 through dG12•rC13. These axes are shown schematically in Chart 2. In all three axes, the C1′ and N1 (for pyrimidine) or N9 (for purine) atoms were used to define the local helix axis. The helical parameters of the two hybrid segments and the central pure DNA segment were calculated separately with the program NEWHEL93 (kindly provided by Professor R. E. Dickerson, UCLA) based on these three local helical axes.

Figure 3 shows the backbone dihedral angles, the glycosidic torsion angles and the sugar conformations for the solution structure of the (cgcgTATACGCG)$_2$ duplex. For purposes of comparison, we also show the corresponding average values reported for the single crystal structure of the (gcgTATACGC)$_2$ duplex (Wang et al., 1982) and the values for B-form DNA, A-form DNA, and A-form RNA duplexes. The backbone angles in the solution structure indicate significant A-form structure, particularly in the α, β, γ, and ϵ torsion angles; the α and ϵ torsion angles all lie between the typical values for A-form DNA and A-form RNA, and the β and γ angles are closer to the typical values for A-form RNA and quite close to the values found in the X-ray structure of the (gcgTATACGC)$_2$ duplex (Wang et al., 1982).

Although the α, γ, and ϵ torsion angles appear to indicate an A-DNA type of structure for this duplex, and the β angle is A-RNA-like, the chimeric nature of this duplex is reflected in the δ and ζ torsion angles. The RNA residues, and the junction residue T5, adopt δ and ζ torsion angles closer to A-form while the DNA residues A6 through G12 have δ and ζ torsion angles that are between A-form and B-form geometries, but closer to B-form. The chimeric structure is also reflected in the χ angles. Figure 3 shows that, with the exception of the 5′-terminal RNA residue c1, the RNA χ torsion angles are A-form-like while the χ torsion angles for

Table 3: RMSD, NOE *R* Factor, and Energy Terms (in kcal) for Structures from "AAA" and "ABA"[a]

	"AAAf"	"ABAf"	"AAAh"	"ABAh"
RMSD with "AAAf"		0.59	0.31	0.52
RMSD with "ABAf"			0.51	0.15
RMSD with "AAAh"				0.48
RMSD with "ABAh"				
NOE *R* factor (*R*)[b]	16.8%	16.8%	16.8%	16.8%
NOE *R* factor (*R**)[b]	12.6%	12.6%	12.6%	12.6%
total energy	−91.03	−88.48	−90.62	−89.01
bond energy	28.97	28.97	29.10	29.04
angle energy	147.82	147.79	148.27	148.02
dihedral energy	257.94	258.75	258.10	257.99
H-bond energy	−1126.19	−1122.55	−1134.71	−1128.68
nonbond energy	−240.94	−241.63	−242.90	−242.85
Coulomb energy	841.37	840.14	851.52	847.46
NOE forcing potential	37.82	37.92	39.31	39.08

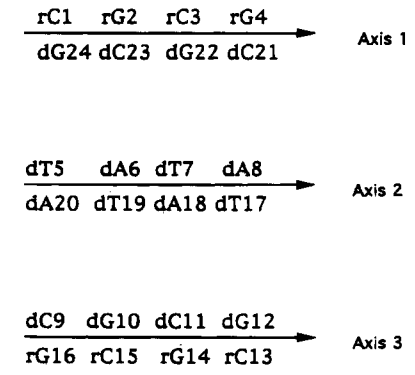
[a] "AAA" is the initial structure with classical A form throughout. "ABA" is the initial structure consisting of classical A form for the hybrid parts and classical B-form for the pure DNA part. The corresponding refined structures are denoted as "AAAf" and "ABAf". "AAAh" and "ABAh" are the structures after higher temperature (1000 K) anealing as described in the text. [b] *R* factors are defined as

$$R = \sum \frac{|I_c - I_e|}{I_e}$$

$$R^* = \begin{cases} \sum \frac{I_e^l - I_c}{I_e} & \text{if } I_c < I_e^l \\ 0 & \text{if } I_e^l < I_c < I_e^u \\ \sum \frac{I_c - I_e^u}{I_e} & \text{if } I_c > I_e^u \end{cases}$$

where I_e^l and I_e^u represent the lower and upper bounds of experimental NOE intensity, respectively. I_e is the average experimental NOE intensity. I_c is the back-calculated NOE intensity.

