RNA (2002), *8*:981–996. Cambridge University Press. Printed in the USA.

DOI: 10.1017.S1355838202024044

Comparative studies of frameshifting and nonframeshifting RNA pseudoknots: A mutational and NMR investigation of pseudoknots derived from the bacteriophage T2 gene 32 mRNA and the retroviral *gag-pro* frameshift site

YUE WANG,[1] NORMA M. WILLS,[2] ZHIHUA DU,[1] ANUPAMA RANGAN,[1] JOHN F. ATKINS,[2] RAYMOND F. GESTELAND,[2] and DAVID W. HOFFMAN[1]

[1]Department of Chemistry and Biochemistry, Institute for Cell and Molecular Biology, University of Texas at Austin, Austin, Texas 78712, USA
[2]Department of Human Genetics, University of Utah, Salt Lake City, Utah 84112-5330, USA

ABSTRACT

Mutational and NMR methods were used to investigate features of sequence, structure, and dynamics that are associated with the ability of a pseudoknot to stimulate a –1 frameshift. In vitro frameshift assays were performed on retroviral *gag-pro* frameshift-stimulating pseudoknots and their derivatives, a pseudoknot from the gene *32* mRNA of bacteriophage T2 that is not naturally associated with frameshifting, and hybrids of these pseudoknots. Results show that the *gag-pro* pseudoknot from human endogenous retrovirus-K10 (HERV) stimulates a –1 frameshift with an efficiency similar to that of the closely related retrovirus MMTV. The bacteriophage T2 mRNA pseudoknot was found to be a poor stimulator of frameshifting, supporting a hypothesis that the retroviral pseudoknots have distinctive properties that make them efficient frameshift stimulators. A hybrid, designed by combining features of the bacteriophage and retroviral pseudoknots, was found to stimulate frameshifting while retaining significant structural similarity to the nonframeshifting bacteriophage pseudoknot. Mutational analyses of the retroviral and hybrid pseudoknots were used to evaluate the effects of an unpaired (wedged) adenosine at the junction of the pseudoknot stems, changing the base pairs near the junction of the two stems, and changing the identity of the loop 2 nucleotide nearest the junction of the stems. Pseudoknots both with and without the wedged adenosine can stimulate frameshifting, though the identities of the nucleotides near the stem1/stem2 junction do influence efficiency. NMR data showed that the bacteriophage and hybrid pseudoknots are similar in their local structure at the junction of the stems, indicating that pseudoknots that are similar in this structural feature can differ radically in their ability to stimulate frameshifting. NMR methods were used to compare the internal motions of the bacteriophage T2 pseudoknot and representative frameshifting pseudoknots. The stems of the investigated pseudoknots are similarly well ordered on the time scales to which nitrogen-15 relaxation data are sensitive; however, solvent exchange rates for protons at the junction of the two stems of the nonframeshifting bacteriophage pseudoknot are significantly slower than the analogous protons in the representative frameshifting pseudoknots.

Keywords: frameshifting; NMR; pseudoknot; retrovirus

INTRODUCTION

In decoding mouse mammary tumor virus (MMTV) messenger RNA, 23% of the ribosomes translating the *gag* gene shift frame shortly before the stop codon at a specific site to enter the protease-encoding gene, *pro*. Near the end of the *pro* gene, 8% of these ribosomes again shift to the −1 frame to enter the *pol* gene to synthesize a fusion protein that includes reverse transcriptase (Jacks et al., 1987; Moore et al., 1987). There is no independent ribosome entry to either the *pro* or *pol* genes; instead their translation is solely mediated by ribosomes that initiate translation at the start of the *gag* gene. The framing of the *gag*, *pro*, and *pol* genes of human endogenous retrovirus-K10 (HERV; Ono et al., 1986) is the same as that of MMTV. In some retro-

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.

viruses, such as murine leukemia virus (MuLV), *gag* and *pol* are in the same reading frame and readthrough of the UAG *gag* termination codon (decoded as glutamine) is required for synthesis of the gag-pol polyprotein (Yoshinaka et al., 1985).

A pseudoknot located several (6 to 8) nucleotides 3′ to the heptanucleotide frameshift site is required for programmed −1 frameshifting in infectious bronchitis virus (IBV), mouse mammary tumor virus (MMTV), simian retrovirus (SRV), and beet western yellows virus (BWYV), where the efficiency of −1 frameshifting in these RNA viruses is ~25%, ~23%, 23%, and 1%, respectively (Jacks et al., 1987; Brierley et al., 1989; Garcia et al., 1993; ten Dam et al., 1994). Pseudoknots are known to cause translating ribosomes to pause; however, the pause alone is not sufficient to cause frameshifting (Tu et al., 1992; Somogyi et al., 1993; Lopinski et al., 2000; Kontos et al., 2001). Other features of frameshift-stimulatory pseudoknots must act directly or via an accessory factor(s) to promote efficient frameshifting. One such feature, a bend of approximately 68 degrees between the two helices, has been implicated in a derivative of the MMTV *gag-pro* pseudoknot (Shen & Tinoco, 1995; Chen et al., 1996). The bend is a consequence of an unpaired adenosine nucleotide between the two stems and constraints imposed by the stacking of nucleotides in loop 2. Another derivative of the MMTV pseudoknot that does not promote frameshifting also contains a bend between the helices (Kang et al., 1996), although it is in a different orientation. Mutational studies of pseudoknots derived from an infectious bronchitis virus (IBV) mRNA −1 frameshift site have shown that introducing an unpaired adenosine at the junction of stacked helical stems does not by itself create a frameshifting pseudoknot, and have suggested the importance of specific interactions between stem 1 and loop 2 (Liphardt et al., 1999). A crystallographic structure of the frameshift-promoting pseudoknot in BWYV revealed rotation (48 degrees) and bending of the stems (25 degrees) at the helical junction (Su et al., 1999). Additional features that may contribute to frameshifting activity in the BWYV pseudoknot include a quadruple base interaction between a nucleotide in loop 1 and the stem regions, noncanonical interactions between nucleotides in loop 2 and stem 1 (Su et al., 1999) and pH-dependent loop 1 to stem 2 tertiary interactions (Nixon & Giedroc, 2000).

A recent NMR study of a frameshift-stimulating pseudoknot derived from the SRV *gag-pro* site (Michiels et al., 2001) revealed a structure that in some respects is reminiscent of a pseudoknot from the gene *32* mRNA of bacteriophage T2 (Du et al., 1996; Holland et al., 1999). which is not believed to be associated with frameshifting. The SRV and bacteriophage T2 pseudoknots each contain two stems that are stacked; in each case bending at the junction of the stems was determined to be modest, if present. In the SRV pseudoknot, the stacked base pairs at the junction of the two stems have a helical twist of 49 degrees (Michiels et al., 2001), compared to 32.7 degrees in an ideal A-helix; in the bacteriophage T2 pseudoknot this same helical twist was found to be 50 ± 6 degrees (Holland et al., 1999). In each pseudoknot, the bases of the loop 2 nucleotides are stacked in the minor groove of stem 1, and the single nucleotide of loop 1 is embedded in the major groove of stem 2 with no apparent tertiary interactions with the stem 2 nucleotides. The apparent structural similarities between the SRV and bacteriophage T2 pseudoknots raise some important questions: Could the bacteriophage T2 pseudoknot serve as a stimulator of frameshifting? And, if not, what are the features that distinguish the frameshifting from nonframeshifting pseudoknots?

Identifying the physical features of pseudoknots that are important in stimulating frameshifting is of significant interest (reviewed by Giedroc et al., 2000), and requires information from other active and inactive structures. In the present study, a combination of mutational and NMR methods were used to further investigate the relationships between primary sequence, structure, dynamics, and frameshifting activity, using a set of pseudoknots with sequences derived from the *gag-pro* −1 frameshift site of HERV, the bacteriophage T2 gene *32* mRNA pseudoknot, and a hybrid of these pseudoknots. These results, when combined with previously reported functional assays of frameshift-stimulating pseudoknots derived from other viral mRNAs (such as MMTV, SRV, and IBV). provide additional insights into the requirements for an efficient −1 frameshift signal.

