An examination of coaxial stacking of helical stems in a pseudoknot motif: The gene 32 messenger RNA pseudoknot of bacteriophage T2

JASON A. HOLLAND, MARK R. HANSEN,¹ ZHIHUA DU, and DAVID W. HOFFMAN

Department of Chemistry and Biochemistry, Institute for Cell and Molecular Biology,
University of Texas at Austin, Austin, Texas 78712, USA

ABSTRACT

The RNA pseudoknot located at the 5' end of the gene 32 messenger RNA of bacteriophage T2 contains two A-form helical stems connected by two loops, in an H-type pseudoknot topology. A combination of multidimensional NMR methods and isotope labeling were used to investigate the pseudoknot structure, resulting in a more detailed structural model than provided by earlier homonuclear NMR studies. Of particular significance, the interface between the stacked helical stems within the pseudoknot motif is described in detail. The two stems are stacked in a coaxial manner, with an approximately 18° rotation of stem1 relative to stem2 about an axis that is parallel to the helical axis. This rotation serves to relieve what would otherwise be a relatively close phosphate–phosphate contact at the junction of the two stems, while preserving the stabilizing effects of base stacking. The ability of the NMR data to determine pseudoknot bending was critically assessed. The data were found to be a modestly precise indicator of pseudoknot bending, with the angle between the helical axes of stem1 and stem2 being in the range of $15 \pm 15^\circ$. Pseudoknot models with bend angles within this range are equally consistent with the data, since they differ by only small amounts in the relatively short-range interproton distances from which the structure was derived. The gene 32 messenger RNA pseudoknot was compared with other RNA structures with coaxial or near-coaxial stacked helical stems.

Keywords: NMR spectroscopy; RNA pseudoknot; RNA structure

INTRODUCTION

A pseudoknot is an RNA structural motif that is formed when a sequence of nucleotides within a single-stranded loop region base pair with a complementary sequence outside that loop (for reviews, see ten Dam et al., 1992; Pleij, 1994). Most of the pseudoknots identified today are of the so-called H-type (or Hairpin-type) in which nucleotides from a hairpin loop base pair with single-stranded regions outside of the hairpin. This topology represents the simplest way to form a pseudoknot structure, featuring two stem regions (stem1 and stem2) and two connecting loops (loop1 and loop2). The two helical stems have the potential to stack coaxially to form a single quasicontinuous helical structure (Pleij et al., 1985; Puglisi et al., 1990; Du et al., 1996; Du &

Hoffman, 1997). Since the early studies of RNA pseudoknots within the tRNA-like structure at the 3' end of a turnip yellow mosaic virus (TYMV) RNA (Rietveld et al., 1983; Kolk et al., 1998), a large body of evidence has accumulated indicating that pseudoknots are present within many RNA classes, including messenger RNAs, ribosomal RNAs, catalytic and self-splicing RNAs, RNA aptamers, and viral RNAs. RNA pseudoknots can play a variety of important functional roles. In three translationally regulated mRNAs, namely *Escherichia coli* S15 mRNA (Portier et al., 1990; Benard et al., 1994; Philippe et al., 1995), *E. coli* α -mRNA (Deckman & Draper, 1987; Tang & Draper, 1989) and T-even bacteriophage gene 32 mRNA (McPheeters et al., 1988), pseudoknots are believed to mediate the binding of their encoded proteins, thereby autoregulating protein expression. Pseudoknot structures are also present at functionally important positions within the coding regions of mRNAs, exemplified by ribosomal frameshifting (e.g., ten Dam et al., 1990; Chen et al., 1995; Du et al., 1997) or translational read-through (Wills et al., 1991, 1994), essential mechanisms employed by retroviruses in regulation the expression of viral proteins.

Reprint requests to: David W. Hoffman, Department of Chemistry and Biochemistry, Institute for Cell and Molecular Biology, University of Texas at Austin, Austin, Texas 78712, USA; e-mail: dhoffman@mail.utexas.edu.

¹Present address: Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309-0215, USA.

The structure of the bacteriophage T2 gene 32 mRNA pseudoknot has previously been investigated using homonuclear NMR methods, and a low resolution structural model was presented that is consistent with the NMR data available at the time (Du et al., 1996). In addition, a mutational study was used to assess the sequence, stem-length, and loop-length requirements for forming a similar pseudoknot structure (Du & Hoffman, 1997). A substantial number of RNA pseudoknots described in the literature have nucleotide sequences that suggest a structural similarity to the gene 32 mRNA pseudoknot, and it has been proposed that these RNAs are members of a pseudoknot family related by a common structural motif (Du et al., 1996; Du & Hoffman, 1997). Features of this pseudoknot family are a stem2 with either 6 or 7 bp, a loop1 consisting of a single nucleotide, and a variable length stem1 and loop2. One member of this pseudoknot family, located downstream of the *gag-pro* frameshift site in simian retrovirus type-1 (ten Dam et al., 1994, 1995), has recently been investigated by NMR methods with results supporting its structural similarity to the gene 32 mRNA pseudoknot (Du et al., 1997).

The goals of the present work are threefold. First, a structural model of the RNA pseudoknot will be presented, based on the results of multidimensional NMR experiments performed on samples enriched in ^{13}C , ^{15}N , or selectively deuterated. Second, the precision of the NMR-derived structural model will be critically assessed, by describing the complete range of all structures that are consistent with the NMR data, while maintaining reasonable molecular geometry. Third, the improved structural model will be compared with other RNA structures that have previously been shown to contain coaxial or near-coaxial stacked stems. The present study is intended to provide further insight into the specific features that stabilize the pseudoknot, complementing recent thermodynamic investigations of structure-stability relationships in RNA pseudoknots (Qiu et al., 1996; Gluick et al., 1997; Theimer et al., 1998).

RESULTS

NMR spectroscopy of the RNA pseudoknot

A necessary step in the analysis of any macromolecular structure by NMR methods is the assignment of each observed resonance to the nucleus from which it arises. In the case of the bacteriophage T2 gene 32 mRNA pseudoknot (Fig. 1), many of the proton resonances were previously assigned using standard two-dimensional proton NMR methods (Du et al., 1996). In the present work, these assignments were confirmed and significantly extended using a combination of multidimensional NMR methods and isotopically labeled RNA samples. Resonance assignments for nonexchangeable base and ribose H1', H2', H3', and H4' protons

FIGURE 1. Schematic diagram of the gene 32 messenger RNA pseudoknot from bacteriophage T2. The diagram shows the numbering of the nucleotides referred to in the text, and is drawn in such a way as to show the base pairing that occurs within the pseudoknot.

were 91% complete, with most of the missing assignments being in the nucleotides of the regular A-helical stems, where the resonance overlap was the greatest.

NMR spectroscopy of ^{13}C - and ^{15}N -labeled RNA pseudoknot

A series of NMR experiments that exploit ^{13}C and ^{15}N chemical shifts and relatively large ^1H - ^{13}C , ^1H - ^{15}N , and ^{13}C - ^{13}C coupling constants were performed on a sample of the RNA pseudoknot enriched in ^{13}C and ^{15}N . These experiments have recently become relatively common in the analysis of RNA structure (cf. Allain & Varani, 1995; Pardi, 1995; Jiang et al., 1996; Varani et al., 1996; Dieckmann & Feigon, 1997; Wijmenga & van Buuren, 1998). The resonance assignments previously obtained using homonuclear two-dimensional methods (Du et al., 1996) served as a starting point for the present work. An HCCH-COSY spectrum (Fesik et al., 1990; Kay et al., 1990) was then used to confirm and extend the ribose ^1H chemical shift assignments, and determine the ^{13}C chemical shifts of the ribose resonances (Fig. 2). Characteristic heteronuclear chemical shifts were used to confirm the nucleotide type associated with proton resonances (Varani & Tinoco, 1991). The introduction of the ^{13}C isotope also permitted the ^1H - ^1H NOE spectrum to be resolved into a third dimension using the ^{13}C chemical shift, significantly reducing resonance overlap (Fig. 3), useful in complet-

FIGURE 2. A single plane of the HCCH-COSY spectrum of ^{13}C -enriched gene 32 mRNA pseudoknot. This spectrum permits scalar-coupled resonances within the same ribose ring to be identified and resolved in three dimensions. The plane shown corresponds to a ^{13}C chemical shift of 74.4 ppm (C2' and C3' carbons).

ing the resonance assignments as well as in identifying additional structural constraints. ^1H - ^{15}N correlated spectra provided unambiguous identification of the ribose and base protons that are part of the same nucleotide (Sklénar et al., 1994), based on their mutual correlation to the N9 nucleus (in purines) or N1 nucleus (in pyrimidines).

