doi:10.1006/jmbi.2000.3931 available online at http://www.idealibrary.com on IDEAL® *J. Mol. Biol.* (2000) **300**, 1143–1154

JMB

Solution Conformation of a Bulged Adenosine Base in an RNA Duplex by Relaxation Matrix Refinement

Varatharasa Thiviyanathan[1], Anton B. Guliaev[2], Neocles B. Leontis[2] and David G. Gorenstein[1]*

[1]*Sealy Center for Structural Biology and Department of Human Biological Chemistry and Genetics, University of Texas Medical Branch Galveston, TX, 77555-1157, USA*

[2]*Department of Chemistry and Centre for Biomoleculor Sciences, Bowling Green State University, Bowling Green OH 43403-0213, USA*

**Corresponding author*

Bulges are common structural motifs in RNA secondary structure and are thought to play important roles in RNA-protein and RNA-drug interactions. Adenosine bases are the most commonly occurring unpaired base in double helical RNA secondary structures. The solution conformation and dynamics of a 25-nucleotide RNA duplex containing an unpaired adenosine, r(GGCAGAGUGCCGC): r(GCGGCACCUGCC) have been studied by NMR spectroscopy and MORASS iterative relaxation matrix structural refinement. The results show that the bulged adenosine residue stacks into the RNA duplex with little perturbation around the bulged region. Most of the bases in the RNA duplex adopt C_3'-*endo* conformation, exhibiting the N-type sugar pucker as found in the A form helices. The sugars of the bulged residue and the 5′ flanking residue to it are found to exhibit C_2'-*endo* conformation. None of the residues are in *syn* conformation.

Keywords: bulged RNA; MORASS; relaxation matrix refinement; bulged adenosine; kinked RNA

Introduction

Bulge-containing RNA structures are found to occur in biologically significant stable RNA molecules, including large and small ribosomal RNAs, catalytic RNAs and genomic RNA of certain viruses. Bulged bases in RNA helices are potentially significant in RNA folding, and are thought to play important roles in specific protein and drug binding to the RNA and RNA-RNA interactions. For example, the bulged adenosine base in helix II of *Escherichia coli* 5 S ribosomal RNA is important for the binding of ribosomal protein L18 (Christiansen *et al.*, 1985), and a bulged adenine base is essential for the binding of the viral coat protein to the bacteriophage R17 genomic RNA (Borer *et al.*, 1995). Solution structure of a drug-RNA complex shows that the aminoglycoside antibiotic paromomycin binds to the target RNA to a pocket created by an A:A mismatch and a single bulged adenosine base (Fourmy *et al.*, 1996). Bulged nucleotides have been suggested to facilitate conformational transitions in double-stranded RNA helices (Leontis & Moore, 1986; Wimberly *et al.*, 1993). Bulge-induced kinks in RNA helices are significant in the tertiary folding of RNA (Golke *et al.*, 1994; Zacharias & Hagerman, 1995). These kinks provide potential recognition sites for the binding of proteins and drugs. Chemical probing, crystal studies, and NMR experiments in solution of bulged RNA duplexes show that bulges cause widening of minor grooves (Gohlke *et al.*, 1994; Werner *et al.*, 1995; Portman *et al.*, 1996). A conserved bulged adenosine molecule in the hepatitis delta virus anti-genomic self-cleaving RNA has been shown to increase the rate of refolding, from an inactive to an active ribozyme structure (Perrotta *et al.*, 1999).

Several experiments, including X-ray crystallographic and NMR studies, have been performed to address the preferred conformations of bulged residues in DNA and RNA duplexes. The naturally occurring frame-shift mutations and the issues of how cells repair and recognize such structures have prompted interest in the fine structure of bulged bases in DNA. NMR studies of DNA duplexes containing a bulged adenosine base have revealed that the extra adenosine stacks into the duplex (Kalnik *et al.*, 1989; Nikonowicz *et al.*, 1989;

Abbreviations used: MD, molecular dynamics; rMD, restrained MD; MORASS, multiple Overhauser relaxation analysis and simulation; NOESY, nuclear Overhauser effect spectroscopy; TOCSY, total COSY.

E-mail address of the corresponding author: david@nmr.utmb.edu

0022-2836/00/051143–12 $35.00/0

1990; Rosen *et al*., 1992). In the genomic RNA of bacteriophage R17 a bulged adenosine base is stacked into the RNA duplex (Borer *et al*., 1995). However, X-ray crystallographic studies on both DNA (Joshua-Tor *et al.*, 1988, 1992) and RNA (Cate *et al*., 1996; Valegard *et al*., 1994; Ennifar *et al*., 1999) duplexes revealed that the bulged adenosine is looped out of the helix. In the reported X-ray crystal structure of a 13-mer DNA duplex, bulged A is stacked into the helix in one strand while it is looped out in the opposite strand (Joshua-Tor, *et al*., 1992). Energy-based conformational analysis of single-base bulges in RNA also showed that a stacked conformation is the energetically most favorable form for adenosine bulge (Zacharias & Sklenar, 1999).

The secondary structure of the RNA duplex containing the bulged adenosine base described here is shown schematically in Figure 1. The G-C rich sequence is defined in such a way that alternate base-pairing is not available for the bases on either strands, and bulge migration is not expected to occur in this sequence. G · C base-pairs are placed both at the end of the duplex as well as flanking the bulged base. Two A · U base-pairs placed on each side of the bulge serve as spectroscopic markers.

Results

Assignment of non-exchangeable proton resonances

Sequential assignments were based on the aromatic-H1′ connectivities and were confirmed by the aromatic-aromatic cross-peaks and aromatic-H2′/H3′ connectivities (Wüthrich, 1986; Varani & Tinoco, 1991). The conventional aromatic H8/H6-anomeric H1′ NOE connectivities were observed in both strands and are shown in Figure 2. Although neither an *A* form nor *B* form geometry was initially assumed, the sequential assignment of this duplex followed that for a right-handed helix.

The pyrimidine bases were identified by the presence of strong intra-residue H6-H5 NOESY cross-peaks and by comparing the NOESY spectra with the TOCSY spectra. With the exception of the H5/H6 cross-peaks of C18 and C24, all intra-residue pyrimidine cross-peaks were well resolved. Additional inter-residue NOESY cross-peaks from the H5 of pyrimidines to the H6/H8 base proton of the $i-1$ residue were also present. Inter-strand NOE connectivities for the H2 protons of A4, A6, and A19 were also present. The adenosine H2 resonances were assigned by their strong cross strand NOEs to base-paired uridine imino (H3) protons, and by their sequential NOE to the H1′ of the 3′ residue. Weak sequential H1′(n) → H1′($n+1$) connectivities were also observed, and were used to confirm the H1′ assignments. This connectivity was not observed between C20-H1′ and C21-H1′. Sugar H2′ resonances were assigned by comparing the intensities of intra-sugar NOEs. The sugar H2′ proton is closer to the H1′ of the same sugar than any other protons, irrespective of the sugar pucker. Sequential H2′(n) → H1′($n+1$) NOE peaks were observed along both strands. This distance is approximately 4 Å in RNA helices. These sequential H2′(n) → H1′($n+1$) NOEs thus provide a useful check of the assignment of the H2′ protons. A short mixing time (50 ms) NOESY was also helpful in assigning the H2′ protons. Sugar H2′ protons of the 5′ terminal residues resonate downfield from most of the H2′s, probably due to the absence of ring currents from the 5′ residue bases. The H2′s of the 3′-terminal residues are shifted upfield compared to other sugar H2′ protons, probably due to the absence of phosphate groups at the 3′ ends. *A* form RNA duplex NOE connectivities were observed for all the bases in both strands, except for C13 and C25 (the bases at the ends of the helices), and for G5 and A6. In *A* form helices, the sequential H2′ to H8/H6 NOE cross-peak is stronger than the intra-residue peak.