RESULTS

Mutational analyses of pseudoknots and –1 frameshifting

The HERV-K10 gag-pro frameshifting pseudoknot

HERV is closely related to the SRV and MMTV retroviruses, and encodes a full-length Gag protein and a functional protease, with the protease gene occurring in the −1 reading frame relative to the Gag protein (Schommer et al., 1996). In the present work, the HERV *gag-pro* slippery sequence and pseudoknot (Fig. 1) were tested for their ability to stimulate −1 frameshifting by inserting these sequences between two open reading frames (ORFs). Glutathione-S-transferase (GST) is encoded by the upstream ORF, and the downstream region encodes rabbit β-globin in the −1 reading frame relative to the GST (Matsufuji et al., 1995), Production of the GST-β-globin fusion protein is therefore dependent on −1 frameshifting (Fig. 2). In this system, the HERV slippery sequence and pseudoknot were found to stimulate −1 frameshifting with an efficiency of

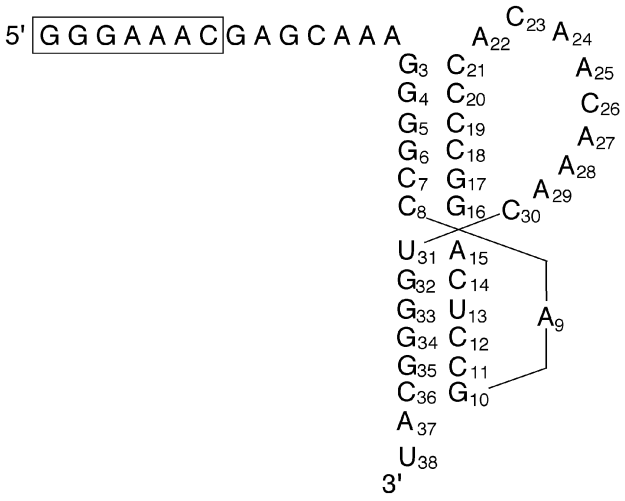

FIGURE 1. Schematic diagram of the slippery sequence and pseudoknot at the *gag-pro* frameshift site of the human endogenous retrovirus K-10 (HERV). The nucleotides within the pseudoknot region are numbered to be consistent with the NMR spectra.

18.5% ± 0.6% (Table 1; Fig. 2). The background rate for the slippery sequence alone is 3% (Table 1; Fig. 2), similar to the background level of frameshifting on the slippery heptanucleotide sequence reported for IBV and MMTV (Brierley et al., 1989; Chamorro et al., 1992). In the present study, when an in vivo dual-luciferase reporting system was used (Grentzmann et al., 1998), the −1 frameshifting efficiency of the HERV slippery site and pseudoknot was found to be 22.6% ± 1.5%, with a background rate for the slippery sequence alone of 1.1% ± 0.1%. We conclude that the HERV slippery sequence and pseudoknot are able to stimulate −1 frameshifting with an efficiency similar to that of other closely related retroviruses, including SRV-1 (ten Dam et al., 1994, 1995) and MMTV (Chen et al., 1995).

The effect of mutations of some of the loop 2 nucleotides on the frameshifting ability of the HERV pseudoknot was investigated (Table 1; Fig. 2). All five of the loop 2 mutants investigated were able to stimulate −1 frameshifting; however, four of the five mutants were less efficient than the wild-type sequence. Deletions of 1, 2, and 3 nt from loop 2 reduced frameshifting efficiency to 9%, 6%, and 6%, respectively. This implies that either the identity of one or more of the deleted nucleotides is important, or that the length of loop 2 is critical, or both. This result is consistent with recent mutational studies of IBV (Liphardt et al., 1999) and SRV (ten Dam et al., 1995) that suggest the identities of at least some of the loop 2 nucleotides are important for frameshifting activity.

Liphardt et al. (1999) identified a requirement for adenosine as the end nucleotide of loop 2 in their mutational analysis of the IBV frameshifting pseudoknot. In the present study, we found that changing the analogous HERV loop 2 nucleotide (C30) to A resulted in a frameshifting activity of 14%, or 74% of wild-type activity (note that the nucleotide in the SRV-1 pseudoknot analogous to C30 is A). Changing HERV loop 2 nucleotide

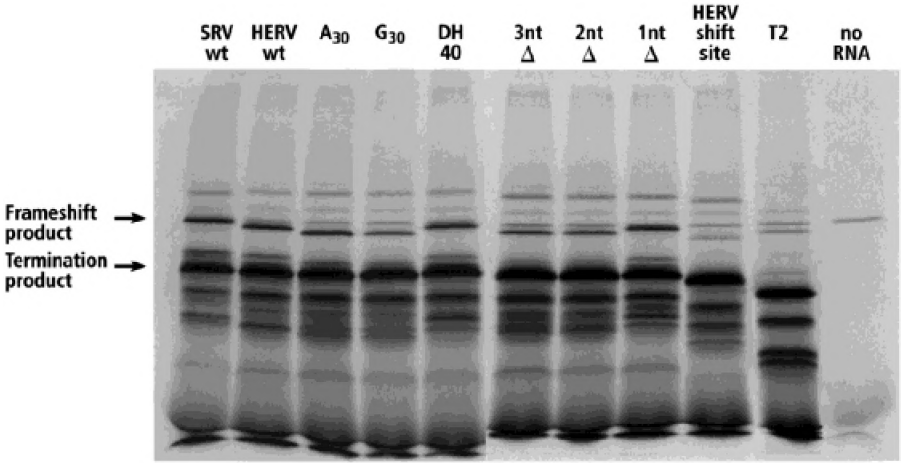

FIGURE 2. An SDS-polyacrylamide gel showing results of in vitro frameshift assays. Arrows indicate the positions of the termination and frameshift products. SRV wt contains the SRV slippery sequence and wild-type SRV pseudoknot; HERV wt contains the HERV slippery site and wild-type HERV pseudoknot; A30 contains a C-to-A change in loop 2 of the HERV pseudoknot; G30 contains a C-to-G change in loop 2 of the HERV pseudoknot; the hybrid pseudoknot contains the HERV slippery site and an HERV-T2 hybrid pseudoknot (see Fig. 4); 3ntΔ, 2ntΔ, and 1ntΔ lack 3, 2, and 1 nt in loop 2, respectively, of the HERV pseudoknot; HERV shift site contains the HERV slippery site but lacks the downstream pseudoknot; T2 contains the MMTV slippery sequence, A AAA AAC, and a 6-nt spacer followed by the wild-type bacteriophage T2 pseudoknot; no RNA shows the translation products in the reticulocyte lysate without added RNA. The value of the background band near the frameshift products was taken into account when determining frameshift efficiencies. The termination band in the T2 lane is shorter than in the other lanes because the ribosome encounters a 0-frame stop codon earlier within the T2 pseudoknot sequence.

TABLE 1. The table shows constructs assayed for in vitro frameshifting activity, performed using the GST-β-globin assay system.

Construct[a]	Frameshifting	Relative frameshifting
HERV wild type	19%	100%
Deletion 1 nt, A28, in loop2	9	47
Deletion 2 nt, A27 and A28, in loop2	6	32
Deletion 3 nt, C26, A27, and A28, in loop2	6	32
C30 → A	14	74
C30 → G	7	37
HERV without pseudoknot	3	16
MMTV with bacteriophage T2 pseudoknot	2	11
HERV with HERV-T2 hybrid pseudoknot	19	100

[a]HERV wild type: HERV slippery site, G GGA AAC, followed by the wild-type HERV pseudoknot; deletions in loop 2: HERV slippery site followed by the HERV pseudoknot lacking 1 nt, A28; 2 nt, A27 and A28; or 3 nt, C26, A27, and A28 in loop 2; C30 → A: HERV slippery site followed by HERV pseudoknot with A at the 3′ end of loop 2; C30 → G: HERV slippery site followed by HERV pseudoknot with G at the 3′ end of loop 2; HERV without pseudoknot: HERV slippery site without pseudoknot; MMTV with bacteriophage T2 pseudoknot: MMTV *gag-pro* slippery sequence, A AAA AAC, followed by 6-nt spacer and bacteriophage T2 pseudoknot; HERV with HERV-T2 hybrid pseudoknot: HERV slippery site followed by HERV 7-nt spacer and HERV-T2 hybrid pseudoknot (DH40) shown in Figure 3B.