Resonance assignments were further confirmed using samples that were selectively labeled with deuterium. Samples were prepared that were selectively deuterium labeled with either adenosine or guanosine, for the purpose of distinguishing the H8 resonances of these two nucleotide types. This is a general problem in NMR studies of RNA, because the resonances to which correlations from H8 are most easily observed (C8, N7, and N9) do not occur in unique regions of the ^{13}C or ^{15}N spectrum. Two-dimensional NOE spectra of the selectively deuterated pseudoknot samples were compared with the spectrum of the fully protonated sample to confirm the purine H8 resonance assignments (Fig. 4). Although rather labor intensive, the experiments were worthwhile: in three instances (involving H8 resonances of loop nucleotides G1, A22, and G27), errors in the previously reported assignments were detected (Du et al., 1996).

Assignments of the ^{15}N nuclei that are not directly bonded to protons were made using a set of ^1H - ^{15}N correlated spectra. We note that there has been some confusion in the literature regarding the assignments of

adenosine N1 and N3. In relatively early work, N1 was reported as being downfield of N3 (e.g., Markowski et al., 1977; Büchner et al., 1978; James et al., 1981). More recent work has reported N1 as being upfield of N3 (e.g., Michnika et al., 1993; Sklénar et al., 1994). In the present work, the downfield resonance is assigned as N1. To support this assignment, a series of ^{15}N - ^1H correlated spectra of adenosine triphosphate (ATP) were obtained over a range of pH values. It was found that the chemical shift of the ATP resonance assigned to N1 is quite sensitive to pH, varying from 157 ppm at pH 3.2 to 226 ppm at pH 6.5, while the N3, N7, and N9 resonances only exhibit relatively small chemical shift changes of 5–8 ppm within this same pH range. The large chemical shift change of the adenosine N1 resonance with varying pH is consistent with its protonation; the pK_a of N1 is approximately 3.5 (Saenger, 1984). Note that the chemical shifts N1 and N3 of A35 in the pseudoknot, which is located in a single-stranded region, closely match the chemical shifts of N1 and N3 in the ATP mononucleotide, when each sample is at pH 6.5 (Fig. 4).

Chemical shifts of ^{15}N have previously been proposed to be influenced by hydrogen bonding in DNA (Gao & Jones, 1987; Wang et al., 1991) and RNA oligomers (Gaffney et al., 1993; Aboul-ela et al., 1995). The results of the present experiments are consistent with this hypothesis, although the effects appear to be somewhat modest. The N3 nuclei of the uridines that are

FIGURE 3. A single plane of the 3-dimensional HMQC-NOESY spectrum of ^{13}C -enriched gene 32 mRNA pseudoknot, which permits ^1H - ^1H NOE peaks to be resolved by the chemical shift of the ^{13}C nucleus to which one of the protons is bonded. The spectrum is folded once in the ^{13}C dimension, so that each plane corresponds to two ^{13}C chemical shifts. The plane shown contains ^{13}C chemical shifts of 74.1 ppm ($\text{C}2'$ and $\text{C}3'$ carbons) and 138.9 ppm ($\text{C}8$ and $\text{C}6$ carbons). Internucleotide NOE peaks are labeled by the number of each of the nucleotides within the primary sequence and the type of proton involved. Intramolecular NOE peaks are labeled by the type of proton involved. Intramolecular pyrimidine H5-H6 NOE peaks are labeled by simply the nucleotide number.

proton donors in Watson-Crick base pairs were found to resonate slightly downfield of N3 nuclei in unpaired uridines, whereas most of the N1 and N3 resonances of the adenosines that are proton acceptors in Watson-Crick base pairs were found to have relatively upfield chemical shifts (Fig. 5). The results for A8 (which is not base paired) and A20 (base paired, at the end of stem1) suggest that ^{15}N chemical shift is not an unambiguous indicator of hydrogen bonding. Perhaps other effects such as base stacking or fraying at the end of a helix contribute to the ^{15}N chemical shift as well.

Structure determination

NMR-derived structural constraints

Nuclear Overhauser effect (NOE) spectra were used to obtain a set of structural constraints, summarized in Figure 6. The constraints consist primarily of inter-proton distance bounds derived from ^1H - ^1H two-dimensional NOE and ^{13}C - ^1H - ^1H three-dimensional HMQC-NOE spectra. In addition, imino and amino protons that participate in hydrogen bonds within Watson-Crick base pairs were assigned to fixed distances from

the oxygen or nitrogen atoms that serve as proton acceptors in the hydrogen bonds. Ribose groups that were clearly in the $\text{C}2'$ *endo* or $\text{C}3'$ *endo* conformation, identified by characteristic ^1H - ^1H coupling constants, were constrained as such. χ torsion angles were constrained to remain within a 40° range centered on the ideal value for *anti* conformation, except in the case of nucleotide G1, which was restrained to remain in an 80° range of *syn*, based on its exceptionally strong H8-H1' NOE cross peak. The NOE cross peaks involving the resonances associated with the central region of each of the two stems were characteristic of regular A-form helical RNA. These A-form helical regions of the molecule (namely, the central five Watson-Crick base pairs in stem2 and four of the base pairs in stem1) were therefore constrained as such during the structure analysis (Fig. 6).

In some recent NMR studies of RNA molecules, ^{13}C - ^{31}P and ^{31}P - ^1H scalar couplings and ^{31}P chemical shifts have sometimes been used to obtain constraints on the allowed ranges for torsion angles along the ribose-phosphate backbone (e.g., Allain et al., 1995; Address et al., 1997). ^{13}C - ^{31}P scalar couplings can, in principle, be obtained from the fine structure of correlation peaks in ^{13}C - ^1H CT-HSQC spectra, and ^{31}P chemical shifts

FIGURE 4. Sections of two-dimensional ^1H - ^1H NOE spectra of the gene 32 mRNA pseudoknot. **A:** A region of the spectrum of the fully protonated pseudoknot; **B:** An identical region of the spectrum of a sample that is deuterated at all of the adenosine H8 positions. Comparison of **A** and **B** permits clear confirmation of the adenosine H8 resonance assignments. **C:** A region of the spectrum of the fully protonated pseudoknot. **D:** An identical region of the spectrum of a sample that is deuterated at all of the guanosine H8 positions. Comparison of **C** and **D** permits clear confirmation of the guanosine H8 resonance assignments.

can be obtained from ^1H - ^{31}P correlated spectra. These spectra were obtained for the gene 32 mRNA pseudoknot, however, in the present work the resonance line-widths and extensive spectral overlap made the quantitative analysis of the ^{13}C - ^{31}P and ^{31}P - ^1H scalar couplings and ^{31}P chemical shift assignments relatively unreliable. For this reason, it was considered most conservative to not use ^{13}C - ^{31}P and ^{31}P - ^1H scalar couplings and ^{31}P chemical shifts to derive torsion angle constraints in the present work.