Chart 2



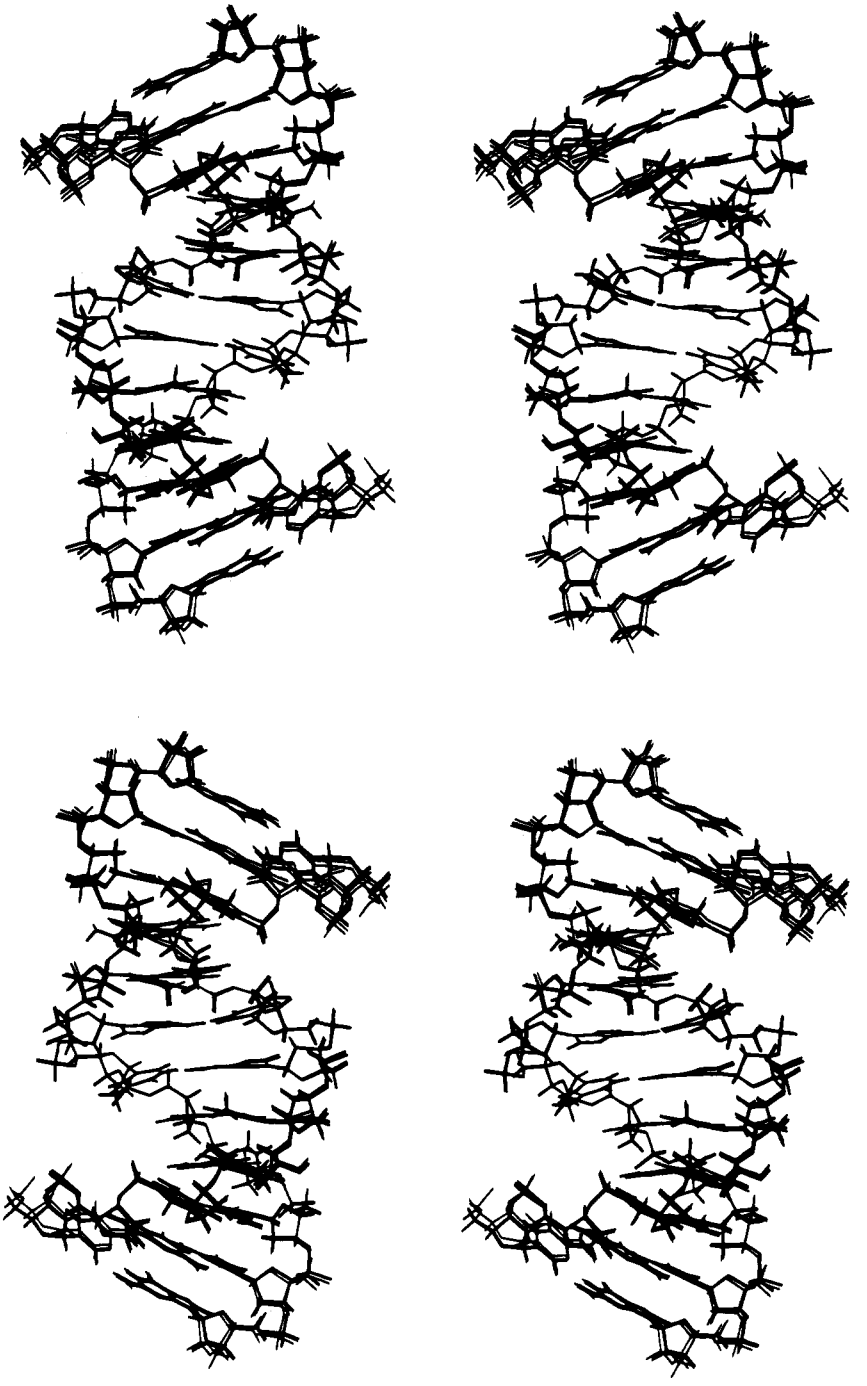


FIGURE 2: Cross-eye stereoviews into the minor groove (upper) and major groove (lower) of the superimposed AAAf, ABAf, AAAh, and ABAh final structures derived from refining different starting structures against the experimental data. The structures AAAh and ABAh were obtained by further annealing the structures AAAf and ABAf from 1000 to 200 K as described in the text. The four structures were overlaid using a best-fit RMS superposition of the heavy atoms.

the DNA residues are intermediate but closer to B-form DNA. The sugar conformations (*P* values) of the RNA residues are all C3′-endo, while the junction DNA residue T5 assumes a C4′-exo conformation. The DNA residues A6 through G12 assume sugar conformations between A-form and B-form but closer to B-form.