Although partly overlapping, the other sugar protons (H3′, H4′, H5′ and H5″) were assigned from the 2D TOCSY spectra. Assignment of H3′ resonances were confirmed by their medium-intensity NOEs to H8/H6 of their own residues and the sequential residue in the 3′-direction. A $^{1}H/^{31}P$ heteronuclear TOCSY experiment (Kellog, 1992) was carried out to help with the assignment of H3′ and H4′ protons. Due to the poor dispersion of the phosphorus resonances (within a range of 0.7 ppm), connecting the H3′ and H4′ protons from one base to the next was not possible. Intra-sugar NOEs were also used to confirm the assignments of the sugar protons. The H1′-H4′ distance is shorter than the H1′-H3′ distance, irrespective of the sugar pucker. The H4′-H5′/H5″ scalar cross-peaks were either overlapped or too weak to be seen in the TOCSY, due to the small coupling constants (less than 3 Hz). Stereospecific assignments for the H5′/H5″ were not made. It is assumed that the H5′ protons resonate downfield from the H5″ (Varani & Tinoco, 1991).

Sequential aromatic-aromatic cross-peaks were observed for all the residues in the first strand (residues G1 through to C13) with the exception of the U8-G9, because of the almost identical chemical shift of the U8-H6 and G9-H8 resonances. In the second strand (residues G13 through C25), sequential aromatic-aromatic cross-peaks were seen for residues G14 through C24, including a weak NOE

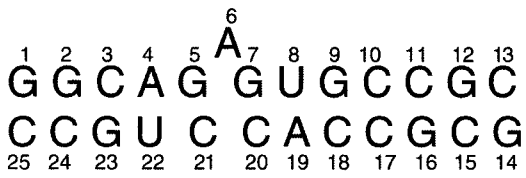

Figure 1. Sequence of the RNA duplex with the bulged adenosine.

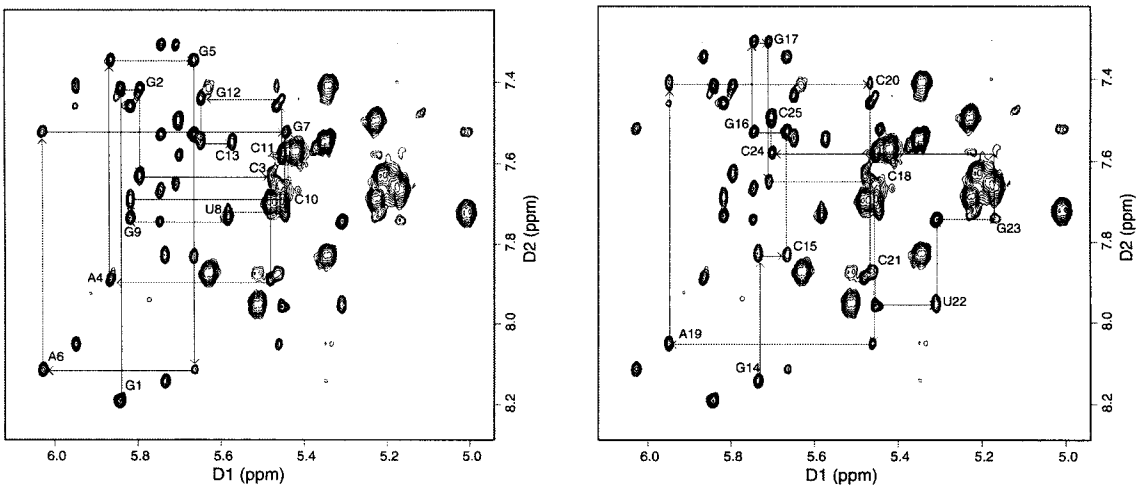

Figure 2. Sequential aromatic-H1′ NOESY connectivities for the (a) top and (b) bottom strands.

peak between C20-H6 and C21-H6, opposite the bulged adenosine, A6.

Exchangeable imino protons

Imino proton spectra of the RNA duplex in water were collected as a function of temperature. The two downfield peaks, assigned to U22 and U8, broaden and disappear first from the spectrum as the temperature increases. The imino resonances corresponding to G7 and G5, the bases flanking the bulged residue, also disappear at lower temperatures than other Gs. The resolved imino protons broaden in a sequential manner on raising the temperature and have been assigned by one and two-dimensional NOE experiments. Uridine imino protons were identified by intense cross-peaks in the imino-aromatic region of the NOESY spectra, which corresponds to the interaction between the uridine imino and H2 proton of adenosine to which it is paired. As expected, two intense peaks were observed corresponding to A4/U22 and U8/A19 base-pairs. The uridine imino protons served as starting points for the sequential assignment of imino and amino protons. For the assignment of the imino region, we also used imino guanosine to amino cytosine to H5/H6 cytosine walk. The proton chemical shifts are deposited at the Bio MagResBank (http://bmrb.wisc.edu) (Accession number: 4745) and also available in Table S1 (supplementary material).

Sugar puckering and glycosidic torsion angles

The absence of most H1′-H2′ cross-peaks in the TOCSY spectra indicates that most of the bases are adopting predominantly N-type sugar puckering as found in *A* form helices. However the bulged adenosine residue (A6) and the guanosine 5′ to it (G5), show H1′-H2′ scalar cross-peaks, indicating that these residues adopt C_2'-*endo* conformation. Previous observations, both in NMR and in crystals, show that in bulged RNA duplexes the sugars of the bulged residue and the 5′ flanking residue often exhibit C_2'-*endo* conformation (Wimberly *et al.*, 1993; Borer *et al.*, 1995; Portmann *et al.*, 1996). Theoretical calculations also resulted in the same conclusion (Zacharias & Sklenar, 1999). The residues at each end of the helix also showed weak H1′-H2′ scalar peaks, indicating a mixed C_3'-endo and C_2'-*endo* sugar puckering for these bases due to dynamics at the ends of the helix. The inter-residue intensities between H8/H6 to H3′ are greater than H8/H6 to H2′, indicating the glycosidic torsion angles are *anti*. None of the residues appear to be in *syn* conformation, since the intensity of the inter-residue H8-H1′ NOE cross-peaks are weaker relative to the intensity of the intra-residue H5-H6 cross-peaks of cytosines.

Evidence for helical stacking of bulged adenosine

Evidence of the stacked nature of the bulged adenosine residue is first presented on the basis of an observed downfield shift of the H8 and H2 protons of the bulged A6 with the increase in temperature (Nikonowicz *et al.*, 1990). Above the melting temperature, base-pairing and stacking interactions are disrupted in the duplex. Aromatic proton chemical shifts are sensitive to the degree of stacking, and the observed change in the chemical shifts of A6 protons is due to the disruption of stacking upon melting the RNA duplex.