Modeling of the pseudoknot structure

One goal of the present work is to determine a model of the pseudoknot structure. A second and equally im-

portant goal is to determine the complete range of all possible structures that are consistent with the NMR-derived structural constraints, while maintaining reasonable molecular geometry in terms of bond lengths and angles. This second goal provides insight into the precision of the structural model. To accomplish these two goals, a large number of starting structures (84 in total) of the pseudoknot were generated for use in a simulated annealing protocol subject to NMR-derived structural constraints. The set of starting structures included unfolded molecules with randomized torsion angles, as well as pseudoknots differing radically in the conformations of the loop regions, the conformations of the nucleotides at the junction of the two stems, and pseudoknots with a variety of bend angles between the

FIGURE 5. A section of the ^{15}N - ^1H HMBC spectrum showing cross peaks correlating H2 with N1 and N3 for the ten adenosines in the gene 32 mRNA pseudoknot at pH 6.5 and 20°C. The ^{15}N chemical shifts of the N1 and N3 resonances of adenosine triphosphate at pH 6.5 and 20°C are also indicated.

two helical stems. For example, the set of starting structures included pseudoknots that were bent at angles of -50° and $+50^\circ$ (and intermediate values), with the bending occurring by rotation about each of three possible axes, as well as starting structures with stem1 translated relative to stem2.

Each of the starting structures for the pseudoknot was subjected to a simulated annealing protocol within the X-PLOR program suite (Brünger, 1993). Starting structures that reached a minimum value of the X-PLOR energy function during the simulated annealing procedure were accepted as possible correct models. When starting models failed to reach the energy minimum within a single round of simulated annealing, the procedure was repeated up to ten times, using different values for the initial trajectories of the atoms. If the energy minimum was reached, the structure was added to the set of possible correct models. A final filter was then applied to screen the possible correct models: The angle formed by the three atoms comprising each hydrogen bond within the twelve Watson-Crick base pairs was evaluated. If this angle exceeded 17° for hydrogen bonds involving imino protons or exceeded 23° for hydrogen bonds involving amino protons, the structure was rejected. Structures rejected by this final filter were subjected to up to ten more rounds of simulated annealing, the hydrogen bonds were re-evaluated, and if the angle and energy criteria were satisfied, the structure was accepted as a possible cor-

rect model. The 17° and 23° criteria are based on studies of observed hydrogen bond angles in nucleic acid crystal structures (Voet & Rich, 1970; Saenger, 1984).

The procedure described above resulted in a set of 19 acceptable structures for the pseudoknot structure being generated. We believe that these are a fair representation of the full range of structures that are simultaneously consistent with the experimental NMR data and have reasonable molecular geometry. Views of these structures are shown in Figures 7 through 9, and are discussed below.

Description of the structure

Junction of the two helical stems

The structure of the C7-G16 and U28-A15 base pairs at the junction of the two helical stems of the pseudoknot is well defined by the NMR data, because of the substantial number of obtained structural constraints for this region of the molecule. The range of structures that are consistent with the NMR data and have the minimum value of the X-PLOR energy function is relatively small, with the root mean square deviation (rmsd) for the four junction nucleotides being 0.75 Å for the set of 19 structures. The structural models obtained for this region of the pseudoknot do not appear to be strongly influenced by the choice of constraint set. For example,

Xplor Distance Constraints

FIGURE 6. A summary of the distribution of the NMR-derived interproton distance constraints used in the structural analysis of the gene 32 mRNA pseudoknot. ^1H - ^1H NOEs were converted to allow interproton distance ranges as described in Materials and Methods. The central five Watson–Crick base pairs in stem2 and four of the base pairs in stem1, which were clearly A-form helical (identified by characteristic regular NOE patterns) were constrained as such.

when the set of NMR-derived constraints contains no lower distance bounds for the internuclear distances (other than van der Waals distances), a nearly identical set of structures was obtained, with no significant difference in the rmsd of the atomic coordinates.

The C7–G16 and U28–A15 base pairs are each of the Watson–Crick type, with the ring planes of all four of the bases nearly parallel. The conformations of the ribose rings of these four nucleotides are C3' *endo*, as occurs in regular A-helical RNA. However, the stacking of C7–G16 upon U28–A15 differs from what occurs in an A-form helix: The C7–G16 base pair is rotated by $50 \pm 6^\circ$ relative to U28–A15, which is significantly greater than the 32° rotation per base pair rise in regular A-form helical RNA (Fig. 7a). This overrotation of approximately 18° serves to relieve what would otherwise be an unacceptably close contact between the phosphates of A8 and U28 at the junction of the two stems, while preserving much of the base stacking (particularly the stacking of the purines A15 and G16) and the accompanying stabilizing effects. Since C6–G17 and G13–C29 are stacked adjacent to C7–G16 and U28–A15, respectively, in a near A-helical manner, the net effect is that stem1 is rotated relative to stem2 about an axis that is parallel to the helical axis. The axis of rotation of stem1 relative to stem2 does not pass through the cen-

ter of the base pairs, but is located near the ribose backbone of A15 and G16, resulting in the interproton distances between A15 and G16 being very near to the normal A-helical values (Fig. 7a). This is consistent with the intensities of the NOE cross peaks being between the H8 proton of G16 and the ribose protons of A15 being similar in intensity to the normal sequential base-to-ribose NOE cross peaks in an A-form helix, and also accounts for the A-helix-like sequential imino-to-imino NOE between G16 and U28 (Du et al., 1996). In contrast, the interproton distances between nucleotides C7 and U28 are relatively long, accounting for the absence of observed NOE cross peaks between the H6 proton of C7 and the ribose protons of U28. These observations are entirely consistent with the results reported by Puglisi et al. (1990) in an NMR study of an RNA pseudoknot with coaxial stems, where A-helix-like sequential NOE cross peaks were observed between protons of the nucleotides in positions analogous to A15 and G16.

The bend angle between the helical axes of the two stems was calculated with the assistance of the program NEWHEL93 (R.E. Dickerson, University of California, Los Angeles). Bend angle was defined as the angle between the line defining the helix axis of stem1 (calculated using the coordinates of the U3–A20, G3–C19, A5–U18, and C6–G17 base pairs) and the line defining the helix axis of stem2 (calculated using the coordinates of the central five base pairs of stem2). It was found that pseudoknot structures that appear quite similar at the C7–G16 and U28–A15 base pairs at the junction of the two stems can have significantly different bend angles. For example, the six structures superimposed in Figure 7 appear similar, however when the helical axes for these same six structures were calculated, it was found that the bend angles are distributed over a substantial range (displayed in Fig. 8). In some of the structures, the helix axes of stem1 and stem2 are nearly parallel, while in other structures the angle between the helix axes is as great as 30° . These observations provide clear examples of a strength of NMR as well as an inherent weakness: A set of short-range interproton distance ranges are suitable for determining a local structure, while structures that differ significantly in a large-scale feature, such as bending, can satisfy the same set of short-range restraints. Although a feature of the global structure (i.e., bending) differs significantly in the six models superimposed in Figure 8, the structures are quite similar on a local level (as can be seen in Fig. 7). We conclude that distances derived from NOE data can not be used to determine a precise bend angle, unless unreasonably precise values are assigned for the distance bounds.