The helical parameter values also reflect the chimeric structure of the $(cgcgTATACGCG)_2$ duplex. Figure 4 shows plots of selected helical parameters calculated using the three local helical axes described in Chart 2. Except for buckle and propeller twist, all of these parameters are sensitive to the choice of local helical axes. The helical parameters in Figure 4 reflect the structural discontinuity in the solution structure at the point where the hybrid segment meets the pure DNA segment. This discontinuity is particularly striking for the parameters rise, twist, *x*-displacement, inclination, and buckle. Thus, the local structure of this duplex has striking chimeric features in that the central DNA segment assumes a B-type helical structure while the flanking DNA·RNA hybrid segments assume a structure closer to A-form. This is in complete contrast to the observations in the crystalline state. The structure of the peripheral hybrid segments is actually quite close to the H-form structure recently described for pure hybrid duplexes (Fedoroff et al., 1993).

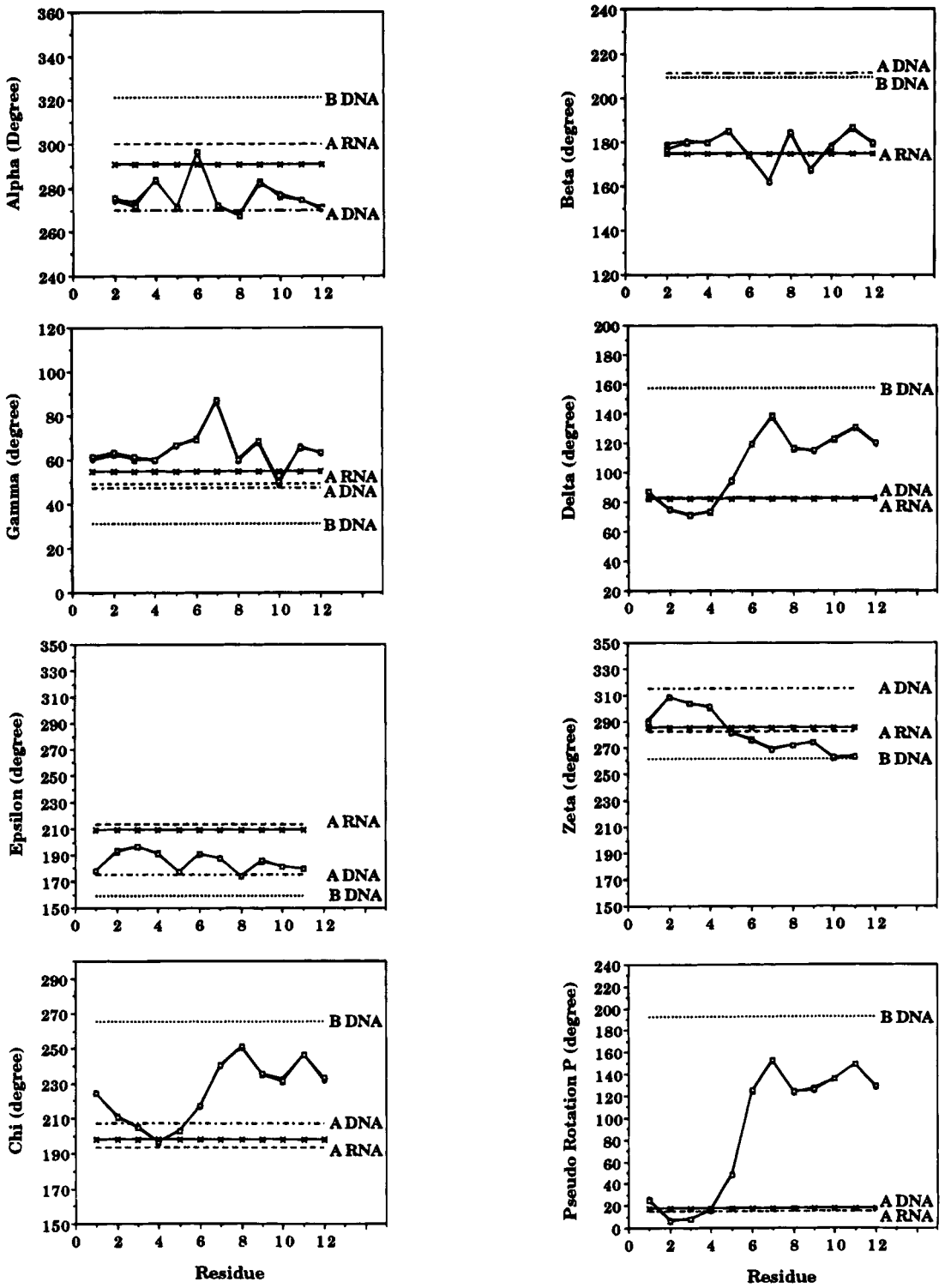


FIGURE 3: Backbone torsion angles and glycosidic torsion angles in the final refined (cgcgTATACGCG)$_2$ structures. The corresponding average values for the single crystal structure (Wang et al., 1982) are indicated by the (×) symbols, and typical values for B-DNA, A-DNA, and A-RNA are also shown for comparison.

DISCUSSION

The solution structure of (cgcgTATACGCG)$_2$ is strikingly different from the uniformly A-form duplex found in the crystalline state for the related duplex (gcgTATACGC)$_2$ (Wang et al., 1982). We attribute these differences to the effects of partial dehydration of the phosphate groups during crystallization (Saenger et al., 1986) and/or crystal lattice packing forces (Dickerson et al., 1994). In contrast to the A-form structure in the crystalline state, the solution structure of this three-segment chimera contains discrete peripheral sections with hybrid (H-form) geometry and a central section with B-form geometry. These results are similar to the chimeric structure found for the two-segment Okazaki fragment duplex (gcgTATACCC)·(GGGTATACGC) in solution (Salazar et al., 1994).