The observation of sequential NOEs between G5 sugar protons (H1′/H2′/H3′) and the A6-H8 proton and the NOEs between A6 sugar protons and G7-H8 indicates that the bulged A6 residue is stacked into the duplex. Moreover, sequential aromatic-aromatic cross-peaks were observed for G5-A6-G7. However no cross-peak was observed between C20-H6 and C21-H6. A very weak NOE cross-peak was observed between C20-H1′ and the C21-H6. Sequential $H1'(n) \rightarrow H1'(n+1)$ NOE

Table 1. Refinement results summary

Iter[a]	$\%rms_{vol}$ [b]	*R*-factor	$Q^{1/6}$ factor	$E_{(Total)}$	$E_{(Const)}$	Force[c]/%error[c]
Model	180.4	0.5635	0.1209	-	-	3.0/25
1	148.5	0.5299	0.1090	−1031	7.2	4.0/23
2	125.3	0.5154	0.0994	−1032	8.5	5.0/21
3	104.4	0.4485	0.0874	−1031	21.6	6.0/19
4	96.0	0.4675	0.0921	−1022	36.7	7.0/17
5	84.5	0.4033	0.0740	−1028	44.7	8.0/15
6	70.6	0.3979	0.0755	−1021	78.2	9.0/13
7[d]	73.9	0.4022	0.0738	−1021	82.0	10.0/10
⟨SA⟩[e]	74.2 ± 8.4	0.4072 ± 0.08	0.0742 ± 0.008	-1028 ± 8	82.6 ± 6.4	10.0/10

[a] MORASS iteration number.
[b] The average of the %rms differences between the experimental and theoretical volumes.
[c] Flat-well potential function parameters (force constant in kcal/mol Å^2; Error is the permitted % error on the constraining distances).
[d] Values for the final averaged structure.
[e] ⟨SA⟩ Values for the ten structures with lowest constraint energy.

connectivities were observed for residues C18-A19-C20, and between C21-U22, but not between C20 and C21. In the 250 ms mixing time NOESY, a weak cross-peak was observed between C20-H6 and C21-H5. However, this cross-peak was not seen in the 50 ms NOESY spectra. These observations suggest an increased spatial separation of C20 and C21, possibly due to a wedging effect of the facing A6 residue that is stacked into the helix, resulting in increased distance between C20 and C21.

Several inter-nucleotide NOE cross-peaks observed for the H2 proton of A6 residue indicate that this bulged adenine base is stacked into the helix. The H2 proton of A6 shows NOE cross-peaks to the H1′ sugar protons of A6 and G7. The crucial evidence for helical stacking comes from the inter-strand NOE cross-peak observed between the H2 of the bulged A6 and the H1′ of C21 on the opposite strand. Moreover, the amino proton of A6 showed weak NOESY cross-peaks to the amino protons of residues C20 and C21, on the opposite strand.

Structure calculation

Distance restraints obtained from integrated NOESY volumes using MORASS program were used in the structure calculation. A total of 419 restraints were used. The structure was calculated by an iterative complete relaxation matrix refinement method using MORASS/AMBER. The progress of the iterative refinement was monitored by several key indicators (described in Materials and Methods), which are summarized in Table 1. Both experimental and theoretical %rms (volume) values start relatively high and gradually settle down to lower values with an increased percentage of volume merging between the experimental and theoretical volumes. As expected, the *R*-factor and

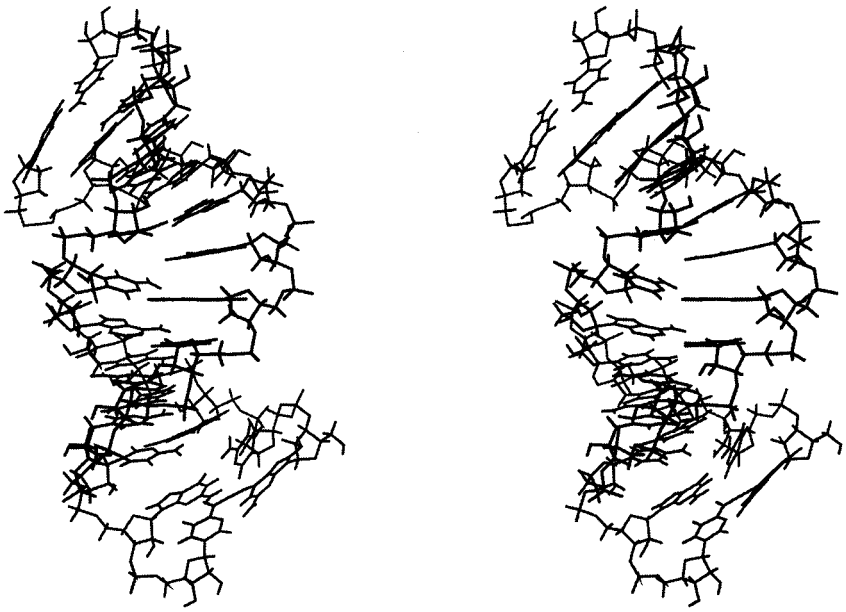

Figure 3. Stereoview of the final averaged structure. The structure was obtained with dihedral restraints on the backbone.

the $Q^{(1/6)}$-factor were also found to decrease as the refinement progressed. The structures clearly improved at each refinement step. The energetics of the molecule, such as the total energy, constraint energy and potential energy, were also monitored during the iterative refinement. Both the total energy and the constraint energy increased in value as the error bars and force constants on the restraints were tightened during the molecular dynamics refinement.

A stereoview of the final averaged structure is shown in Figure 3, and an ensemble of ten structures with lowest constraint energy is shown in Figure 4. The overall structure of this bulged duplex is well defined by NMR. The heavy atom rmsd for the ten structures (shown in Figure 4) is 1.3 Å and the rmsd for the local structure around the bulge, shown in Figure 4(b), (for residues G5-A6-G7, and C20-C21) is 1.1 Å. Structural refinement statistics are shown in Table 2.

To test the quality of the refinement, 2D NOESY volumes were calculated based on the final averaged structure, and were compared to the experimental 2D NOESY volumes. As can be seen in the plot in Figure 5, most of the calculated volumes fit very well with the experimental volumes, indicating the high quality of the final structure and the refinement process. The MORASS program was used to calculate the volumes from the final structure. Significant deviations between calculated and experimental NOESY volumes were seen for residues at the termini, where fraying of the strands and the resulting dynamics are not accounted for in the volume calculation. Agreement of experimental and calculated volumes indicates that not only the distance restraints are satisfied but also that the relaxation matrix approach is satisfied (Wu *et al*., 1997).

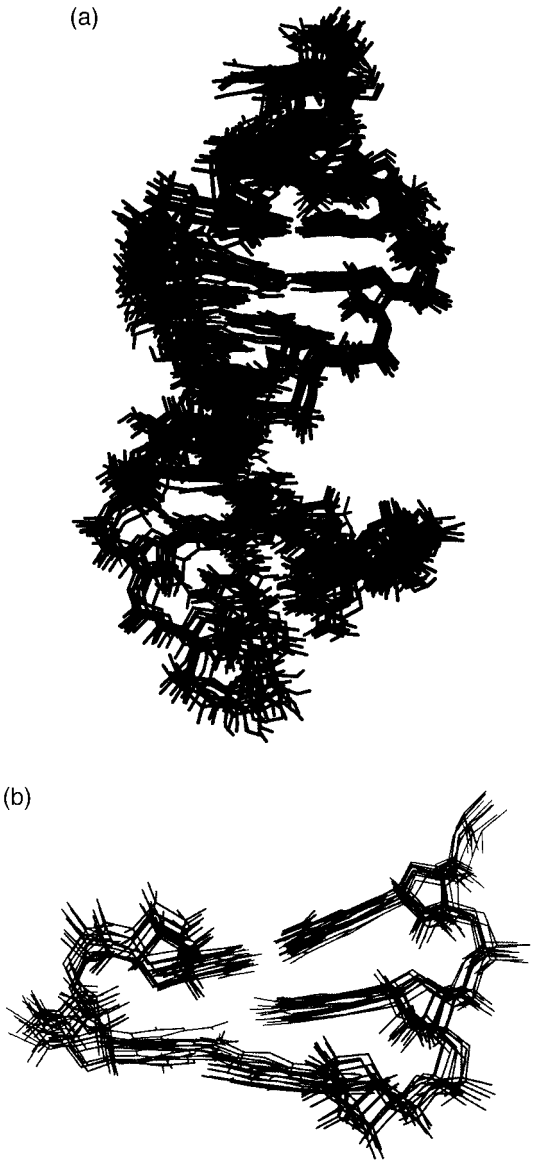

Figure 4. (a) Overlay of ten final structures with lowest constraint energy. (b) Overlay of the bulged region from the ten final structures with lowest constraint energy. The structures were obtained from r-MD with backbone dihedral restraints.