The HERV wild type was used as the standard for determining relative frameshifting. Frameshift efficiencies are rounded to the nearest whole number and have an uncertainty of approximately 1%.

C30 to G resulted in a frameshifting activity of 7%, or 37% of wild-type activity (the nucleotide in the bacteriophage T2 pseudoknot analogous to C30 is G; Table 1; Fig. 2). Therefore, as was the case in IBV, the identity of the last nucleotide in loop 2 can have a significant effect on the frameshift efficiency of the HERV pseudoknot.

Frameshifting ability of the bacteriophage T2 gene 32 mRNA pseudoknot

The gene *32* mRNA pseudoknot from bacteriophage T2 naturally functions as a regulatory and protein binding site and is not known to be associated with −1 frameshifting. Because it has been extensively investigated by NMR methods (Du et al., 1996; Holland et al., 1999) and has some significant structural similarities to a known frameshifting pseudoknot (Michiels et al., 2001), the bacteriophage pseudoknot was tested for its ability to function as a stimulator of −1 frameshifting activity when located downstream of a retroviral slippery sequence. When the bacteriophage T2 pseudoknot followed the MMTV slippery sequence, A AAA AAC, by a 6-bt spacer, the frameshifting efficiency was 2%. The wild-type spacing of 7 nt between slippery sequence and pseudoknot was not tested, because this would have introduced stop codons in the −1 reading frame of the bacteriophage T2 pseudoknot sequence. For comparison, in the case of MMTV slippery sequence and the MMTV pseudoknot, decreasing the spacer from 7 to 6 nt reduced frameshifting from 23% (wild-type) to 9% (data not shown). Although the suboptimal spacing may therefore have prevented the bacteriophage T2 pseudoknot from exerting its full effect, the bacteriophage T2 pseudoknot does not appear to be an efficient stimulator of −1 frameshifting. This result supports a hypothesis that the retroviral pseudoknots have properties that make them particularly efficient stimulators of −1 frameshifting.

Frameshifting ability of a hybrid mRNA pseudoknot

One approach to identifying pseudoknot features important for promoting −1 frameshifting is to make hybrid pseudoknots between active and inactive structures. Inspection of the *gag-pro* frameshift site of several retrovirus mRNAs reveals that stem 1 of the pseudoknot is relatively well conserved, with usually at least three G-C base pairs at the 5′ end, in contrast to stem 2, which is more variable in sequence, although its length is usually 6 or 7 bp. The length and sequence of loop 2 varies widely among the retrovirus species. These observations suggest that the sequence of stem 1 is an important determinant of frameshift-stimulating ability. To test this hypothesis, we attempted to convert the bacteriophage T2 gene *32* mRNA pseudoknot to a frameshift-active form by giving it a retrovirus-like stem 1 sequence. This hybrid pseudoknot (construct "DH40," Fig. 3B) contains the natural sequence of stem 1 of the SRV-1 and HERV pseudoknots, whereas the nucleotides near the junction of the helical stems are the same as in the bacteriophage T2 gene *32* mRNA pseudoknot. Interestingly. this hybrid pseudoknot was found to stimulate frameshifting as well as the wild-type HERV pseudoknot (Table 1), a result consistent with stem 1 being an important determinant of frameshifting activity.

Frameshifting ability of mutant pseudoknots derived from the hybrid

Eleven additional pseudoknots derived from the HERV-bacteriophage T2 hybrid were tested for frameshifting activity, with sequences designed to address the effects of changing the identity of the loop 2 nucleotide closest to the junction of the stems from G to A (thus making the hybrid more like SRV, and less like the bacteriophage T2 sequence), inserting an unpaired "wedged" adenosine at the junction of the two stems (as occurs in the MMTV *gag-pro* pseudoknot), and changing the stem 2 base pair closest to the junction of the stems from U-A to G-C or G-U; these changes were evaluated individually and in various combinations (Table 2), All 11 of these mutant pseudoknots were reasonably active as stimulators of frameshifting, with efficiencies ranging between 15% and 24%, as deter-

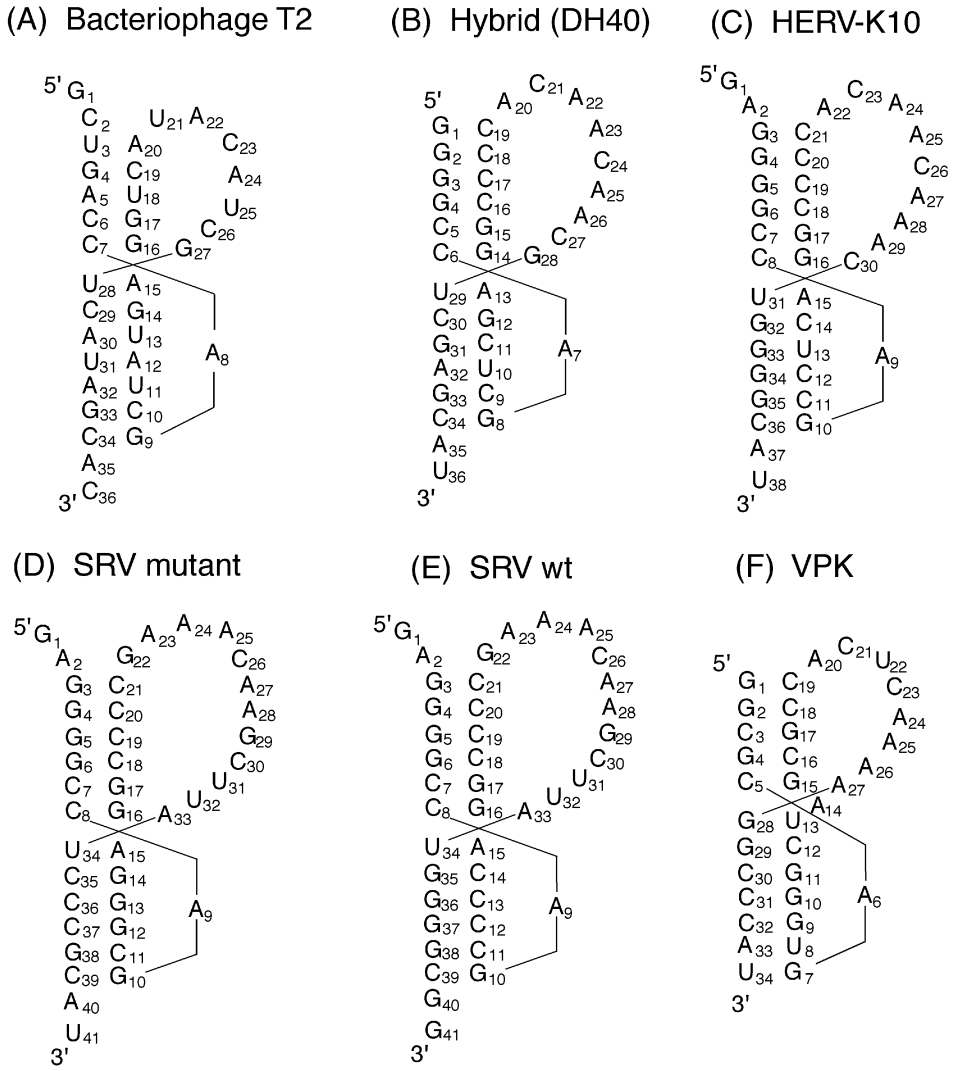

FIGURE 3. Schematic diagrams of the RNA pseudoknots discussed in this work. **A:** The gene *32* mRNA pseudoknot of bacteriophage T2. **B:** A designed hybrid frameshifting pseudoknot, construct "DH40." **C:** The pseudoknot downstream from the *gag-pro* frameshift site in the human endogenous retrovirus K-10 (HERV). **D:** A functional mutant of the pseudoknot downstream from the SRV *gag-pro* frameshift site. **E:** The pseudoknot downstream from the *gag-pro* frameshift site in the simian retrovirus-1 (SRV). **F:** VPK, a functional mutant derived from the pseudoknot downstream from the MMTV *gag-pro* frameshift site. The diagrams are drawn in such a way as to show the base pairing that occurs within the pseudoknots.