Nucleotide A8 of Loop1

Since the C7–G16 and G9–C34 Watson–Crick base pairs are clearly defined by the NMR data, and stable as

FIGURE 7. Views of the structure of nucleotides C7, A15, G16, and U28 forming the two base pairs at the junction of the two helical stems of the gene 32 mRNA pseudoknot. **A:** Six superimposed structures of the C7-G16 and A15-U28 base pairs, representative of the full range of structures that are consistent with the NMR-derived structural constraints and have reasonable molecular geometry. For comparison, the C7-G16 and A15-U28 base pairs are shown in the lower part of the figure, as they would be if they were part of a single ideal A-form helix. The distances between phosphorous atoms are indicated in the pseudoknot and in the standard A-form helix. In the pseudoknot, the C7-U28 interphosphorous distance is significantly extended, to relieve what would otherwise be a close contact between the phosphate groups of A8 and U28. **B:** A different view of the same four nucleotides in the same six superimposed structures as shown in **A**, but rotated by 90°. Again, a view of the four nucleotides as they would be in an ideal A-form helix is shown for comparison in the lower part of the figure.

indicated by their imino proton exchange rates, it becomes a requirement that the single nucleotide A8 must cross the major groove of the seven-base-pair length of stem2. Although the observation of some internucleotide NOEs involving A8 were reported in our earlier work (Du et al., 1996), with the current improved resonance assignments we conclude that only two intranucleotide (and no internucleotide) NOEs are reliably observed and assigned. In particular, no long-range NOEs to the H2 proton of A8 could be unambiguously identified, as might be expected if A8 participates in loop-stem interactions (Kolk et al., 1998). As a result, the position of A8 is not as well defined as other parts of the pseudoknot structure (rmsd of 4.2 Å), and a relatively wide range of positions inside the major groove of stem2 are consistent with the NMR-derived constraints and X-PLOR energy function. A mutational study of the pseudoknot has shown that any nucleotide (C, G, or U) can substitute at the A8 position, providing evidence against any essential tertiary interactions (Du & Hoffman, 1997), however a recent thermodynamic study reveals that pyrimidine substitutions of A8 (and the analogous nucleotide in the T4 pseudoknot) are destabilizing by up to 1.2 kcal mol⁻¹ in free energy at 37 °C (Theimer et al., 1998). Interestingly, the upfield ¹⁵N chemical shift observed for N1 of A8, but not N3, suggests that N1 may participate in a tertiary hydrogen

bond, which could account for the destabilization of the T2 and T4 gene 32 mRNA pseudoknots (Theimer et al., 1998).

Although the secondary structure of the pseudoknot shown in Figure 1 suggests that a rather large distance must be spanned by A8 to link C7 with G9 at opposite ends of stem2, this is not the case. The natural twist of the A-form helix of stem2 is such that the 3' end of C7 and the 5' end of G9 are relatively close together, and the distance between them can be spanned by the single A8 nucleotide crossing the major groove of stem2. We have inspected the bond angles of the A8 nucleotide in the accepted structures, and found that they occur within the ranges previously observed in crystallographic studies of RNAs. Therefore, there is no requirement that the A8 nucleotide be "strained" in spanning the major groove of the seven-base-pair stem2.

3' end of the pseudoknot (nucleotides G9 and C33–C36)

The G9-C34 Watson–Crick base pair is stacked adjacent to C10-G33 in the A-form helical part of stem2, with planes of the bases of G9 and C34 near parallel to the planes of C10 and G33 (Fig. 9). Near A-helical base stacking continues through A35 and C36, extending from the 3' end of stem2, even though these nucleo-

FIGURE 8. Two orthogonal views of the same six superimposed structures of the gene 32 mRNA pseudoknot as shown in Figure 7. Only the five base pairs of stem1 and the seven base pairs of stem2 are shown; stem1 is at the top of the molecule as viewed. The superpositioning was performed using the nucleotides of the central five Watson–Crick base pairs in stem2. The helix axes for the two stems in each of the six structures are indicated as lines. The helix axes of stem2 are nearly overlapping, since the superpositioning was performed using the nucleotides of this stem. The bend angles for the pseudoknots shown are between 8 and 30°. Pseudoknots with bend angles between 0 and 30° are equally consistent with the NMR data as well as the X-PLOR energy function.

tides are not base paired. The conformations of nucleotides G9, C34, and A35 are relatively well defined by the NMR data (the rmsd in the coordinates is 0.74 Å). The position of nucleotide C36 is not as well defined by the NMR-derived constraints (with an rmsd of 2.6 Å), however, it does appear that C36 deviates somewhat from ideal A-form helical geometry, as some unusual NOE cross peaks are observed, such as A35 H8 to C36 H5, H1' and H2', and A35 H1' to C36 H2'. The stacking of A35 below the G9–C34 base pair accounts for the protection of the G9 imino proton from fast exchange with the solvent (Du et al., 1996). A35 should be considered to be an integral part of the minimal pseudoknot structure, since deletion of the unpaired nucleotides at the 3' end of the pseudoknot significantly reduced stability (Qiu et al., 1996; Theimer et al., 1998). The ribose conformations are C3' *endo* for G9, C34, and A35, whereas C36 is in a mixed conformation.

Loop2 and the 5' end of the pseudoknot (nucleotides 1, 2, and 21–27)

In the previous homonuclear NMR study of the pseudoknot, the assignments of only about half of the nuclei in

this region of the molecule were obtained (Du et al., 1996). In the present work, resonance assignments for the ^1H and ^{13}C nuclei for these nucleotides are nearly complete, and several corrections to the previous assignments were made. Despite the increase in the number of resonance assignments, the structure of loop2 is not as well defined by the NMR data as other regions of the molecule (with a 3 Å rmsd for the coordinates of nucleotides 1, 2, and 21–27), and the description of this region of the molecule must remain qualitative. The reasons for the relatively poor definition of this region of the molecule are twofold. First, distance constraints for the amino and imino protons could not be obtained because of the fast exchange rate of these protons with the solvent. Second, the majority of the 66 assigned NOE peaks are either intranucleotide or between sequential nucleotides, and therefore are not as useful in defining tertiary structural features. However, the topology of the pseudoknot, defined by the clearly observed Watson–Crick base pairings in the stems, enforces the requirement that loop2 cross the distance from A20 to U28, located at opposite ends of stem1. This by itself imposes substantial constraints as to the possible locations of the loop2 nucleotides.

FIGURE 9. **A:** A stereo pair of a structural model of the gene 32 mRNA pseudoknot. The structure is one of the six shown in Figure 8. The bend angle of this pseudoknot model is 12° , defined as the angle between the helix axes of stem1 and stem2. **B:** A ribbon diagram showing the path of the sugar-phosphate backbone in the same structural model of the gene 32 mRNA pseudoknot. The phosphate backbone is color ramped from the 5' end (red) to the 3' end of the molecule (blue).

All of the assigned NOE cross peaks for the loop2 nucleotides are consistent with the following features. The first nucleotide of loop2, U21 is stacked on A20, defined by eight NOE cross peaks associated with nonexchangeable protons. The phosphate backbone changes direction at nucleotide A22, a conclusion supported by a weak nonsequential NOE cross peak between the H5 protons of U21 and C23. NOE cross peaks between the base and ribose protons of the sequential nucleotides C23, A24, U25, and C26 are consistent with the bases of these nucleotides being stacked. Nonsequential NOE cross peaks between A24 and nucleotide C19 of stem1 (specifically, from A24 H1' to C19 H5 and H6, and A24 H2' to C19 H6) are consistent with this loop2 nucleotide being in the minor groove of stem1. The position of G27 at the 3' end of loop2 is similar to that described for the nucleotide in the analogous position in a TYMV RNA pseudoknot, with its base turned at nearly a 90° angle to the planes of the bases in stem1 (Kolk et al., 1998). Near the 5' end of the gene 32 pseudoknot, an NOE cross peak between H2' of G1 and H8 of A22 indicates that these two nucleotides are near each other within the structure. G1 is in the *syn* conformation, indicated by a very strong NOE cross peak between its H8 to H1' protons. Only intranucleotide NOE cross peaks are observed for C2, suggesting that it is bulged, or at least not stacked between the adjacent nucleotides G1 and U3.