Refinement with torsional angle restraints

To test the effect of restraints on the backbone dihedral angles, we carried out two independent sets of refinements. For the first set of refinements, no restraints were used on the backbone dihedral angles. Sugar puckering restraints on residues G5 and A6 were used as to restraint these two sugars to be in *B* form. Glycosidic angles were not restrained. For the second set, along with the sugar puckering restraints on G5 and A6, torsional restraints were also used on backbone dihedrals and glycosidic angles as described in Materials and Methods.

The rmsd between the two final averaged structures obtained from the two sets of refinements, shown in Figure 6, is 1.2 Å. Analysis of helical parameters (Table 3) and torsional angles for the two final averaged structures revealed that major differences were seen only at the ends of the helix. No significant difference in energies were observed between the two final averaged structures.

Discussion

Preferred conformation of the bulged adenosine

Why do X-ray and NMR analyses disagree over the conformation of adenine bulge duplexes? In all X-ray crystal structures of *A* bulge DNA and RNA duplexes, the adenine base is flipped out of the helix, while in NMR solution structures adenine is stacked into the helix. The fact that results of our studies and those of other NMR studies on bulged adenine bases differ from the results of X-ray crystallographic studies may be attributed to the differ-

Table 2. Structure determination statistics

A. *Distance Restraints*			
Inter-residue	267		
Inter-residue	146		
Inter-strand	6		
B. *Structure statistics*		SA	⟨SA⟩
Distance violations (Å)	>1.0	1	1.2 ± 0.1
	>0.5	4	4.6 ± 0.8
	>0.2	14	16 ± 1
Largest distance violation (Å)		1.09	1.17
Angle violations[a]	>5°	None	None
	>2°	None	None
C. *Mean deviation from ideal covalent geometry*			
Bond angles (deg.)		2.41	2.48 ± 0.4
Bond lengths (Å)		0.011	0.012 ± 0.001
D. *rmsd for all heavy atoms (relative to the mean structure)*			
ten lowest energy			
All residues	1.32 Å		
Bulged region	1.09 Å		

[a] For the r-MD set using torsional restraints on backbone angles.

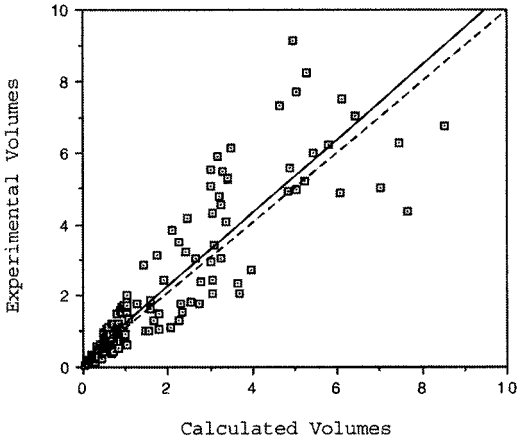

Figure 5. Plot of the experimental volumes *versus* theoretical volumes. Theoretical volumes were calculated from the final averaged structure refined without dihedral restraints. Linear regression is shown as a continuous line and the linear fit through the origin with a unity slope is shown in broken line.

ences in solution and solid-state conditions. The following reasons may be attributed for the difference in the results. First, a looped out structure may be stabilized by inter-molecular contacts in crystal. A 13-mer DNA duplex, d(CGAGAATTCGCG)$_2$, was studied by both X-ray (Joshua-Tor *et al.*, 1988) and NMR (Nikonowicz *et al.*, 1990). In order for the 13-mer to stack end-to-end in the crystal, it must be straight. Wedging in of the bulge base forces the structure to kink and consequently the duplex adopts in solution a bent geometry which is not favored for stacking and crystallization. Although the crystal structure for this 13-mer was originally proposed to have the bulge adenosine in an extrahelical conformation, later analysis of the X-ray structure showed that the bulge adenosine is indeed in an intercalating conformation (Joshua-Tor *et al.*, 1992). However, the duplex receiving the bulge adenosine molecule is from a different molecule in the crystal and the bulge adenosine base connects one duplex to another. Clearly in dilute solution, such intermolecular interactions will not be a factor and the duplex can adopt its low-energy (stacked in) conformation. Secondly, exposing the aromatic base to solvent is energetically costly, and therefore, an unfavorable conformation in solutions. Thirdly, the divalent metal ions, such as Mg^{2+}, may stabilize the looped out adenine in crystals. All reported NMR studies were performed in solutions containing no Mg^{2+} or other divalent ions, whereas most crystals were grown in Mg^{2+}-containing media.

From the NMR data of our RNA bulge, we see no evidence of either a second conformation or rapid equilibration with a significantly populated looped out form. Note that the rmsd of the heavy atom coordinates are in fact lower in the bulge region than for the molecule as a whole. In addition the NOEs in the bulge region are all 100% relative intensity to those in the stem, indicating that a significant population of the extrahelical conformation is not present (we make the assumption that NOEs to a flipped out, conformationally flexible base would be negligible). Thus within the signal-to-noise ratio of the experiment we can estimate that less than 10% of the bulge RNA is in a flipped out conformation.

Bulged adenosine is inserted into the duplex with little perturbation

The ability to trace NOE connectivities between the base aromatic protons and their own and the 5′-flanking sugar H1′, H2′, and H3′ protons along the (G5-A6-G7) segment and the (C20-C21) segment shows that the bulged adenosine A6 is accommodated into the helix with minimal disruption of the duplex. Inter-residue aromatic-aromatic and sugar-aromatic NOESY cross-peaks were observed for the residues flanking the bulged residue. These cross-peaks provide further evidence that the bulged residue and the neighboring residues are stacked upon each other very well. There is no line broadening seen for the resonances of A4, G5, A6, G7 or U8 protons, indicating little or no dynamics around the bulged residues (other than possible very fast dynamic averaging). Analysis of helical parameters further supports the conclusion that the bulged residue is inserted with little or no perturbation. The helical parameters follow that of an *A* form helix. However, deviations from a regular *A* form duplex are observed at the bulged site. The bulged residue A6, and the 5′-flanking residue G5, have C_2'-*endo* sugar pucker, while the rest of the residues (except the ones at the ends of the strands) exhibit C_3'-*endo* sugar puckering. The observed change in the sugar puckering

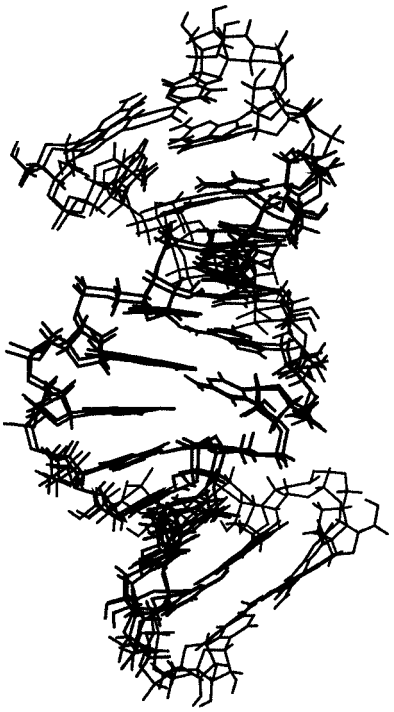

Figure 6. Overlay of the final averaged structures from r-MD with and without backbone dihedral restraints.

around the bulge site result in increased entropy, and thereby may offset the free energy change associated with the insertion of the bulged residue into the duplex (Ennifar *et al*., 1999).