mined in our in vitro assays. Changing the 3′ nucleotide of loop 2 from G to A (construct 53-2) resulted in little change in frameshifting activity relative to the reference hybrid (Table 2). Inserting an adenosine wedge into the junction of the two stems (construct 59-7, Table 2) resulted in a decrease in frameshift efficiency from 19% to 15%. The most significant increases in frameshifting activity were seen in constructs where the stability of the stem 1/stem 2 interface was affected by replacing U-A with G-C. Without a wedged adenosine, frameshifting increases to 23% (construct 57-5), with a wedged adenosine (construct 63-3), the activity, 20%, is similar to the hybrid. When there is a G-U base pair in stem 2, the frameshifting efficiency is comparable to the parent hybrid when there is an adenosine at the 3′ end of loop 2 (construct 55-2). Frameshifting is also increased to 24% by having a G-C base pair in stem 2, an adenosine at the 3′ end of loop 2, and a wedged adenosine at the junction of the stems (construct 61-2); interestingly, this arrangement is predicted to most resemble the junction of the two stems in the MMTV pseudoknot. It is noted, however, that a G-U base pair in stem 2 will also suffice (construct 61-11) in the same context as construct 61-2. We conclude that strengthening the junction of the stems by increasing the stability of stem 2 can increase frameshifting. The presence of an adenosine wedge can increase frameshifting in a favorable environment of a strengthened stem 2 and an adenosine at the 3′ end of loop 2. The results reinforce the significance of the junction of the stems for frameshifting. Interestingly, of all the pseudoknots evaluated for frameshifting ability, the only pseudo-

TABLE 2. Results of in vitro frameshift assays for a series of pseudoknots derived from the HERV-bacteriophage T2 hybrid pseudoknot (sequence DH40, Figure 3B) performed using the GST-β-globin assay system.[a]

Construct	3′ end loop 2	Base stem 2	Wedged base	Frameshifting efficiency	Relative efficiency
HERV-T2 hybrid	G	U-A		19%	100%
53-2	**A**	U-A		20	105
55-1	**A**	**G-C**		22	116
55-2	**A**	**G-U**		20	105
57-5	G	**G-C**		23	121
57-4	G	**G-U**		18	95
59-7	G	U-A	**A**	15	79
59-2	**A**	U-A	**A**	16	84
61-2	**A**	**G-C**	**A**	24	126
61-11	**A**	**G-U**	**A**	24	126
63-3	G	**G-C**	**A**	20	105
63-5	G	**G-U**	**A**	16	84

[a]These constructs were designed to further address the effect of changes at the junction of the pseudoknot stems, including the effect of inserting a wedge base into the hybrid pseudoknot structure. Relative frameshift efficiencies are tabulated as a percentage of the frameshifting efficiency of the hybrid pseudoknot. Frameshift efficiencies are rounded to the nearest whole number and have an uncertainty of approximately 1%. Sequence differences from the hybrid pseudoknot are highlighted by bold letters.

knot found to essentially fail as a frameshift stimulator was the natural bacteriophage T2 sequence.

Comparative NMR studies of frameshifting, nonframeshifting, and hybrid pseudoknots

Structural features of the bacteriophage T2 and SRV mutant pseudoknots were previously investigated using homonuclear and heteronuclear NMR methods (Du et al., 1997; Holland et al., 1999; Michiels et al., 2001). In the present study, comparative NMR investigations of the HERV and hybrid frameshifting pseudoknots are added, and in addition, the dynamic properties of the bacteriophage T2 and representative frameshifting pseudoknots are investigated using NMR methods and compared. We note that the frameshifting pseudoknots are often found to be less amenable to NMR analysis than the bacteriophage pseudoknot, primarily due to chemical shift degeneracy as a result of the repetitive G-C rich stem sequences; Michiels et al. (2001) have also recently commented on this aspect of the frameshifting pseudoknot NMR spectra. The preparative yields of the bacteriophage T2 and SRV pseudoknots (with sequences shown in Fig. 3) were uniformly excellent in transcription reactions; however the yields for the HERV and hybrid pseudoknots were significantly lower, limiting our NMR studies to ^{1}H and ^{15}N spectra in these cases. Nevertheless, proton and ^{15}N NMR data were obtained that are suitable for confirming the base pairings schemes for the pseudoknots, and providing significant insight into their structural and dynamic features.

NMR investigation of the hybrid RNA pseudoknot

The hybrid frameshifting pseudoknot (Fig. 3B) was designed so that its sequence is the same as the bacteriophage T2 pseudoknot near the junction of the two stems and near the 3′ end of stem 2, while retaining sequence similarity to the HERV and SRV pseudoknots in other regions of the molecule. The resonance assignment problem for the stem resonances in the hybrid was greatly simplified by making comparisons with the previously assigned spectra of the bacteriophage T2 and SRV mutant pseudoknots, as similar chemical shifts and NOE intensities were observed in regions where the sequences are the same. Imino proton resonances and NOE cross peaks characteristic of Watson–Crick base pairs were observed for all base pairs of the hybrid pseudoknot (Figs. 4 and 5) except for G1-C19 and A13-U29. The nonobservation of the G1 and U29 imino resonances does not necessarily indicate the absence of the base pairs, as it is not unusual for imino protons at the ends of RNA helices to exchange rapidly enough with the solvent so as to be unobservable by NMR. NOEs involving nonexchangeable protons are a more reliable indicator of base pairing, and in the case of the hybrid pseudoknot, 27 NOE cross peaks involving the protons of the four nucleotides (C6, A13, G14, and U29) at the junction of the helical stems were observed, in all cases with intensities and chemical shifts similar to those of the analogous nucleotides (C7, A15, G16, and U28) of the bacteriophage pseudoknot, where the imino resonances characteristic of Watson–Crick base pairs are detected. For example, chemical shifts of the H2, H5, H6, H8, H1′ protons of the four nucleotides at the junction of the stems of the hybrid pseudoknot differ from those of the analogous nucleotides in the bacteriophage T2 pseudoknot by an average of only 0.06 ppm. Of particular significance, the NOE cross peaks involving the H2 proton of nucleotide A13 (located at the center of the junction of the two stems) and the H1′ protons of G14 (in stem 1) and C30 (in stem 2) in the hybrid are similar in intensity and chemical shift to the NOE cross peaks arising from the analogous nucleotides of the bacteriophage T2 pseudoknot. The observed NOEs involving protons at the junction of the stems support a pseudoknot structural model, rather than a stem-loop or mixture of stem-loop structures. The presence of the G8-C34 Watson–Crick base pair within the hybrid pseudoknot, located at the 3′ end of stem 2, is significant in that it enforces the requirement that loop 1 consists of a single nucleotide spanning the major groove of stem 2. The chemical shifts and NOE cross peaks involving the protons of nucleotides G8, C34, and A35 near the 3′ end of the hybrid frameshifting pseudoknot are nearly identical to the corresponding chemical shifts and NOE cross peaks