DISCUSSION

The NMR results described in the present work permit a description of the gene 32 mRNA pseudoknot that is significantly more detailed than that provided by the previous study based on homonuclear NMR data (Du et al., 1996). However, all regions of the pseudoknot structure are not equally well defined, a result of the uneven distribution of NMR-derived structural constraints. A particularly interesting (and well-defined) feature of the pseudoknot is the junction of the stems, where the two A-form helices are coaxially stacked to form a single quasicontinuous helix. Coaxial stacking of helical stems is a recurrent theme in RNA structure, and has been observed in tRNAs, the hammerhead ribozyme, and at the 3' end of plant viral genomic RNA, in addition to the present NMR study of the gene 32 mRNA pseudoknot. A comparison of the coaxial stacking of helices that occurs in these different RNA structures reveals some striking similarities.

One parameter that is useful in describing the stacking of two coaxial RNA stems is the bend angle between the helix axes. In the present work, we have concluded that NMR data is a modestly precise indicator of pseudoknot bending, with bend angles of $15 \pm 15^\circ$ being consistent with the NMR-derived constraints and the X-PLOR energy function. For comparison purposes, we obtained the coordinates of several RNA

molecules with coaxial stems from the Protein Data Bank, including the hammerhead ribozyme and several tRNA structures, and analyzed the angle between the helical axes of the coaxial stems using the program NEWHEL93 (R.E. Dickerson, University of California, Los Angeles). The hammerhead ribozyme, which contains non-Watson–Crick base pairing at the junction of two coaxial stems, has an angle between the lines defining the helix axes of 9° (PDB reference number 1hnh; Pley et al., 1994). The angle between the helix axes of the coaxially stacked T-arm and amino acid acceptor stem in yeast tRNA^{phe} is $13\text{--}16^\circ$ in three independent crystal structures, whereas the angle between the helix axes of the stacked D-arm and anticodon arm is $24\text{--}26^\circ$ (PDB reference numbers 4tra, 4tna, and 6tna; Westoff et al., 1988; Hingerty et al., 1978; Sussman et al., 1978, respectively). This larger bend angle can be attributed to non-Watson–Crick interactions among the stacked bases at the junction of the helical stems. The angle between the helix axes of the stacked T-arm and amino acid acceptor stem in tRNA^{asp} is between 4 and 7° , as found in two different crystal structures, whereas the angle between the helix axes of the stacked D-arm and anticodon arm is $15\text{--}16^\circ$ (PDB reference numbers 2trna and 3trna; Westhof et al., 1988). We conclude that in terms of its bend angle of $15 \pm 15^\circ$, the quasi-continuous A-helix formed by the two stems of the gene 32 mRNA pseudoknot is similar to these other RNAs with coaxial or near-coaxial stacked stems.

Although there have not yet been any X-ray crystallographic structures published for RNA pseudoknots, several pseudoknots have now been investigated by NMR methods, and in most cases they have been described as having coaxial stacked stems. These coaxial pseudoknots include the first pseudoknot characterized by NMR methods (Puglisi et al., 1990), as well as a pseudoknot-containing structure at the 3' end of a TYMV with three coaxial stacked A-helical stems (Kolk et al., 1998). A pseudoknot associated with -1 frameshifting in the mouse mammary tumor virus *gag-pro/pol* mRNA has been described as having a bend angle of approximately 68° between the two stems, which was attributed to the effect of an unpaired adenosine nucleotide (Shen & Tinoco, 1995; Chen et al., 1996). The pseudoknot associated with the -1 *gag-pro/pol* frameshift site in simian retrovirus-1 is reported to be similar to the coaxial gene 32 mRNA pseudoknot described in the present study (Du et al., 1997). Coaxial stacking of helical stems is clearly a common theme in RNA structure, consistent with base stacking being a significant stabilizing force in nucleic acids.

Another parameter that is useful in describing the stacking of helical stems is the angle of rotation of stem1 relative to stem2 about an axis parallel to the helix axes. The rotation-per-base-pair rise at the junction of the T-arm and amino acid acceptor stem in five different tRNA crystal structures is between 38 and

46°. This is greater than the 32° rotation-per-base-pair rise that occurs in an ideal A-form helix, and is similar to the rotation-per-base-pair rise at the junction of the helical stems in the gene 32 mRNA pseudoknot ($50 \pm 6^\circ$ at the step from the A15-U28 to C7-G16 base pairs). The overrotation at the junction of the pseudoknot stems relieves what would otherwise be a close phosphate-phosphate contact, while keeping the position the C7 nucleotide in stem1 near the minimum distance from G9 in stem2, so that the single nucleotide of loop1 can easily span the major groove of the seven-base-pair stem2.

In summary, the present study has provided an improved structural model of the pseudoknot within the gene 32 mRNA of bacteriophage T2. The model provides further insight into the structural details and stabilizing features of this pseudoknot motif, which appears to be widespread in nature. Atomic coordinates for the structure of the gene 32 mRNA pseudoknot shown in Figure 9 have been submitted to the Protein Data Bank.

MATERIALS AND METHODS

RNA synthesis and purification

The 36-nt gene 32 mRNA pseudoknot with the sequence shown in Figure 1 was transcribed *in vitro* using T7 RNA polymerase and synthetic DNA templates, as previously described (Du et al., 1996). The RNA product was separated from transcripts of incorrect size by electrophoresis on 20% polyacrylamide gels under denaturing conditions in 8 M urea. RNA was visualized by UV shadowing, and removed from the gel using a BioRad model 422 electroeluter. The RNA was further purified by repeated ethanol precipitation, and finally passed through a Sephadex G25 gel filtration column in 1 mM phosphate buffer and lyophilized. A typical yield was 1 mg of purified RNA per 10 mL transcription volume. Samples of ^{13}C and/or ^{15}N enriched pseudoknot were prepared by including isotopically labeled NTPs in the *in vitro* transcription reactions. Labeled ribonucleotides were isolated from *E. coli* grown on minimal media using ^{15}N ammonium chloride as the nitrogen source for the pseudoknot enriched in ^{15}N , or ^{13}C sodium acetate and ^{15}N ammonium chloride as the carbon and nitrogen sources, respectively, for pseudoknots enriched in both isotopes (Hoffman & Holland, 1995). Ribonucleotide triphosphates were prepared essentially as described by Batey et al. (1992) and Nikonowicz et al. (1992).

Deuteration of ATP and GTP

ATP and GTP deuterated at the H8 position were prepared by placing 1 g of ATP and GTP in separate tubes and dissolving each in 8 mL of 99.9% D_2O . The pH of each solution was raised to 12 by the addition of 2 M KOD, and the solutions placed in an oven at 65 °C for 16 h. NMR analysis of the two solutions showed that at this point approximately 85% of the ATP and 95% of the GTP had deuterium incorporated at

the H8 position. The two solutions were lyophilized after lowering the pH to 7.5 with HCl. The partially deuterated purines were then redissolved in D_2O , the pH raised to 12, and they were subjected to another heating cycle. This second heating cycle resulted in approximately 99% and 97% deuteration of the H8 positions on GTP and ATP, respectively, as judged by proton NMR. ^{31}P NMR spectra and HPLC showed that there was little degradation of the triphosphates. The pH of each ribonucleotide solution was lowered to 7 using HCl and diluted to 40 mM with H_2O before use in *in vitro* transcription reactions. The deuterated nucleotides were used to make two pseudoknot samples. Each pseudoknot contained either deuterated adenosine or deuterated guanosine, but not both. Quantities of the RNA pseudoknots suitable for NMR studies were obtained from 100 mL of *in vitro* transcription. Yields from the transcriptions were approximately 0.8 mg of RNA per 10 mL reaction.