Compression of major groove and the expansion of the minor groove at the bulge site is evident by the smaller roll for the base-pairs flanking the bulged residue compared to the rest of the base-pairs in the duplex (Table 3). Despite the expansion of the minor groove and the compression of the major groove, no severe unwinding of the helix is observed. Helical twist remains $\sim 33^\circ$ throughout the helix, although a slightly higher twist is observed near the bulge. Expansion of the helical backbone near the bulge region is reflected by the short helical rise around the bulge, 2.8($\pm$0.2) compared to 3.5($\pm$0.4) for the rest of the helix. However, the inter-strand P-P distance remains the same throughout the helix.

Observation of imino resonances between base-paired residues flanking the bulged residue shows that the bulged residue is inserted with little or no perturbation and the bulged region is well defined. This is supported by the Cartesian rmsd of 1.1 Å for the bulged region among the ten structures with lowest constraint energies. Although the bulged residue is inserted into the helix with little perturbation, the wedge-like insertion introduces a kink in the helical axis of the duplex that could serve as potential recognition sites for proteins and drugs that bind to specific regions of RNA.

The NMR structure shows a normal *anti* conformation for the stacked in A6, further supporting the conclusion that the A6 is inserted into the duplex with little pertubation.

The stacked adenine base induces a kink of $\sim 28^\circ$ ($\pm 4^\circ$) on the helical axis. Pronounced kinking at bulge sites were observed in gel retardation, chemical probing and NMR experiments. The observed kinking of 28° in this RNA duplex is higher compared to the 15-16° kinking per bulged residue generally observed in RNA and DNA duplexes. Fluorescence resonance energy transfer (FRET) studies on a DNA and RNA duplexes with a seven-residue bulge showed a kinking of 90° (Gohlke *et al*., 1994). Gel retardation studies also revealed a kinking of $\sim 15^\circ$ per bulged residue (Zacharias & Hagerman, 1995).

Effect of dihedral restraints on MD refinement

We have argued against using tight restraints on the backbone dihedral angles during the MD refinement. Previously, we have reported (Gorenstein, 1994) that large amplitude ε and ζ torsional angle fluctuations can occur during restrained-MD for *B*-DNA duplexes. Fluctuations for these two important torsional angles indicate transitions between low energy B_{I} ($\varepsilon = g-$; $\zeta = t$) conformations to higher energy B_{II} ($\varepsilon = t$; $\zeta = g-$) conformations. These results (Gorenstein, 1994) clearly demonstrated that there are significant variations in the relative populations of B_{I} and B_{II} conformations for duplex deoxyoligonucleotides in solution. Therefore, placing no restraints on the backbone dihedral angles will allow a greater degree of conformational flexibility and eliminates artificial claims of higher precision in the refined structures. In contrast to DNA duplexes, RNA duplexes generally exhibit greater variation (and dynamics) in torsional angles α and γ (Saenger, 1984) on the 5′-side of the phosphate group.

We have shown earlier from simulation studies that helical parameters and backbone torsional angles may be accurately reproduced by the hybrid matrix methodology (Kaluarachchi *et al*, 1991). The perfect data simulations suggest that, indeed, inter-proton distance restraints in principle derived from NOESY function to adequately define both the local helical parameters and the backbone torsional conformation. Since no direct distance constraints fix the position of phosphate ester atoms that define α and ζ, reproduction of these two angles may be poorer than the other angles. A greater level of variation is seen between the two final averaged structures for α and γ, while no significant differences are seen between the two structures for the other torsional angles. This suggests that these torsional angles are dynamically averaged in solution. However, alternatively adding more volume restraints to the MORASS calculations could result in better reproduction of nearly all the backbone torsional angles throughout the sequence (Kaluarachchi *et al.*, 1991).

Table 3. Inter-base-pair helical parameters for the final refined structures (A) with dihedral restraints and (B) without dihedral restraints

A. *Helical parameters for the final structure refined using dihedral restraints*						
Base-pair step	SHF	SLD	RIS	TLT	ROL	TWS
G2-C24/G1-C25	0.77	−2.0	3.0	−0.99	9.2	28.99
C3-G23/G2-C24	0.01	−1.8	3.5	7.51	0.8	37.46
A4-U22/C3-G23	1.61	−1.9	3.1	8.12	1.8	32.66
G5-C21/A4-U22	−0.26	−2.1	2.6	3.52	2.3	34.01
U8-A19/G7-C20	0.76	−1.8	2.8	−0.94	1.7	33.93
G9-C18/U8-C20	−0.33	−1.8	3.2	8.07	−1.6	31.96
C10-G17/G9-C18	0.42	−2.2	3.4	3.35	12.1	35.48
C11-G16/C10-G17	0.68	−2.2	3.2	4.38	4.8	30.69
G12-C15/C11-G16	0.94	−1.5	4.1	10.19	22.4	31.92
C13-G14/G12-C15	0.28	−2.2	3.7	−8.87	7.5	35.26
B. *Helical parameters for the final structure refined without dihedral restraints*						
Base-pair step	SHF	SLD	RIS	TLT	ROL	TWS
G2-C24/G1-C25	0.83	−2.6	4.7	1.18	41.2	30.31
C3-G23/G2-C24	−0.25	−2.4	3.4	3.21	−3.8	38.48
A4-U22/C3-G23	1.13	−2.2	3.1	5.73	1.6	29.65
G5-C21/A4-U22	0.02	−2.0	2.7	2.69	2.0	34.61
U8-A19/G7-C20	0.45	−1.8	2.9	3.01	0.2	33.85
G9-C18/U8-C20	0.25	−1.7	3.6	6.31	2.9	33.21
C10-G17/G9-C18	0.21	−2.2	3.2	1.76	11.9	32.58
C11-G16/C10-G17	0.52	−2.1	3.2	1.61	5.5	30.55
G12-C15/C11-G16	−0.06	−1.4	5.1	5.04	30.7	23.73
C13-G14/G12-C15	0.13	−1.23	2.9	−9.49	14.2	26.31

Abbreviations: SHF, shift; SLD, slide; RIS, rise; TLT, tilt; ROL, roll; TWS, twist.

Because of the crankshaft-type motion of the A_I to A_{II} conformational transition, the actual length of the sugar-phosphate backbone is not significantly different in the two conformations. Thus the NOEs will not be greatly perturbed and use of fixed distance constraints (within a flat-well energy function) will still allow sampling of the different A_I/A_{II} conformations. For DNA duplexes, a similar effect is found for the B_I *versus* B_{II} conformation groups. Thus, the mixed conformational dynamics observed for a number of the phosphate groups in the RNA bulge duplex likely more accurately represents the true depiction of the structure. If we were to constrain the backbone into the more normal (A_I for RNA or B_I for DNA) torsional angles, without explicit experimental data supporting these torsional angles, then we could well be misrepresenting the relative dynamics of these molecules. Indeed for DNA duplexes (Kaluarachchi *et al.*, 2000) analysis of coupling constants, NOESY distance restrained structures and restrained molecular dynamics (r-MD), strongly supports a picture of considerable dynamic flexibility along the sugar phosphate backbone.

Relaxation matrix refinement

In larger biomolecules, such as this RNA duplex, the two-spin approximation leads to systematic errors in estimating inter-proton distances, since indirect magnetization transfer *via* other protons (spin-diffusion) contributes significantly to NOE intensities even at short mixing times. The two-spin approximation requires that the NOESY derived distances be obtained from vanishingly short experimental mixing times where the build-up of NOESY intensity is proportional to the inverse sixth power of the inter-proton distance and the effects of spin diffusion are minimal. Since most of the structurally important longer range NOEs are not observed at these short mixing times, the use of the two-spin approximation has raised concern over the accuracy of the refined NMR structures derived by this methodology. Unfortunately, in many RNA NMR structures refined from $^{15}N/^{13}C$-edited NOESY spectra, the two-spin analysis is widely utilized. The MORASS complete relaxation matrix procedure utilizes an incremental replacement of the theoretical volumes with the experimental values as the iterative refinement proceeds. This allows a more robust Eigenvalue/Eigenvector solution to the Bloch relaxation equations, and more precise (and hopefully accurate) depiction of the true solution structure.