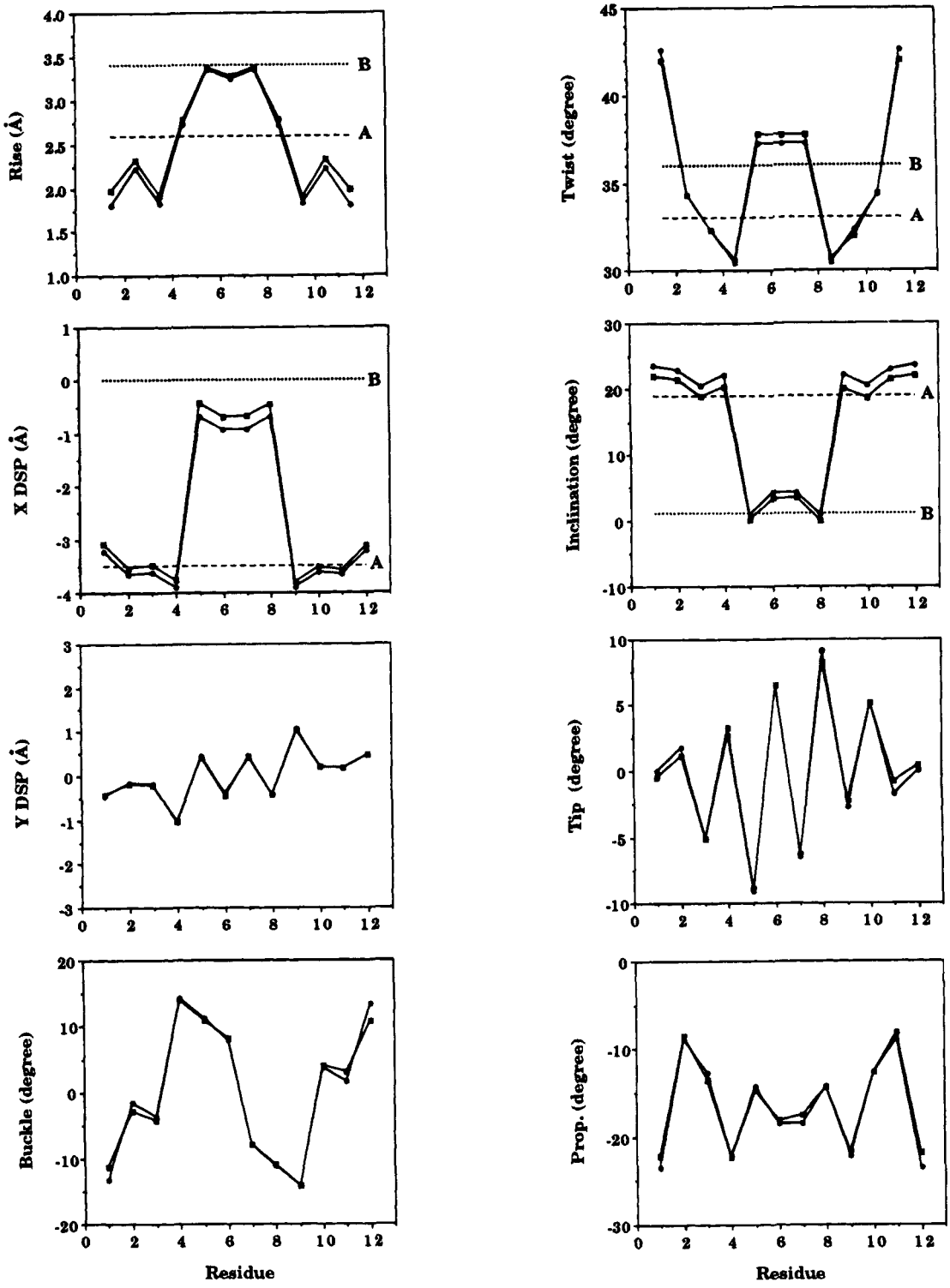


FIGURE 4: Helical parameters calculated using the three local helical axes described in Chart 2 for the refined final (cgcgTATACGCG)$_2$ structures. Typical rise, twist, *x*-displacement, and inclination values for standard A-form and B-form helix geometries are also shown in the top four plots.

The bend angle between the hybrid segment and the pure DNA segment, calculated from the three helical axes described in Chart 2, is ca. 23.5 ± 1° and the direction of the bend is into the major groove. This bend angle is close to the 26° value predicted for hybrid−DNA junctions in the modeling studies of Selsing et al. (1979b) and is slightly larger than the ca. 18.1° bend angle reported for the chimeric Okazaki fragment duplex (gcgTATACCC)•(GGGTATACGC) in solution (Salazar et al., 1994). However, it must be pointed out that the precise value for the bend angle depends on how the local helical axes are chosen. Nevertheless, the magnitude of H−B junction bending reported for this double chimera and in our previous single chimera study (Salazar et al., 1994) is very similar to that found in the modeling studies of Selsing et al. (1979a,b).

With respect to the hybrid duplex−DNA duplex junction, only the T5•A20 (or the symmetrical T17•A8) DNA base pair is involved in the structural transition from hybrid duplex