NMR spectroscopy

NMR spectra of the RNA pseudoknots were collected at 500 MHz using a Varian Unity-Inova spectrometer. Samples typically contained 10–20 mg of RNA dissolved in 550 μL of 10 mM Na/K phosphate buffer in either 90% H_2O /10% D_2O or 99.9% D_2O at pH 6.0–6.8, unless otherwise stated. Spectra in 90% H_2O /10% D_2O were acquired using the jump and return method for solvent suppression (Plateau & Guéron, 1982). Spectra in 99.9% D_2O were obtained using presaturation for solvent suppression. Most spectra in D_2O solvent were acquired at either 20 or 25 °C. Data acquired at different temperatures between 10 and 30 °C resulted in relatively small temperature-dependent shifts in proton resonance frequencies; these shifts were in some cases useful in identifying nearly overlapping resonances. Spectra in 90% H_2O /10% D_2O solvent were acquired at 1, 10, and 20 °C.

^1H homonuclear NMR spectra

Two-dimensional NOE spectra in D_2O solvent were obtained using mixing times of 60, 100, 200, and 280 ms. Spectra were typically obtained with 4,500 Hz sweep width, a 3-s recycle time, and contained 2,048 complex points in the acquisition dimension and 400 complex points in the indirect dimension. Two-dimensional NOE spectra in H_2O solvent were obtained using a mixing time of 200 ms, with a sweep width of 12,204 Hz, and contained 2,048 complex points in the acquisition dimension and 300 complex points in the indirect dimension. $\text{H1}'$ to $\text{H2}'$ coupling constants were estimated from the separation of the antiphase components in a high resolution 2QF-COSY spectrum obtained with a sweep width of 2,600 Hz, and 2,048 complex points in the acquisition dimension and 350 complex points in the indirect dimension. ^1H - ^1H TOCSY spectra were obtained with mixing times of 60 and 100 ms. A three-dimensional homonuclear NOESY-TOCSY spectrum (Vuister et al., 1988; Hoffman & White, 1995) was acquired in D_2O solvent at 25 °C, with a sweep width of 4,500 Hz in each of the three dimensions, and 100 tiers of 256 by 1,024 real points, with an NOE mixing time of 200 ms and a TOCSY mixing time of 38 ms. Eight scans were acquired for each free induction decay, with a recycle time of 1.4 s between scans.

Heteronuclear NMR spectra

All ^{13}C correlated spectra were obtained using a ^{13}C - ^{15}N double-labeled sample of the gene 32 mRNA pseudoknot. One-bond ^{13}C - ^1H correlations were obtained from a set of three CT-HSQC spectra (Santoro & King, 1992), with constant time delays of 22, 46, and 70 ms, using 256, 550, and 800 real points, respectively, in the indirect dimension. A sweep width of 6,308.2 Hz was used for ^{13}C , centered at 84 ppm, so that the spectrum was folded once in the indirectly detected ^{13}C dimension. Additional ribose ^1H and ^{13}C resonance assignments were obtained using a three dimensional ^{13}C - ^1H HCCH-COSY spectrum (Fesik et al., 1990; Kay et al., 1990), obtained with sweep widths of 2,000 Hz for the proton dimensions and 6,425 Hz centered at 85 ppm for the ^{13}C dimension. A three-dimensional ^{13}C - ^1H NOESY-HMQC spectrum (Marion et al., 1989; Zuiderweg & Fesik, 1989) consisting of 64 tiers of 256 by 512 real points was obtained with an NOE mixing time of 200 ms. Sixteen scans were acquired for each free induction decay, with a recycle time of 1.5 s between scans. The sweep within each ^1H dimension was 4,000 Hz, and the ^{13}C sweep width was 6,425 Hz, centered at 85 ppm, so that the spectrum was folded once in the ^{13}C dimension.

All ^{15}N correlated spectra were obtained using a uniformly ^{15}N -enriched sample. Long-range ^{15}N - ^1H correlations were detected using two-dimensional heteronuclear multiple bond coherence (HMBC) spectra, acquired by the method of Bax & Marion (1988). Six hundred FIDs of 4,096 real points were acquired for each spectrum. Sweep widths in the proton and nitrogen dimensions were 5,000 and 6,000 Hz, respectively, with the center of the nitrogen spectrum at 197 ppm. Spectra were acquired at 20 °C, with a 30-ms magnetization transfer delay chosen to optimize detection of the relatively large 15-Hz couplings, and at 28 °C with a 60 ms delay chosen to optimize detection of smaller two- and three-bond couplings. The total acquisition times were 38 and 70 h for the spectra acquired at 20 and 28 °C, respectively. One bond ^1H - ^{15}N correlations were obtained from an HMQC spectrum acquired in 90% $\text{H}_2\text{O}/10\%$ D_2O solvent using the jump and return method for solvent suppression.

^1H , ^{13}C , and ^{15}N chemical shifts are referenced to internal 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) as recommended by Wishart et al. (1995). ^1H chemical shifts are referenced to the DSS at 0 ppm. The 0-ppm ^{13}C and ^{15}N reference frequencies were determined by multiplying the 0-ppm ^1H reference frequency by 0.251449530 and 0.101329118, respectively. Chemical shifts are reported for a sample in 10 mM sodium/potassium phosphate buffer, pH 6.5, 20 °C.

Structure determination

A total of 176 NOE-derived distance constraints distributed as shown in Figure 6 were used in defining the structure of the gene 32 mRNA pseudoknot. Seventy-four of these constraints were derived from an NOE spectrum acquired using a relatively short mixing time of 60 ms to minimize the effects of spin diffusion on the NOE intensities. Depending on its intensity, each NOE was placed into one of four categories ranging from very strong to weak, and assigned to interproton distance bounds as follows: Strong (1.8–3.5 Å), medium (2.5–4.5 Å), weak (3.5–5.5 Å) and very weak (4.5–6.5 Å). Ninety-three NOEs were detected in spectra with mixing times

of up to 200 ms, and these were assigned to distance ranges of 4.5–6.5 Å. These NOE/distance assignments were internally verified by comparing with NOE cross peaks arising from the regular A-form stem regions of the molecule, where the interproton distances are known. Nine distance constraints were derived from peaks observed in the three-dimensional ^1H - ^1H - ^{13}C NOESY-HMQC spectrum, but obscured by overlap in all of the two-dimensional data sets. These were assumed to correspond to an interproton separation of less than 5.0 Å for the NOE transfer. No intraribose or intrabase NOE-derived distance constraints were used. Torsion angle constraints were used to restrict riboses 3–7, 9–20, and 28–35 to the C3' *endo* conformation, based on their small H1' to H2' coupling constants. The ribose of G1 was restricted to the C2' *endo* conformation, based on its intraribose coupling constants. The torsion angles of the other riboses were not restricted. The glycosidic torsion angle χ of G1 was restricted to a *syn* ($60 \pm 40^\circ$) conformation. The χ angles for all other nucleotides were restricted to an *anti* ($-170 \pm 20^\circ$) conformation. Hydrogen bonds within the 12 Watson–Crick base pairs were introduced as distance constraints between the atoms participating in the hydrogen bond. No angle constraints were used to force the bases within each pair to be coplanar. Within stem1, distance constraints were used to set the interatomic distances between nucleotides 3–6 and 17–20 to ideal A-form helical values (Arnott et al., 1972). Distance constraints were used to set the interatomic distances between nucleotides 10–13 and 29–33 in stem2 to ideal A-form helical values. If the C6-G17 and G13-C29 base pairs are defined using the NMR data rather than restricted to A-helical geometry, there is not a substantial change in the relative positions determined for the C7-G16 and A15-U28 base pairs at the junction of the helical stems.