Materials and Methods

The NTPs were obtained from Sigma and the T7 RNA polymerase enzyme was prepared from the over-producing *E. coli* strain BL-21 containing plasmid pAR219 obtained from Brookhaven National laboratory, NY (Davanloo *et al.*, 1984). Other reagents used in the transcription reaction were purchased from Pharmacia. RNAses along with the RNA sequencing kit were obtained from Pharmacia LKB laboratories. Radioactive [γ-^{32}P]ATP was obtained from Dupont. DNA templates were obtained in purified form from Midland Certified Reagents, TX.

RNA synthesis and purification

The RNA strands, r(GGCAGAGUGCCGC) and r(GCGGCACCUGCC) were synthesized separately by *in vitro* run-off transcription of the appropriate DNA templates using T7 RNA Polymerase (Millgan *et al*., 1987). The transcription reaction contained 40 mM Tris-HCl (pH 8.1), 20 mM $MgCl_2$, 5 mM DTT, 1 mM spermidine, 4 mM each of ATP, CTP, GTP and UTP, 10 mM GMP, 0.01% (w/w) Triton X-100, and 112 nM template and 50 μg/ml T7 RNA polymerase incubated at 37°C for overnight. The RNA products were precipitated using cold ethanol and the strands were purified by polyacrylamide gel electrophoresis under denaturing conditions (8 M urea). The RNA strands were extracted from the gel by electroelution, and concentrated by ultrafiltration using Centricon-3 centrifugal concentrator devices (Amicon). The sequences of the RNA strands were verified by established enzymatic sequencing methods using RNAs U2, BC and T1 (for details, see Thiviyanathan, 1992).

Sample preparation

The RNA strands, dissolved in buffer containing 100 mM NaCl, 10 mM sodium phosphate (pH 6.8) and 0.02 mM EDTA, were mixed in 1:1 stochiometric ratio calculated from their molar extinction coefficients (Puglisi & Tinoco, 1989) and were annealed by heating the mixture to 80°C and slowly cooling back to room temperature. Single-strand contaminants were removed by running the annealed sample down a Sephadex-100 size-exclusion column. Purity of the sample was tested again by running an analytical gel. Samples used for non-exchangeable proton NMR experiments were lyophilized several times from ^2H_2O before dissolution in 99.96% ^2H_2O (obtained from Cambridge Isotope Laboratories).

NMR experiments

Two dimensional NOESY and TOCSY spectra at room temperature in ^2H_2O were collected at 500 MHz on a Varian VXR-500 instrument. Phase-sensitive two dimensional NOESY spectra were collected with quadrature detection using a spectral width of 5000 Hz and mixing times of 50, 250 and 400 ms. The real and imaginary data points were acquired sequentially. Attenuation of the residual HOD resonance was achieved by low power pre-saturation pulse during the relaxation delay. The data sets were collected with 512 t_1 experiments using 2048 complex points in the t_2 dimension and 32 scans per t_1 increment. The relaxation delay was three seconds.

Two-dimensional TOCSY data were collected with a spectral width of 4400 Hz and 100 ms mixing time. The TOCSY data set were collected with 2048 complex points in the t_2 dimension and 256 t_1 increments. Phosphorus spectrum of the RNA duplex in ^2H_2O at 25°C was obtained at 81 MHz on a Varian XL-200 instrument. The $^1H/^{13}P$ hetero TOCSY experiment (Kellog, 1992) was performed at 240 MHz on a Varian UnityPlus instrument.

Data processing and analysis

The 2D data sets were processed using Felix program (version 95; Biosym Inc., San Diago, CA) to give a data matrix of 2048 × 1024 real data points after zero-filling in both dimensions. Only the real part of the matrix was stored. A 90° shifted sine bell apodization function was used in both dimensions. Baseline correction was done using the FLATT routine provided in the Felix program (Guntert & Wüthrich, 1992). Peaks were picked using the automated peak-picking routine available in the Felix program. The linewidth was set at 10 Hz for the automated peak-picking and the box sizes were manually adjusted for each peak of interest.

Distance and torsional angle restraints

Distance restraints were calculated from the integrated NOESY cross-peaks between non-exchangeable protons, using the iterative hybrid relaxation matrix program, MORASS (Post *et al*., 1990; Meadows *et al*., 1997). The intra-residue H5-H6 volume of pyrimidine bases were used as the ruler to derive distances from NOESY volumes. The cross-peaks between exchangeable protons were qualitatively classified as weak (1.9-5.5 Å), medium (1.9-4.5 Å), and strong (1.9-3.5 Å). A total of 419 distance restraints derived from integrated NOESY cross-peaks were used for the structural refinement. To ensure Watson-Crick base-pairing in the helix, as shown by the assigned imino peaks, hydrogen bond restraints were added between the base-pairs in the duplex. A total of 12 hydrogen bond restraints were used at an equilibrium distance of 1.9 (±0.19 Å) with a force constant on each restraint of 15 kcal/mol $Å^2$. Only one hydrogen bond restraint was applied to either A · U (between N1 of A and H3 of U) or G · C base-pairs (between H1 of G and N3 of C) to allow variable propeller twist between the base-pairs during the refinement.

C'_2-*endo* sugar puckering restraints were introduced for the sugars of G5 and A6 residues, based upon the observed H1′-H2′ scalar cross-peaks. A C'_3-*endo* conformation was assigned to all the ribose sugars of residues except G1, G5, A6, C13, G14, and C25, based on the absence of H1′-H2′ scalar cross-peaks indicating a weak $^3J_{H1'-H2'}$. The sugars of the terminal residues (G1, C13, G14, and C25) were not restrained. Two independent sets of MD refinements were carried out to compare the effect of using restraints on backbone dihedral angles during MD calculations. In the first set, no backbone dihedral restraints were used. Sugar puckering restraints were used on sugars of A6 and G5 to restraint them to be in C'_2-*endo* conformation. Glycosidic angles were not restrained. For the second set, dihedral restraints were used as described below. Since all of the $^3J_{P-H5'}$ and $^3J_{P-H5''}$ coupling constants were small (less than 5 Hz) β angles were restrained to be in the *trans* conformation (Wijmenga *et al*., 1993). Additionally, loose dihedral constraints (±30°) were used for α and ζ, derived from standard values for an *A* form helix (Saenger, 1984) for all residues except residues G5, A6 and the residues at the termini, based upon the narrow range of ^{31}P chemical shifts found for all phosphate groups (Gorenstein, 1994) and the absence of H1′-H2′ cross-peaks in the TOCSY spectra. Glycosidic torsional angle χ was restrained to be *anti* (180(±40)°) for all residues, based upon the lack of strong intra nucleotide H1′-H8/H6 NOE cross-peaks. The ε angles were derived from H3′(n)-P(n + 1) couplings, and are restrained to be *trans* for all residues (210(±40)°). The γ angles were restrained to be in *gauche*$^+$ (60(±40)°) based on small coupling constants between H4′/H5′ and H4′/H5″. The δ angles were restrained, however, C'_3-*endo* sugar puckering restraints were used for all residues except, G5 and A6 which were restrained to be C'_2-*endo*, and the terminal residues for

which no restraints were used. The force constant was set to 25 kcal/mol deg^2 on each dihedral restraint.