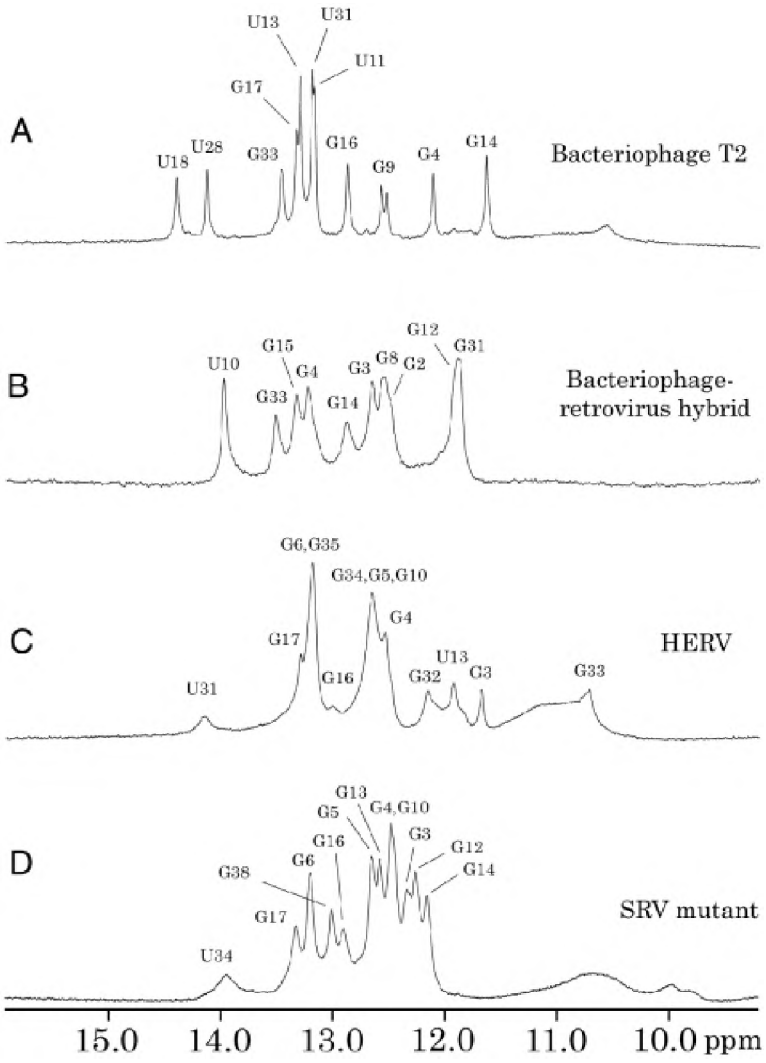

FIGURE 4. Sections of the one-dimensional proton NMR spectra of the RNA pseudoknots obtained in 90% H_2O/10% D_2O solvent at 20 °C. Assigned imino resonances are labeled. Shown are spectra of: the imino proton spectrum of the gene *32* mRNA pseudoknot of bacteriophage T2, with sequence shown in Figure 3A (**A**); the hybrid frameshifting pseudoknot with designed sequence shown in Figure 3B (**B**); the imino proton spectrum of the 38 nt pseudoknot from human endogenous retrovirus-K10 (HERV) *gag-pro* frameshift site, with sequence shown in Figure 3C (**C**); the imino proton spectrum of a functional mutant of the SRV-1 pseudoknot, with sequence shown in Figure 3D (**D**).

in the analogous nucleotides (G9, C34, and A35) in the bacteriophage T2 pseudoknot, indicating that the two pseudoknots are structurally similar in the vicinity of the terminal base pair of stem 2 as well; the chemical shifts of the corresponding H2, H5, H6, H8, imino, and amino protons of these 3 nt differ by an average of only 0.03 ppm.

The local structure at the 4 nt (C6, A13, G14, and U29) at the junction of the stems of the hybrid pseudoknot can be defined using distances derived from 27 NOE cross peaks identified in the homonuclear NMR data. Although this is somewhat fewer constraints than were obtained for the analogous region of the bacteriophage T2 pseudoknot (where 44 NOE cross peaks were identified in homonuclear and heteronuclear NMR data), the local structure at the C6-G14 and U29-A13 base pairs in the hybrid was found to be quite well defined (Fig. 6), with the root mean square deviation for the four junction nucleotides being 0.8 Å for a set of 20 accepted structures. In the hybrid pseudoknot, the helical twist between the C6-G14 and A13-U29 base pairs is approximately 50 degrees, or 18 degrees greater than the helical twist in regular A-form helical RNA. This overrotation is indistinguishable to that observed in the bacteriophage T2 (Holland et al., 1999) and SRV (Michiels et al., 2001) pseudoknots, and serves to relieve what would otherwise be an unacceptably close contact between the phosphates of A7 and U29 at the junction of the two stems, while preserving much of the stabilizing effects of base stacking. The ring planes of the bases of C6-G14 and U29-A13 in the hybrid are nearly parallel. Structures with bend angles within the range of 0 to 28 degrees were found to satisfy the distance bounds for the four junction nucleotides in the hybrid; this range of bend angles is not significantly different than that which is reported for the bacteriophage T2 (Holland et al., 1999) and SRV (Michiels et al., 2001) pseudoknots.

We conclude that in terms of several important structural features, the hybrid, bacteriophage T2, and SRV pseudoknots are similar; these features include the local structure at the junction of stems 1 and 2, loop 1

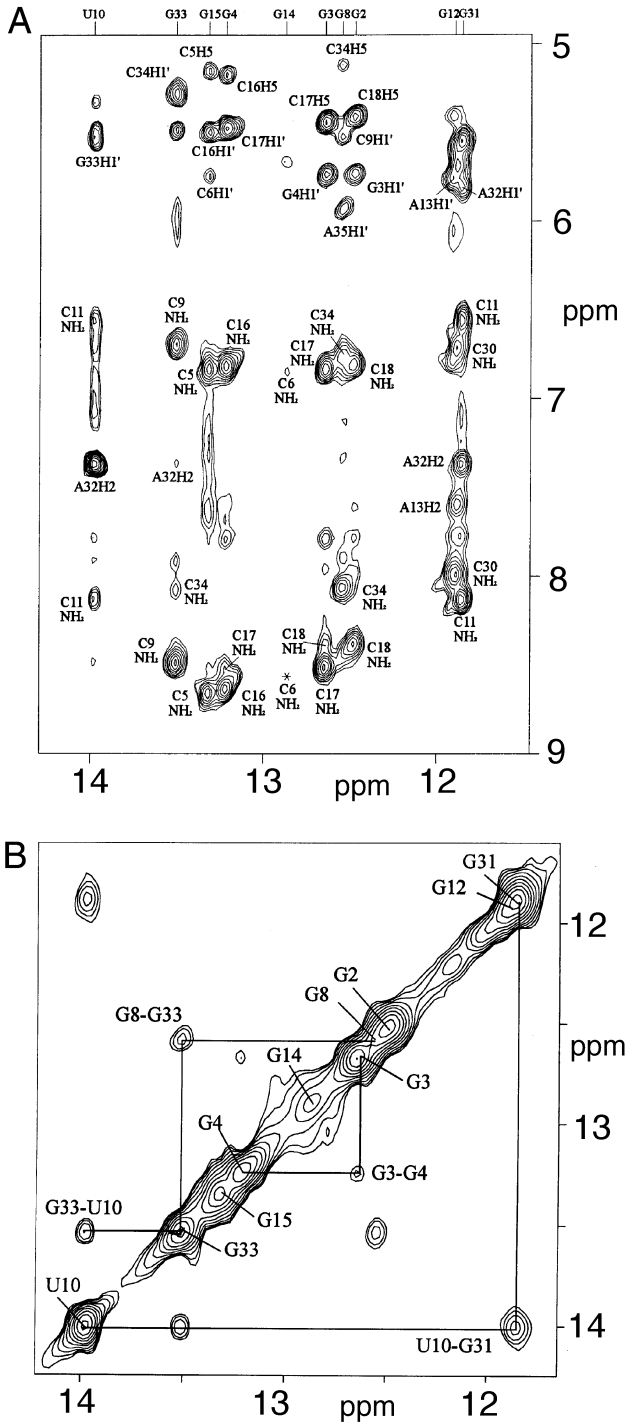

FIGURE 5. **A:** A section of the 500-MHz 2-D NOESY spectrum of the hybrid (DH40) frameshifting pseudoknot with sequence shown in Figure 3B, obtained in 90% H_2O/10% D_2O solvent, 10 mM phosphate buffer, pH 6.8, at 20 °C. The NOE mixing time was 200 ms. Cross peaks are observed between the imino protons of the adjacent base pairs in the helical stems. **B:** A section of the same spectrum showing NOE cross peaks involving the imino resonances. Cross peaks are labeled by residue type and number.

consisting of a single nucleotide, overrotation of the base pairs at the junction of the stems, little or no bending between the stems, and the local structure near the 3′ end of stem 2.