The structure determination of the pseudoknot was carried out using the hybrid distance geometry-simulated annealing and energy minimization protocols within the X-PLOR version 3.1 program suite (Brünger, 1993), with the goal of identifying the full range of structures that are consistent with the distance and angle constraints describe above, while having reasonable molecular geometry, consistent with a minimum value of the X-PLOR energy function. Square-well potentials were used for interproton distance constraints and torsion angle constraints, so that there was no penalty for interproton distances or torsion angles that were within the defined bounds. A total of 84 starting structures were used, including extended structures that contained randomized torsion angles, as well as structures differing radically in the conformations of the loop regions, the junction of the two stems, and in the bend angle between the helical stems. After simulated annealing, models of the pseudoknot structure were considered acceptable if the following criteria were met: (1) the minimum value of the X-PLOR energy function was reached; (2) no interproton distance constraint violations exceeded 0.5 Å; (3) no torsion angle constraint violations exceeded 8° ; (4) the angle formed by the three atoms comprising each hydrogen bond involving imino protons was less than 17° ; (5) the angle formed by the three atoms comprising each hydrogen bond involving amino protons was less than 23° . The angle criteria for hydrogen bonds is based on the values observed in nucleic acid crystal structures (Voet & Rich, 1970; Saenger, 1984). The simulated annealing procedure was repeated up to ten times for starting structures that failed to

satisfy the acceptance criteria, using different initial trajectories for the atoms. The procedure described above resulted in the generation of 19 accepted models for the pseudoknot structure. The influence of the choice of distance bounds on the structure determination was evaluated by repeating the simulated annealing procedure using no lower bounds on the NOE-derived interproton distances. This resulted in a set of structures that was distributed over a range that was indistinguishable from that obtained using lower distance bounds.

ACKNOWLEDGMENTS

We thank Dr. David P. Giedroc for many helpful discussions. This work was supported by NIH grant R01-AI40187 (to D.W.H. and D.P.G.), a grant from the American Cancer Society (to D.W.H.) and a grant from the Robert A. Welch Foundation (to D.W.H.).

Received August 14, 1998; returned for revision October 2, 1998; revised manuscript received October 19, 1998

REFERENCES

- Aboul-ela F, Karn J, Varani G. 1995. The structure of the human immunodeficiency virus type-1 TAR RNA reveals principles of RNA recognition by Tat protein. *J Mol Biol* 253:313–332.
- Addess K, Basilion JP, Klausner RD, Rouault TA, Pardi A. 1997. Structure and dynamics of the iron responsive element RNA: Implications for binding of the RNA by iron regulatory proteins. *J Mol Biol* 274:72–83.
- Allain FH-T, Varani G. 1995. Structure of the P1 helix from group 1 self-splicing introns. *J Mol Biol* 250:333–353.
- Arnott S, Hukins DWL, Dover SD. 1972. Optimized parameters for RNA double-helices. *Biochem Biophys Res Commun* 48:1392–1398.
- Batey RT, Inada M, Kujawinski E, Puglisi JD, Williamson, JR. 1992. Preparation of isotopically labeled ribonucleotides for multidimensional NMR spectroscopy. *Nucleic Acids Res* 20:4515–4523.
- Bax A, Marion D. 1988. Improved resolution and sensitivity in ^1H -detected heteronuclear multiple-bond correlation spectroscopy. *J Mag Reson* 78:185–191.
- Benard L, Philippe C, Dondon L, Grunberg-Manago M, Ehresmann B, Ehresmann C, Portier C. 1994. Mutational analysis of the pseudoknot structure of the S15 translational operator from *Escherichia coli*. *Mol Microbiol* 14:31–40.
- Brünger AT. 1993. *X-PLOR version 3.1: A system for X-ray crystallography and NMR*. New Haven, Connecticut: Yale University Press.
- Büchner P, Maurer W, Rüterjans H. 1978. Nitrogen-15 nuclear magnetic resonance spectroscopy of ^{15}N -labeled nucleotides. *J Mag Reson* 29:45–63.
- Chen X, Chamorro M, Lee SI, Shen LX, Hines JV, Tinoco I, Varmus HE. 1995. Structural and functional studies of retroviral RNA pseudoknots involved in ribosomal frameshifting: Nucleotides at the junction of the two stems are important for efficient ribosomal frameshifting. *EMBO J* 14:842–852.
- Chen X, Kang H, Shen LX, Chamorro M, Varmus HE, Tinoco I. 1996. A characteristic bent conformation of RNA pseudoknots promotes –1 frameshifting during translation of retroviral RNA. *J Mol Biol* 260:479–483.
- Deckman IC, Draper DE. 1987. S4-alpha mRNA translation regulation complex. II. Secondary structures of the RNA regulatory site in the presence and absence of S4. *J Mol Biol* 196:323–332.
- Dieckmann T, Feigon J. 1997. Assignment methodology for larger RNA oligonucleotides: Application to an ATP-binding RNA aptamer. *J Biomol NMR* 9:259–72.
- Du Z, Giedroc DP, Hoffman DW. 1996. Structure of the autoregulatory pseudoknot within the gene 32 messenger RNA of bacteriophages T2 and T6: A model for a family of structurally related RNA pseudoknots. *Biochemistry* 35:4187–4198.
- Du Z, Hoffman DW. 1997. An NMR and mutational study of the pseudoknot within the gene 32 mRNA of bacteriophage T2: Insights into a family of structurally related RNA pseudoknots. *Nucleic Acids Res* 25:1130–1135.
- Du Z, Holland JA, Hansen MR, Giedroc DP, Hoffman DW. 1997. Base-pairings within the RNA pseudoknot associated with the Simian Retrovirus-1 gag-pro frameshift site. *J Mol Biol* 271:463–470.
- Fesik SW, Eaton HL, Olejniczak ET, Zuiderweg ERP. 1990. 2D and 3D NMR spectroscopy employing ^{13}C - ^{13}C magnetization transfer by isotropic mixing. Spin system identification in large proteins. *J Amer Chem Soc* 112:885–888.
- Gaffney BL, Goswami B, Jones RA. 1993. Nitrogen-15 labeled oligodeoxy-nucleotides. 7. Use of ^{15}N NMR of probe H-bonding in an $\text{O}^6\text{MeG}\cdot\text{C}$ base pair. *J Amer Chem Soc* 115:12607–12608.
- Gao X, Jones RA. 1987. Nitrogen-15 labeled oligodeoxynucleotides. Characterization by ^{15}N NMR of d[CGTACG] containing $^{15}\text{N}^6$ - or $^{15}\text{N}^1$ -labeled deoxyadenosine. *J Amer Chem Soc* 109:3169–3171.
- Gluck TC, Wills NM, Gesteland R, Draper DE. 1997. Folding of an mRNA pseudoknot required for stop codon readthrough: Effects of mono- and divalent cations on stability. *Biochemistry* 36:16173–16186.
- Hingerty BE, Brown RS, Jack A. 1978. Further refinement of the structure of yeast tRNA^{Phe}. *J Mol Biol* 124:523–534.
- Hoffman DW, Holland JA. 1995. Preparation of carbon-13 labeled ribonucleotides using acetate as an isotope source. *Nucleic Acids Res* 23:3361–3362.
- Hoffman DW, White SW. 1995. The RNA stem-loop of the trans-activation response (TAR) element of equine infectious anemia virus. *Nucleic Acids Res* 23:4058–4065.
- James TL, James JL, Lapidot A. 1981. Structure and dynamic information about double-stranded DNA from nitrogen-15 NMR spectroscopy. *J Amer Chem Soc* 103:6748–6750.
- Jiang F, Kumar RA, Jones RA, Patel DJ. 1996. Structural basis of RNA folding and recognition in an AMP-RNA aptamer complex. *Nature* 382:183–186.
- Kay LE, Ikura M, Bax A. 1990. Proton-proton correlation via carbon-carbon couplings: A three-dimensional NMR approach for the assignment of aliphatic resonances in proteins labeled with carbon-13. *J Amer Chem Soc* 112:888–889.
- Kolk MH, van der Graaf M, Wijmenga SS, Pleij CWA, Heus HA, Hilbers CW. 1998. NMR structure of a classical pseudoknot: Interplay of single- and double-stranded RNA. *Science* 280:433–438.
- Marion D, Driscoll PC, Kay LE, Wingfield PT, Bax A, Gronenborn AM, Clore GM. 1989. Overcoming the overlap problem in the assignment of ^1H NMR spectra of larger proteins by the use of three-dimensional heteronuclear ^1H - ^{15}N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: Application to interleukin-1 beta. *Biochemistry* 28:6150–6156.
- Markowski V, Sullivan GR, Roberts JD. 1977. Nitrogen-15 nuclear magnetic resonance spectroscopy of some nucleosides and nucleotides. *J Amer Chem Soc* 99:713–718.
- McPheeters DS, Stormo GD, Gold L. 1988. Autogenous regulatory site on the bacteriophage T4 gene 32 messenger RNA. *J Mol Biol* 201:517–535.
- Michnicka MJ, Harper JW, King GC. 1993. Selective isotopic enrichment of synthetic RNA: Application to the HIV-1 TAR element. *Biochemistry* 32:395–400.
- Nikonowicz EP, Sirt A, Legault P, Jucker FM, Baer LM, Pardi A. 1992. Preparation of ^{13}C and ^{15}N labelled RNAs for heteronuclear multidimensional NMR studies. *Nucleic Acids Res* 20:4507–4513.
- Pardi A. 1995. Multidimensional heteronuclear NMR experiments for structure determination of isotopically labeled RNA. *Methods Enzymol* 261:350–80.
- Philippe C, Bénard L, Portier C, Westhof E, Ehresmann B, Ehresmann C. 1995. Molecular dissection of the pseudoknot governing the translational regulation of *E. coli* ribosomal protein S15. *Nucleic Acids Res* 23:18–28.
- Plateau P, Guéron M. 1982. Exchangeable proton NMR without line distortion. *J Am Chem Soc* 104:7310–7311.