morphology to DNA duplex structure, with the T5 (T17) sugar taking on a C4′-exo conformation while the A20 (A8) sugar assumes a conformation in the C1′-exo range. Despite the presence of this highly heteromerous (by DNA standards) base pair, normal Watson−Crick hydrogen bonding is maintained. Relatively normal base stacking is also conserved, although it is worth noting that a quite high (negative) buckle value is observed between the junction base pairs g4•C21 (C9•g16) and T5•A20 (A8•T17). Interestingly, a similarly highly buckled (ca. −18°) structure was also observed for the junction residues A7•T14 and C8•g13 in both the solution structure and the crystal structure of the Okazaki fragment (gcgTATACCC)•(GGGTATACGC) (Egli et al., 1992, 1993; Salazar et al. 1994), although the overall structures differed markedly in solution and in the crystalline state.

The (n)P−(n+1)H3′ distance restraints used in the structure refinement are equally consistent with either a B_I or a B_{II} backbone phosphodiester conformation; however, the final structures consistently preferred a B_I backbone conformation ($\epsilon-\zeta \approx -90°$) throughout—presumably due to the combined effects of indirect proton−proton constraints between adjacent residues. Moreover, the one-dimensional proton-decoupled ^{31}P NMR spectrum (data not shown) showed no evidence of the downfield shifts associated with B_{II} phosphodiesters (Chou et al., 1992). These results, and the observation of relatively normal base stacking and base pairing, are in agreement with the modeling predictions of Selsing et al. (1979b), namely, that the transition from the hybrid section to the DNA section should involve only one base pair without a kink at the junction.