Structure refinement

Structure refinement was done using the iterative MORASS/AMBER procedure as described (Kaluarachchi *et al*., 1991). Initial model structure was built using the xLeap visual interface (Schafmeister *et al*., 1995) provided in the AMBER suite. In the starting structure, the bulged adenosine residue was inserted into the helix, as several convincing experimental NOEs indicated that the bulged residue is inserted into the helix. The initial structure was energy minimized for 3000 steps without any restraints. To neutralize the negative charges on phosphate groups 23 Na^+ counter ions were placed around phosphate groups. A rectangular box was added, providing at least 10 Å of explicit TIP3P water molecules yielding 4692 water molecules. Molecular dynamics (MD) simulations were carried out using the SANDER module of AMBER 5.0 (Case *et al*., 1997) with SHAKE applied to all hydrogen atoms and 2 fs time steps. A 10 Å cutoff was applied to the Lennard-Jones interactions. Constant pressure was maintained with isotropic scaling. At first, the water box was subjected to a series of equilibration MD runs while holding the solute fixed (Cheatham *et al,* 1995). During the equilibration steps, when position constraints on solute molecules were gradually relaxed, and the final refinement runs particle mesh Ewald method (Darden *et al*., 1993; Konerding *et al*., 1999) was used to calculate the electrostatic interactions. The resulted structure was used as the starting structure for the subsequent rounds of refinement. The errors (i.e. uncertainties) in the calculated inter-proton distances were described using a flat-well energy penalty function (Gorenstein *et al*., 1990) in which the width of the flat-well is defined as a percentage of the equilibrium inter-proton distance, r_{ij}^{eq}. The flat-well span was decreased from an initial value of ±25 % r_{ij}^{eq} to ca 10 % while the distance restraint force constants were increased accordingly as the structures approached convergence. For each iteration, theoretical volumes were calculated from a model structure based on isotropic tumbling with a correlation time of 3.0 ns. This correlation time was determined by optimizing the match between the theoretical NOE volumes and the experimentally determined NOE volumes. Previous calculations by others (Mujeeb *et al*., 1993) indicate that linear DNA molecules with less than 13 base-pairs behave isotropically. Theoretical volumes were merged with the experimental volumes to form a hybrid matrix, which is diagonalized to yield the complete relaxation rate matrix from which the inter-proton distances are calculated. These distances were used in restrained molecular dynamics (r-MD) refinement using AMBER (version 5.0) (Case *et al*., 1997). The output structure from the r-MD is then used as the model structure in the next iteration. The entire procedure, including data simulation, is repeated until internal consistency among the calculated and observed 2D NOESY spectra and the output structure is reached.

For the r-MD procedure, the starting structure was first energy minimized with the NOE constraints for 3000 steps. A 12 ps of restrained molecular dynamics protocol with temperature annealing was performed on the energy minimized model (temperature is increased from 298 K to 600 K for 4 ps; cooled to 298 K over the next 4 ps; continued at 298 K for the last 4 ps). The averaged structure from the last 4 ps of the r-MD was energy minimized for 3000 steps, and the resulting structure was used as the starting structure for the next iteration.

During r-MD, the non-bonded interaction cutoff distance was set to 8 Å and a distance dependent dielectric constant was used with an integration time step of 1 fs. The charges at the 5′ and 3′-ends of the RNA strands were modified to avoid non-physical electrostatic interactions.

The progress of the iterative refinement process was monitored by several key factors, indicating the match between experimental NOESY volumes and theoretical NOESY volumes calculated from the refined structure. The rms errors in the volumes were used as the first criterion for monitoring the refinements. The %rms (volume) is given by:

$$\%\text{rms (volume)} = \sqrt{1/N \sum_{ij} \left(\frac{v_{ij}^a - v_{ij}^b}{v_{ij}^a} \right)^2} \times 100\%$$

where *a* or *b* can be either the experimental or theoretical 2D volumes to give the %rms(exp) or %rms(the), respectively.

The *R*-factor (Gonzalez, *et al*., 1991) that is similar to the *R*-factor used in X-ray crystallography was also used as a refinement criteria. The *R*-factor is given by:

$$R = \frac{\sum_{ij} |v_{ij}^a - v_{ij}^b|}{\sum_{ij} v_{ij}^a}$$

We have suggested that the %rms (volume) is a very useful measure of quality of fit to the spectra, since it weighs the percentage differences in the theoretical and experimental volumes for both large and small cross-peaks equally. The *R*-factor is regarded as a poorer measure of the quality of the refined structure since it is often dominated by the largest cross-peaks. Another quality of fit, the $Q^{(1/6)}$-factor (Thomas *et al*., 1991), also better reflects the quality of the structure since it weighs more heavily on the weak cross-peaks (longer, inter-residue distances) compared to the *R*-factor. The $Q^{(1/6)}$ factor is defined as:

$$Q^{(1/6)} = \frac{\sum_{ij} \tau_m |(v_{ij}^a)^{1/6} - (v_{ij}^b)^{1/6}|}{\sum_{ij} (1/2)\tau_m |(v_{ij}^a)^{1/6} + (v_{ij}^b)^{1/6}|}$$

The dihedral angles and the helical parameters of the final averaged structures were calculated using CURVES5.3 program (Lavery & Sklenar, 1988). All structures were built and displayed with MIDAS2.1 (Ferrin *et al*., 1988).

Acknowledgments

This work was supported by NIH (AI27744), NIEHS (ES06676), and the Sealy and Smith Foundation grants to D.G.G., and NSF CHE-9302619, NIH grant R15-GM55898 and ACS-PRF (31427-B4) to N.B.L. Building funds for the UTMB NMR center were provided by NIH (1CO6CA59098). The authors thank Dr Susan Cady, Dr Igor Ouporov, and Michael Hills for their assistance with sample preparation, data collection and processing.

References

Borer, P. N., Lin, Y., Wang, S., Roggenbuck, M. W., Gott, J. M., Uhlenbeck, O. C. & Pelczer, I. (1995). Proton NMR and structural features of a 24-nucleotide RNA hairpin. *Biochemistry,* **34**, 6488-6503.

Case, D. A., Pearlman, J. W., Caldwell, T. E., Chetham, T. E., III, Ross, W. S., Simmerling, C. L., Darden, T., Merz, K. M., Stanton, A. L., Cheng, J. J., Vincent, M., Crowley, M., Ferguson, D. M., Radmer, R. J., *et al*, (1997). *AMBER 5. 0*, University of California, San Francisco.

Cate, J. R., Gooding, A. R., Podell, E., Thou, K., Golden, B. L., Kundrot, C. E., Cech, T. & Doudna, J. A. (1996). Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science,* **273**, 1678-1686.

Christiansen, J., Douthwaite, S. R., Christiansen, A. & Garrett, R. A. (1985). Does unpaired adenosine-66 from helix II of *E. coli* 5 S RNA bind to protein L18? *EMBO J.* **4**, 1019-1024.

Cheatham, T. E., Miller, J. L., Fox, T., Darden, T. A. & Kollman, P. A. (1995). Molecular dynamics simulations on solvated biomolecular systems. *J. Am. Chem. Soc.* **117**, 4193-4194.

Davanloo, P., Rosenberg, A. H., Dunn, J. J. & Studier, F. W. (1984). Cloning and expression of the gene for bacteriophage T7 RNA polymerase. *Proc. Natl Acad. Sci. USA,* **81**, 5765-5774.

Darden, T. A., York, D. & Pedersen, L. G. (1993). Particle mesh Ewald - an n.log(n) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089-10092.

Ennifar, E., Yusupov, M., Walter, P., Marquet, R., Ehresmann, B., Ehresmann, C. & Dumas, P. (1999). The crystal structure of the dimerization initiation site of genomic HIV-1 RNA reveals an extended duplex with two adenine bulges. *Structure Fold. Des.* **7**, 1439-1449.

Ferrin, T. E., Huang, C. C., Jarvis, L. E. & Langridge, R. (1988). The MIDAS display system. *J. Mol. Graph.* **6**, 13-37.

Fourmy, D., Recht, M. I., Blanchard, S. C. & Puglisi, J. D. (1996). Structure of the A site of *Escherichia coli* 16 S ribosomal RNA complexed with an aminoglycoside antibiotic. *Science,* **274**, 1367-1371.