Confirmation of base pairing in the HERV pseudoknot

Base pairings for the HERV pseudoknot were confirmed using two-dimensional proton and ^{15}N-^{1}H correlated and NMR spectra, obtained for a 38-nt version of the molecule with sequence shown in Figure 3C, as well as a 36-nt version prepared without the unpaired nucleotides G1 and A2. Again, the assignment of the stem resonances was simplified by the similarity in chemical shifts among the protons of the HERV, SRV mutant, and hybrid pseudoknots in regions where the sequences are the same, specifically, in stem 1, at the junction of stem 1 and stem 2, and at the end of stem 2 nearest the 3′ end of the molecules. In both the 36- and 38-nt versions of the HERV pseudoknot, the wobble-type G33-U13 base pair in stem 2 showed the expected strong NOE cross peak between the G and U imino protons, as well as ^{15}N chemical shifts characteristic of G and U imino nitrogens, near 145 ppm and 160 ppm, respectively. The resonances of the HERV pseudoknot are broader than those of the other pseudoknots (Fig. 4), perhaps an indication of a tendency toward aggregation, or an indication of intermediate exchange on the NMR time scale. As was the case in the hybrid and SRV and mutant pseudoknots (Du et al., 1997), the imino proton resonances of the base pairs at the junction of the HERV pseudoknot stems are very broad, which is attributed to rapid exchange with the solvent (Fig. 4).

Comparison of the dynamics of frameshifting and nonframeshifting pseudoknots

Pseudoknots with similar structural features may differ significantly in terms of their flexibility and dynamics. One of the strengths of NMR spectroscopy is its ability to provide information regarding the motions of biomolecules. NMR-observable parameters related to molecular motion in RNA include ^{15}N relaxation parameters (R_1, R_2, and ^{15}N-^{1}H heteronuclear NOE) and imino proton exchange rates. The possibility that frameshifting and nonframeshifting pseudoknots may differ in their dynamic properties was investigated.

^{15}N relaxation analysis of the RNA pseudoknots

Representative RNA pseudoknots selected for ^{15}N relaxation analysis are (1) the nonframeshifting bacteriophage T2 pseudoknot (Fig. 3A), (2) a frameshifting pseudoknot derived from the SRV *gag-pro* site (Fig. 3D) that has been shown to be an efficient stimulator of −1 frameshifting (ten Dam et al., 1995) and does not contain a wedged base at the junction of its two stems (Du et al., 1997), and (3) a frameshifting pseudoknot derived from the MMTV *gag-pro* site (Fig. 3F) that does

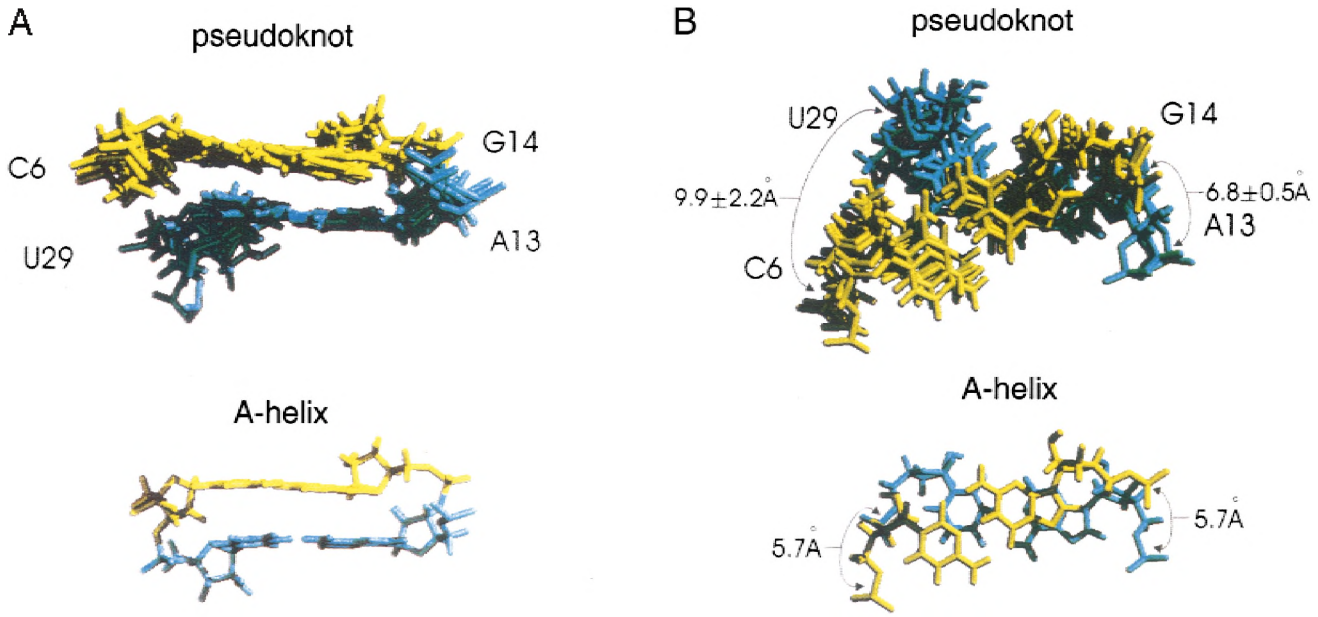

FIGURE 6. Views of the local structure at the junction of the two helical stems of the DH40 hybrid frameshifting pseudoknot with sequence shown in Figure 3B, showing the nucleotides of the C6-G14 and A13-U29 base pairs. **A**: Six superimposed structures of the C6-G14 and A13-U29 base pairs, representative of the full range of structures that are consistent with 27 interproton distances derived from NOE cross peaks and have reasonable molecular geometry. For comparison, the C6-G14 and A13-U29 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. **B**: A different view of the same 4 nt in the same six superimposed structures as shown on the left, but rotated by 90 degrees. Again, a view of the 4 nt as they would be in an ideal A-form helix is shown for comparison in the lower part of the figure. Distances between phosphorous atoms along the RNA backbone are indicated.

contain a wedged adenosine at the junction of its two stems (Shen & Tinoco, 1995), These three pseudoknots were selected because they represent the extremes in terms of frameshifting ability (or lack thereof), extremes of structural types (wedged adenosine or no wedge), and each has reasonably good dispersion in the chemical shifts of its imino resonances and sufficient preparative yield in in vitro transcription reactions so that isotope-enriched samples may be prepared.

For the bacteriophage pseudoknot, ^{15}N relaxation rate data were obtained for the 11 observable imino ^{15}N nuclei within the stems. Relaxation rates were successfully back-calculated using a relatively simple dynamic model, where the molecule was assumed to have an ellipsoidal shape, and only an order parameter (S^2) for each imino group, a ratio between the long and short axes of the diffusion tensor (D_{ratio}), and a global tumbling time (τ_m) were used as variables. Excellent agreement between the observed and calculated relaxation data (Table 3) was obtained using a ratio of 2.13:1.0:1.0 for the long and short axes of the diffusion tensor; this value for D_{ratio} is similar to a value of 2.0:1.0:1.0 calculated using the program DASHA (Orekhov et al., 1995) and the coordinate file "2tpk" from the Protein Data Bank to model the diffusion of the molecule. The R_1 and R_2 relaxation rates and ^{15}N-^{1}H NOEs for the 11 observed imino nitrogen nuclei indicate a global tumbling time (τ_m) of 10.4 ns, which is within the range typical of biomolecules with similar molecular weight. Order parameters for the imino groups are within the range of 0.72 to 0.83 (Table 3), values typical of a well-ordered structure.

Inspection of the data in Table 3 reveals that the R_1 and R_2 relaxation rates of imino ^{15}N nuclei are not the same for all base pairs within the pseudoknot stems. The differences in observed relaxation rates do not necessarily imply differences in internal motions, but can also be accounted for by (1) differences in the chemical shift anisotropy of the G versus U imino ^{15}N nuclei (Akke et al., 1997), and (2) the anisotropic tumbling of the pseudoknot, which causes the ^{15}N relaxation rates to exhibit a dependence on the orientation of each N-H bond vector relative to the major axis of inertia within the molecular structure (Lee et al., 1997; Luginbühl et al., 1997; Lillemoen & Hoffman, 1998). For example, N-H bond vectors that are perpendicular to the long axis of the molecule tumble (on average) faster than N-H vectors that are parallel to the major axis, as illustrated in Figure 7. ^{15}N relaxation rates therefore depend on the angle (theta) between each individual N-H vector and the long axis of the molecule, as well as the internal motions.