For this particular sequence in this type of duplex, the RNA residues influence the conformation of the downstream (3′ side) DNA sugars more strongly than they influence their partner's conformation transversely through base pairing. The same appears to be true for the glycosidic torsion angle χ. At the end of the four ribonucleotides with low A-type χ angles, T5 has a transition χ value of ca. 220°, and it is not until A6 that B-like χ values in the 240−260° range are attained. The sugar conformations in the two hybrid sections of this duplex are more heteromerous than in the simple DNA•RNA hybrid duplex (GTCACATG)•(caugugac) (Salazar et al., 1993; Fedoroff et al., 1994), particularly for the g2•dC23 base pair. These slight differences may simply reflect the sequence dependence of sugar conformations in DNA•RNA hybrid duplexes. Nonetheless, the general markedly heteromerous nature of the sugar conformations in hybrids is conserved in the hybrid segments of this chimeric duplex.

A structural feature of hybrid−DNA chimeric duplexes that should also reflect the discontinuity at the RNA−DNA junction, and may be of particular biological importance, is the minor groove width. The width of the minor groove (cross-strand phosphate−phosphate distances) along the chimeric (cgcgTATACGCG)$_2$ duplex is plotted in Figure 5; these distances are independent of the selection of helical axes. The central DNA section of this duplex has a strikingly narrow minor groove width of ca. 5 Å in the middle of the TATA tract. This feature is reflected in the surprisingly strong cross-strand H2−H1′ NOESY cross peaks which are at least as strong as the H8/H6−H1′ intensities (see the upper left cross peak in Figures 1 and S1-S3). A narrow minor groove of ca. 5 Å is typical of B′-DNA rather than B-DNA,

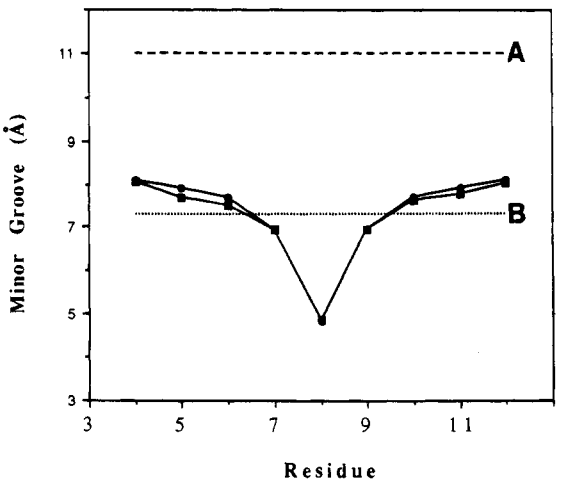


FIGURE 5: Minor groove width as a function of position along the (cgcgTATACGCG)$_2$ duplex. The width is represented by the (n)P−(m+3)P phosphorus−phosphorus distance minus the phosphorus van der Waals radius of 5.8 Å. The values for standard A-form and B-form DNA are plotted for comparison.

where the groove is typically ca. 7 Å (Chuprina et al., 1991), and the overnarrowing is probably a consequence of the flanking hybrid structure on both sides. The minor groove width then widens to greater than 7 Å in going from the pure DNA section to the hybrid section, reaching a value of ca. 8 Å at the junction. Unfortunately, the hybrid segment of this duplex is not long enough to fully establish the width of the hybrid segment minor groove more than one base pair from the junction, since the (n)P-(m+3)P distance is usually used to measure the groove width. However, in an alternating sequence such as this, we predict that the width of the minor groove in a longer flanking hybrid section should approach the values of 8.6−9.0 Å observed for nonchimeric random sequence hybrid duplexes (Fedoroff et al., 1993).

The minor groove width, perhaps together with the magnitude of buckle at the junction (if this feature is found to be conserved in other hybrid−chimeric duplexes), may determine the site of cleavage by RNase H. This enzyme is known to cleave these types of junctions in a sequence-specific manner (Pullen et al., 1992) that may simply reflect the sequence dependence of the minor groove width across the transition from H-form to B-form. Additional structural work with longer flanking hybrid segments is needed in order to determine the effects of sequence on the structure at the junction between the hybrid and DNA segments.