- Pleij CWA. 1994. RNA pseudoknots. *Curr Opin Struc Biol* 4:337–344.
- Pleij CWA, Rietveld K, Bosch L. 1985. A new principle of RNA folding based on pseudoknotting. *Nucleic Acids Res* 13:1717–1731.
- Pley HW, Flaherty KM, McKay DB. 1994. Three dimensional structure of a hammerhead ribozyme. *Nature* 372:68–74.
- Portier C, Philippe C, Dondon L, Grunberg-Manago M, Ebel JP, Ehresmann B, Ehresmann C. 1990. Translational control of ribosomal protein S15. *Biochim Biophys Acta* 1050:328–336.
- Puglisi JD, Wyatt JR, Tinoco I. 1990. Conformation of an RNA pseudoknot. *J Mol Biol* 214:437–453.
- Qiu H, Kaluarachchi K, Du Z, Hoffman DW, Giedroc DP. 1996. Thermodynamics of folding of the RNA pseudoknot of the T4 gene 32 autoregulatory mRNA. *Biochemistry* 35:4175–4186.
- Rietveld K, Pleij CWA, Bosch L. 1983. Three-dimensional models of the tRNA-like 3' termini of some plant viral RNAs. *EMBO J* 2:1079–1085.
- Saenger W. 1984. *Principles of nucleic acid structure*. New York: Springer-Verlag.
- Santoro J, King G. 1992. A constant-time 2D Overbundenhausen experiment for inverse correlation of isotopically enriched species. *J Mag Reson* 97:202–207.
- Shen LX, Tinoco I. 1995. The structure of an RNA pseudoknot that causes efficient frameshifting in mouse mammary tumor virus. *J Mol Biol* 247:963–978.
- Sklenar V, Peterson RD, Rejante MR, Feigon J. 1994. Correlation of nucleotide base and sugar protons in a ^{15}N -labeled HIV-1 RNA oligonucleotide by ^1H - ^{15}N HSQC experiments. *J Biomol NMR* 4:117–122.
- Sussman JL, Holbrook SR, Warrant RW, Church GM, Kim S-H. 1978. Restrained refinement of the structure of yeast tRNAPhe. *J Mol Biol* 123:607–630.
- Tang CK, Draper DE. 1989. Unusual mRNA pseudoknot structure is recognized by a protein translational repressor. *Cell* 57:531–536.
- ten Dam EB, Brierley I, Inglis S, Pleij CWA. 1994. Identification and analysis of the pseudoknot-containing *gag-pro* ribosomal frameshift signal of simian retrovirus-1. *Nucleic Acids Res* 22:2303–2310.
- ten Dam EB, Pleij CWA, Bosch L. 1990. RNA pseudoknots: Translational frameshifting and readthrough on viral RNAs. *Viral Genes* 4:121–136.
- ten Dam EB, Pleij CWA, Draper, D. 1992. Structural and functional aspects of RNA pseudoknots. *Biochemistry* 31:11665–11676.
- ten Dam EB, Verlaan PWG, Pleij CWA. 1995. Analysis of the role of the pseudoknot component in the SRV-1 gag-pol ribosomal frameshift signal: Loop lengths and stability of the stem regions. *RNA* 1:145–154.
- Theimer CA, Wang Y, Hoffman DW, Krisch HM, Giedroc DP. 1998. Non-nearest neighbor effects on the thermodynamics of unfolding of a model mRNA pseudoknot. *J Mol Biol* 279:545–564.
- Varani G, Tinoco I. 1991. Carbon assignments and heteronuclear coupling constants for an RNA oligonucleotide from natural abundance ^{13}C - ^1H correlated experiments. *J Amer Chem Soc* 113:9349–9352.
- Varani G, Aboul-ela F, Allain FH. 1996. NMR investigations of RNA structure. *Prog NMR Spectr* 29:51–127.
- Voet D, Rich A. 1970. The crystal structures of purines, pyrimidines, and their intermolecular complexes. *Prog Nucleic Acid Res Mol Biol* 10:183–265.
- Vuister GW, Boelens R, Kaptein R. 1988. Non-selective three-dimensional NMR spectroscopy. The 3D NOE-HOHAHA experiment. *J Mag Reson* 80:175–185.
- Wang C, Gao X, Jones RA. 1991. Nitrogen-15 labeled oligonucleotides. 2. Solvent isotope effects on the chemical shift of the adenine N1 in an A-T base pair. *J Amer Chem Soc* 113:1448–1450.
- Westhof E, Dumas P, Moras D. 1988. Restrained refinement of two crystalline forms of yeast aspartic acid and phenylalanine transfer RNA crystals. *Acta Crystallogr A* 44:112–123.
- Wijmenga SS, van Buuren BNM. 1998. The use of NMR methods for conformational studies of nucleic acids. *Prog NMR Spectr* 32:287–387.
- Wills NM, Gesteland RF, Atkins JF. 1991. Evidence that a downstream pseudoknot is required for translational read-through of the Moloney murine leukemia virus *gag* stop codon. *Proc Natl Acad Sci USA* 88:6991–6995.
- Wills NM, Gesteland RF, Atkins JF. 1994. Pseudoknot-dependent read-through of retroviral *gag* termination codons: Importance of sequences in the spacer and loop 2. *EMBO J* 13:4137–4144.
- Wishart DS, Bigam CG, Yao J, Abildgaard F, Dyson HJ, Oldfield E, Markley JL, Sykes BD. 1995. ^1H , ^{13}C , and ^{15}N chemical shift referencing in biomolecular NMR. *J Biomol NMR* 6:135–140.
- Zuiderweg ERP, Fesik SW. 1989. Heteronuclear three-dimensional NMR spectroscopy of the inflammatory protein C5a. *Biochemistry* 28:2387–2391.