For the pseudoknot derived from the SRV *gag-pro* −1 frameshift site (Fig. 3D), a relatively simple dynamic model was sufficient to fit the observed and calculated ^{15}N relaxation data, where only an order

TABLE 3. NMR data related to dynamics of the bacteriophage T2 gene *32* mRNA pseudoknot.[a]

Nt	R_1 (s^{-1}) (obs)	R_1 (s^{-1}) (calc)	R_2 (s^{-1}) (obs)	R_2 (s^{-1}) (calc)	NOE	S^2	theta (deg)	K_{HX} (s^{-1})
Stem 1								
G4	1.50 ± 0.04	1.48	10.89 ± 0.3	11.19	0.72 ± 0.04	0.83	110.2	2.4 ± 1.0
U18	1.31 ± 0.04	1.33	10.12 ± 0.3	9.79	0.73 ± 0.04	0.77	76.2	4.2 ± 2.0
G17	1.38 ± 0.04	1.38	10.04 ± 0.3	10.08	0.71 ± 0.04	0.76	78.1	1.6 ± 1.0
G16	1.50 ± 0.04	1.52	11.15 ± 0.3	10.94	0.71 ± 0.04	0.83	83.4	2.1 ± 1.0
Stem 2								
U28	1.40 ± 0.04	1.34	9.19 ± 0.3	9.33	0.76 ± 0.04	0.74	94.1	2.7 ± 1.0
G14	1.39 ± 0.04	1.38	9.91 ± 0.3	9.95	0.71 ± 0.04	0.76	83.6	1.1 ± 0.6
U13	1.26 ± 0.04	1.26	9.21 ± 0.3	9.22	0.69 ± 0.04	0.73	77.2	1.0 ± 0.6
U31	1.25 ± 0.04	1.23	9.03 ± 0.3	9.19	0.76 ± 0.04	0.72	106.8	0.8 ± 0.5
U11	1.22 ± 0.04	1.21	9.37 ± 0.3	9.50	0.70 ± 0.04	0.73	64.5	0.8 ± 0.5
G33	1.31 ± 0.04	1.27	10.10 ± 0.3	10.43	0.72 ± 0.04	0.75	119.9	1.4 ± 1.0
G9	1.29 ± 0.04	1.39	11.50 ± 0.3	11.23	0.72 ± 0.04	0.81	61.5	1.7 ± 1.0

[a]Shown are the observed ^{15}N relaxation rates (R_1 and R_2) and rates calculated using an anisotropic model for the pseudoknot motion, ^{15}N-^1H NOEs, order parameters (S^2), and rates with which the imino protons exchange with solvent (K_{HX}), for guanosine N1 and uridine N3 imino groups. Theta refers to the angle between the imino N-H bond vector and the major axis of inertia of the pseudoknot.

parameter (S^2). a global tumbling time (τ_m) and an axial ratio for the diffusion tensor (D_{ratio}) of a prolate ellipsoid were used as variables. Order parameters for the imino ^{15}N nuclei are within the range of 0.77 to 0.89, values typical of a well-ordered structure. Excellent agreement between observed and calculated ^{15}N relaxation data (Table 4) was obtained using a ratio of 1.82:1.0 between the long and short axes of the diffusion tensor of the frameshifting pseudoknot, a ratio calculated using the program DASHA (Orekhov et al., 1995). ^{15}N relaxation data were not obtained for U34 because of the rapid exchange of the imino proton with solvent. The τ_m of 8.83 ns that best accounts for the ^{15}N relaxation data in the SRV frameshifting pseudoknot is less than the value of 10.4 ns obtained for the bacteriophage T2 pseudoknot. The relatively low value obtained for τ_m of the SRV pseudoknot may be due to a particularly compact loop 2 structure, resulting in a minimum surface area and hydrodynamic drag. Such a compact and ordered loop 2 was reported by Michiels et al. (2001) in their recent structural analysis of a similar frameshifting pseudoknot derived from the SRV sequence.

The VPK pseudoknot (Fig. 3F) is also an efficient stimulator of frameshifting (Shen & Tinoco, 1995), but differs from the SRV pseudoknot in that it contains an unpaired adenosine at the junction of its two stems; this wedged adenosine introduces the significant bend angle between the axes of the stems (Shen & Tinoco, 1995). ^{15}N relaxation data (R_1, R_2, and NOE) was obtained for six of the imino nitrogens; these are G1, G2, G4, G10, G11, and G17, all of which are part of G-C base pairs in the stems; other imino protons exchanged with the solvent too rapidly for relaxation rates to be measured. As was the case with the bacteriophage T2 and SRV pseudoknots, a simple motional model was

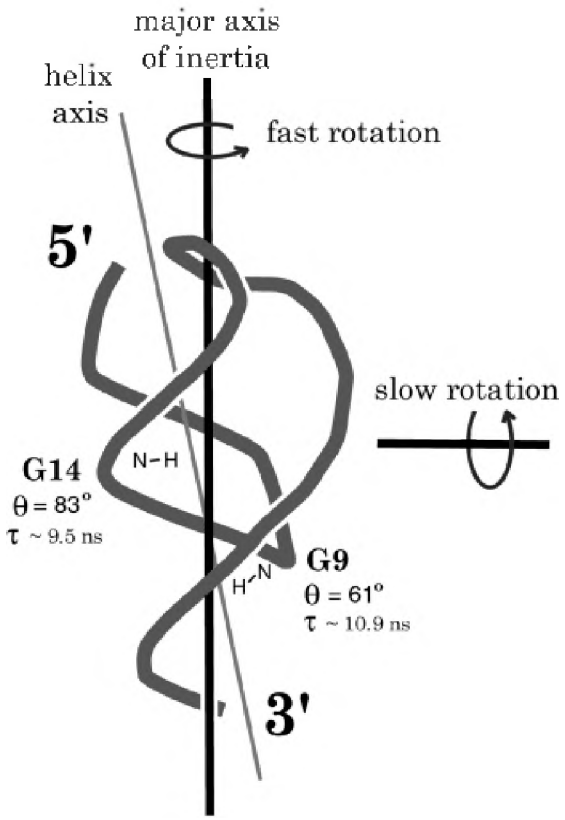

FIGURE 7. A representation of the bacteriophage T2 RNA pseudoknot, derived from Protein Data Bank entry 2tpk (Holland et al., 1999), indicating the direction of the major axis of inertia, the helix axes, and the orientation of the imino N-H bond vectors for nucleotides G9 and G14. The anisotropic rotation of the pseudoknot results in tumbling times that depend on the angle (theta) between the ^{15}N-^1H bond vector and the major axis of inertia. For example, the imino ^{15}N-^1H bond vector of nucleotide G14 is nearly perpendicular to the major axis of inertia, and has a shorter average tumbling than that of G9, accounting for the observed differences in R_1 and R_2 relaxation rates. The ratio R_2/R_1 for G14 is 7.2, consistent with a rotational correlation time of 9.5 ns; the ratio R_2/R_1 for G9 is 8.9, consistent with a rotational correlation time of 10.9 ns.