Gohlke, C., Murchie, A. I. H., Lilley, D. M. J. & Clegg, R. M. (1994). Kinking of DNA and RNA helices by bulged nucleotides observed by fluorescence energy transfer. *Proc. Natl Acad. Sci. USA,* **91**, 11660-11664.

Gorenstein, D. G., Meadows, R. P., Metz, J. T., Nikonowicz, E. & Post, C. B. (1990). 31P and 1H 2-dimensional NMR and NOESY-distance restrained molecular dynamics methodology for defining sequence-specific variations in duplex oligonucleotides; A comparison of NOESY two-spin and relxation rate matrix analysis. In *Advances in Biophysical Chemistry* (Bush, C. A., ed.), pp. 47-124, JAI Press, Greenwich, CT.

Gorenstein, D. G. (1994). Conformation and dynamics of DNA and protein-DNA complexes. *Chem. Rev.* **94**, 1315-1338.

Gonzalez, C., Rullmann, J. A. C., Bonvin, A. M. J. J., Boelens, R. & Kaptein, R. (1991). Toward an NMR R factor. *J. Magnet. Res.* **91**, 659-664.

Guntert, P. & Wüthrich, K. (1992). FLATT - a new procedure for high-quality baseline correction of multidimensional NMR spectra. *J. Magnet. Res.* **96**, 403-407.

Joshua-Tor, L., Rabinovich, D., Hope, H., Frolow, F., Appela, E. & Sussman, J. L. (1988). The three-dimensional structure of a DNA duplex containing looped-out bases. *Nature,* **334**, 82-84.

Joshua-Tor, L., Frolow, F., Appela, E., Hope, H., Rabinovich, D. & Sussman, J. L. (1992). Three-dimensional structures of bulge-containing DNA fragments. *J. Mol. Biol.* **225**, 397-431.

Kalnik, M. W., Norman, D. G., Swann, P. F. & Patel, D. J. (1989). Conformation of adenosine bulge-containing deoxytridecanucleotide duplexes in solution. *J. Biol. Chem.* **264**, 3702-3711.

Kaluarachchi, K., Meadows, R. & Gorenstein, D. G. (1991). How accurately can oligonucleotide structures be determined from the hybrid relaxation rate matrix/NOESY restrained molecular dynamics approach? *Biochemistry,* **30**, 8785-8797.

Kaluarachchi, K., Gorenstein, D. G. & Luxon, B. A. (2000). How do proteins recognize DNA? Solution structure and local conformational dynamics of lac operators by 2D NMR. *J. Biomol. Struct. Dynam.* **11**, 123-133.

Kellog, G. W. (1992). Proton-detected hetero TOCSY experiments with application to nucleic acids. *J. Magnet. Res.* **98**, 176-182.

Konerding, D. E., Cheatham, T. E., Kollman, P. A. & James, T. J. (1999). Restrained molecular dynamics of solvated duplex DNA using particle mesh Ewald method. *J. Biomol. NMR,* **13**, 119-131.

Lavery, R. & Sklenar, H. (1988). The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dynam.* **6**, 63-91.

Leontis, N. B. & Moore, P. B. (1986). NMR evidence for dynamic secondary structure in helices II and III of the 5 S RNA of *E. coli*. *Biochemistry,* **25**, 5736-5744.

Meadows, R., Post, C. B., Luxon, B. A. & Gorenstein, D. G. (1997). *MORASS Program*, University of Texas Medical Branch, Galveston, TX, USA.

Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. (1987). Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucl. Acids Res.* **15**, 8783-8798.

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

Nikonowicz, E., Roongta, V., Jones, C. R. & Gorenstein, D. G. (1989). Two-dimensional H-1 and P-31 NMR spectra and restrained molecular dynamics structure of an extrahelical adenosine tridecamer oligodeoxyribonucleotide duplex. *Biochemistry,* **28**, 8714-8725.

Nikonowicz, E. P., Meadows, R. P. & Gorenstein, D. G. (1990). NMR structural refinement of an extrahelical adenosine tridecamer d(CGCAGAATTCGCG)2 *via* a hybrid relaxation matrix procedure. *Biochemistry,* **29**, 4193-4204.

Perrotta, A. T., Nikiforova, O. & Been, M. D. (1999). A conserved bulged adenosine in a peripheral duplex of the antigenomic HDV self-cleaving RNA reduces kinetic trapping of inactive conformations. *Nucl. Acids Res.* **27**, 795-802.

Portmann, S., Grimm, S., Workman, C., Usman, N. & Egli, M. (1996). Crystal structures of an *A* form duplex with single-adenosine bulges and a confor-

mational basis for site-specific RNA self-cleavage. *Chem. Biol.* **3**, 173-184.
Post, C. B., Meadows, R. P. & Gorenstein, D. G. (1990). On the evaluation of interproton distances for three-dimensional structural analysis by NMR using a relaxation rate matrix analysis. *J. Am. Chem. Soc.* **112**, 6796-6803.
Puglisi, J. D. & Tinoco, I., Jr (1989). Absorbance melting curves of RNA. *Methods Enzymol.* **180**, 304-325.
Rosen, M. A., Live, D. & Patel, D. J. (1992). Comparative NMR study of A-bulge loops in DNA duplexes: intrahelical stacking of *A*, *A-A*, and *A-A-A* bulge loops. *Biochemistry,* **31**, 4004-4014.
Saenger, W. (1984). *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.
Schafmeister, C. E. A. F., Ross, W. S. & Romanovxki, V. (1995). *LeaP Program*, University of California, San Francisco.
Thiviyanathan, V. (1992). Synthesis and physical characterization of the RNA model systems. MSc thesis, Bowling Green State University, Bowling Green, OH.
Thomas, P. D., Basus, V. J. & James, T. L. (1991). Protein solution structure determination using distances from tow dimensional nuclear Overhauser effect experiments. *Proc. Natl Acad. Sci. USA,* **88**, 1237-1241.
Valegard, K., Murray, J. B., Stockley, P. G., Stonehouse, N. J. & Liljas, L. (1994). Crystal structure of an RNA bacteriophage coat protein-operator complex. *Nature,* **371**, 623-626.
Varani, G. & Tinoco, I., Jr (1991). RNA structure and NMR spectroscopy. *Quart. Rev. Biophys.* **24**, 479-532.
Werner, M. H., Clore, G. M., Fisher, C. L., Fisher, R. J., Trina, L., Shiloach, J. & Gorenborn, A. M. (1995). The solution structure of the human ETS-1-DNA complex reveals a novel mode of binding and true side chain interaction. *Cell,* **83**, 761-771.
Wijmenga, S. S., Mooren, M. M. W. & Hilbers, C. W. (1993). NMR of nucleic acids; from spectrum to structure. In *NMR of Macromolecules*, Oxford University Press, Oxford, UK.
Wimberly, B., Varani, G. & Tinoco, I., Jr (1993). The conformation of Loop E of eukaryotic 5 S ribosomal RNA. *Biochemistry,* **32**, 1078-1087.
Wu, M., SantaLucia, J., Jr & Turner, D. H. (1997). Solution structure of $(rGGCAGGCC)_2$ by two dimensional NMR and iterative relaxation matrix approach. *Biochemistry,* **36**, 4449-4460.
Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Zacharias, M. & Hagerman, P. J. (1995). Bulge-induced bends in RNA: quantification by transient electric birefringence. *J. Mol. Biol.* **247**, 486-500.
Zacharias, M. & Sklenar, H. (1999). Conformational analysis of single-base bulges in *A* form DNA and RNA using a hierarchical approach and energetic evaluation with a continuum solvent model. *J. Mol. Biol.* **289**, 261-275.

Edited by I. Tinoco

(Received 30 March 2000; received in revised form 2 June 2000; accepted 6 June 2000)

http://www.academicpress.com/jmb

Supplementary Material is available from JMB Online comprising one Table