ACKNOWLEDGMENT

We thank Dr. Oleg Fedorov for help and discussions during the early stages of this work.

SUPPLEMENTARY MATERIAL AVAILABLE

Three figures showing experimental and back-calculated NOESY spectra at 60, 120, and 180 ms mixing times (4 pages). Ordering information is given on any current masthead page.

REFERENCES

Benevides, J. M., Wang, A. H.-J., Rich, A., Kyogoku, Y., van der Marel, G. A., van Boom, J. H., & Thomas, G. J., Jr. (1986) *Biochemistry 25*, 41−50.

Chou, S.-H, Flynn, P., & Reid, B. R. (1989) *Biochemistry 28*, 2435−2443.
Chou, S.-H., Flynn, P., Wang, A., & Reid, B. R. (1991) *Biochemistry 30*, 5248−5357.
Chou, S.-H., Cheng, J.-W., & Reid, B. R. (1992) *J. Mol. Biol. 228*, 138−155.
Chuprina, V. P., Lipanov, A. A., Fedoroff, O. Yu, Kim, S.-G., Kintanar, A., & Reid, B. R. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 9087−9091.
Dickerson, R. E., Goodsell, D. S., & Neidle, S. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 3579−3583.
Drobny, G., Pines, A., Sinton, S., Weitekamp, D. P., & Wemmer, D. E. (1979) *Faraday Symp. Chem. Soc. 13*, 49−55.
Egli, M., Usman, N., Zhang, S., & Rich, A. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 534−538.
Egli, M., Usman, N., & Rich, A. (1993) *Biochemistry 32*, 3221−3237.
Fedoroff, O. Yu., Salazar, M., & Reid, B. R. (1993) *J. Mol. Biol. 233*, 509−523.
Griesinger, C., Sørensen, O. W., & Ernst, R. R. (1986) *Chem. Phys. 85*, 6837−6852.
Griesinger, C., Sørensen, O. W., & Ernst, R. R. (1987) *J. Magn. Reson. 75*, 474−492.
Haasnoot, C. A. G., Westerink, H. P., van der Marel, G. A., & van Boom, J. H. (1983) *J. Biomol. Struct. Dyn. 1*, 131−149.
Harbison, G. S. (1993) *J. Am. Chem. Soc. 115*, 3026−3027.
Kim, S.-G., Lin, L.-J., & Reid, B. R. (1992) *Biochemistry 31*, 3564−3574.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967−974.
Mellema, J.-R., Haasnoot, C. A. G., van der Marel, G. A., Willie, G., van Boeckel, C. A. A., van Boom, J. H., & Altona, C. (1983) *Nucleic Acids Res. 11*, 5717−5738.
Ogawa, T., & Okazaki, T. (1980) *Annu. Rev. Biochem. 49*, 421−457.
Otting, G., Widmer, H., Wagner, G., & Wuthrich, K. (1985) *J. Magn. Reson. 66*, 187−193.
Pullen, K. A., Ishimoto, L. K., & Champoux, J. J. (1992) *J. Virol. 66*, 367−373.
Saenger, W., Hunter, W. N., & Kennard, O. (1986) *Nature 324*, 385−388.
Salazar, M., Champoux, J. J., & Reid, B. R. (1993a) *Biochemistry 32*, 739−744.
Salazar, M., Fedoroff, O. Yu., Miller, J. M., Ribeiro, N. S., & Reid, B. R. (1993b) *Biochemistry 32*, 4207−4215.
Salazar, M., Fedoroff, O. Yu., Zhu, L., & Reid, B. R. (1994) *J. Mol. Biol. 241*, 440−455.
Selsing, E., & Wells, R. D. (1979a) *J. Biol. Chem. 254*, 5410−5416.
Selsing, E., Wells, R. D., Alden, C. J., & Arnott, S. (1979b) *J. Biol. Chem. 254*, 5417−5422.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286−292.
Wang, A. H.-J., Fujii, S., van Boom, J. H., van der Marel, G. A., van Boeckel, C. A. A., & Rich, A. (1982) *Nature 299*, 601−604.
Zhu, L., & Reid, B. R. (1995) *J. Magn. Reson., Ser. B* (in press).
Zhu, L., Reid, B. R., Kennedy, M., & Drobny, D. P. (1994) *J. Magn. Reson., Ser. A, 111*, 195−202.

BI942008P