TABLE 4. NMR data related to dynamics of the SRV RNA pseudoknot, with sequence in Figure 3D.[a]

Nt	R_1 (s^{-1}) (obs)	R_1 (s^{-1}) (calc)	R_2 (s^{-1}) (obs)	R_2 (s^{-1}) (calc)	NOE	S^2	theta (deg)	K_{HX} (s^{-1})
Stem 1								
G3	1.71 ± 0.02	1.71	10.73 ± 0.4	10.69	0.76 ± 0.08	0.89	121.0	9 ± 2
G4	1.63 ± 0.01		9.77 ± 0.1		(overlap with G10)			
G5	1.66 ± 0.03		9.43 ± 0.2		(overlap with G13)			
G6	1.70 ± 0.03	1.58	8.60 ± 0.1	8.79	0.70 ± 0.04	0.77	93.4	1.7 ± 1.0
G17	1.62 ± 0.04	1.68	9.67 ± 0.2	9.38	0.71 ± 0.06	0.82	82.7	2.6 ± 1.0
G16	1.54 ± 0.06	1.60	9.34 ± 0.3	8.93	0.67 ± 0.12	0.78	82.9	50 ± 40
Stem 2								
U34								300 ± 150
G14	1.55 ± 0.02	1.59	9.67 ± 0.2	9.31	0.67 ± 0.06	0.80	68.8	1.1 ± 0.4
G13	(overlap with G5)							
G38	1.48 ± 0.03	1.50	9.75 ± 0.3	9.34	0.70 ± 0.08	0.78	120.5	2.3 ± 1.0
G12	1.56 ± 0.02	1.55	9.70 ± 0.2	9.80	0.72 ± 0.05	0.81	56.9	2.0 ± 1.0
G10	(overlap with G4)							

[a]Shown are the observed and calculated ^{15}N relaxation rates (R_1 and R_2), ^{15}N-^1H NOEs, order parameters (S^2), and estimated imino proton exchange rates (K_{HX}) for guanosine N1 and uridine N3 imino groups within base-paired nucleotides. Theta is the angle between the imino N-H bond vector and the major axis of inertia of the pseudoknot. The imino resonances of G4 and G10 and G5 and G13 overlap in both the ^1H and ^{15}N dimensions; reported relaxation rates for these nuclei are for the composite (overlapping) peaks.

sufficient to fit the observed and calculated ^{15}N relaxation data. Order parameters for the six imino ^{15}N nuclei were found to be within the range of 0.83 to 0.88, values typical of a well-ordered structure, indicating that the motions of the individual N-H bond vectors are tightly coupled to the overall tumbling of the molecule. The global tumbling time for the VPK pseudoknot was found to be 11.1 ns, slightly longer than that obtained for the bacteriophage T2 and SRV pseudoknots, and may be attributed to hydrodynamic drag due to a less compact loop 2 structure.

In summary, we found that a relatively simple motional model was sufficient to fit the observed and calculated ^{15}N relaxation data for each of the three representative pseudoknots (bacteriophage T2, SRV, and VPK). In all cases, the relaxation data could be back-calculated using a model where the motion of each N-H bond vector is tightly coupled to the overall tumbling of an ellipsoidal molecule. It was not necessary to invoke either chemical exchange or motion on a nanosecond or picosecond time scale to fit the observed and calculated data, indicating that neither the representative frameshifting or nonframeshifting pseudoknots contain complex internal motions on time scales to which the ^{15}N relaxation rate data is sensitive. It should be noted, however, that ^{15}N relaxation data was only obtainable for the stem regions of each structure, where imino proton exchange rates were slow enough for the ^{15}N nuclei to be detected.

Comparison of imino proton exchange rates

Inspection of the ^1H NMR spectra of the bacteriophage T2 and the frameshifting pseudoknots suggests that there are significant differences in the rates at which the imino protons within the base pairs exchange with solvent (Fig. 4), particularly for those base pairs located at the junction of the pseudoknot stems. The relationship between solvent exchange rate and dynamic features of RNA such as the opening frequency of base pairs can be quite complex. When the base pair dissociation constant is much less than one, and the rate of imino proton exchange is significantly greater than the rate of base pair closing, the imino proton exchanges with solvent each time the base pair opens (referred to as an "EX1" mechanism). Alternatively, a base pair may open many times before proton exchange occurs; in this case the rate-limiting step is the base-catalyzed exchange with solvent. The former mechanism has been observed in studies of tRNA (Hurd & Reid, 1980; Johnston & Redfield, 1981; Choi & Redfield, 1995) and a protein–DNA complex (Dhavan et al., 1999), whereas the latter has been observed in several DNAs and drug–DNA complexes, in which case, the rate of base pair opening may be one to two orders of magnitude higher than the imino proton exchange frequency (reviewed by Guéron & Leroy, 1995). The observation of relatively rapid imino proton exchange for a particular base pair within an RNA structure therefore implies that either the base pair has a relatively rapid opening frequency or the efficiency of exchange is accelerated by the presence of a catalyst.

Imino proton exchange rates for the bacteriophage T2 pseudoknot and the representative frameshifting pseudoknots were estimated using a saturation transfer method. For the bacteriophage T2 pseudoknot, imino proton exchange rates for the stem base pairs are all relatively slow, in the range of 0.8 to 4.2 Hz (Fig. 8;

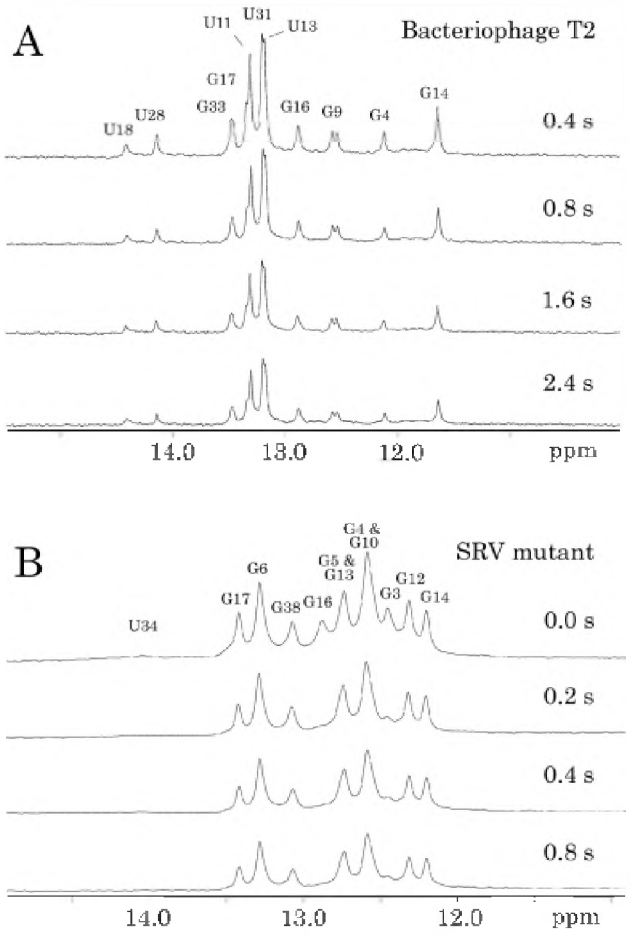

FIGURE 8. Imino proton spectra of the bacteriophage T2 pseudoknot and frameshifting SRV mutant pseudoknot (sequences shown in Figure 3A and D), showing the effect of saturation transfer from the solvent water resonance. **A:** Imino resonances of the bacteriophage T2 pseudoknot are only partially saturated when the solvent resonance is saturated for up to 2.4 s. **B:** In contrast, saturation of the imino resonances of the G16 and U34 of the SRV mutant frameshifting pseudoknot is nearly complete when the solvent resonance is saturated for as little as 0.2 s, indicative of the relatively rapid opening frequency of the A15-U345 and C8-G16 base pairs. The resonance of G3, located at the 5′ end of stem 1, is also significantly influenced by saturation transfer. Spectra were obtained at 20 °C.

Table 3), Interestingly, the exchange rates of imino protons at the junction of the stems are not much different from those in the central regions of the stems. In the SRV frameshifting pseudoknot (Fig. 8; Table 4), exchange rates for most of the base-paired imino protons are in the range of 1 to 2 Hz, with the exceptions of G16 at the junction of the stems (~50 Hz), G3 at the end of stem 1 (~9 Hz), and U34 at the junction of the two stems (~300 Hz), In the VPK frameshifting pseudoknot and the hybrid frameshifting pseudoknot, the imino protons of the base pairs adjacent to the junction of the stems are unobserved, indicating solvent exchange rates of greater than 300 Hz. The data therefore suggest that the opening frequency of the base pairs at the junction of the stems in the bacteriophage T2 pseudoknot is significantly slower than the corresponding base pairs in the frameshifting pseudoknots.

DISCUSSION

A particularly interesting result of the present work is the finding that the bacteriophage T2 pseudoknot is (by a substantial margin) the poorest stimulator of frameshifting among all of the pseudoknots tested. This observation lends support to a hypothesis that the pseudoknots derived from retroviral and coronaviral sequences do indeed have distinctive properties that make them efficient frameshift stimulators. NMR analyses of the nonframeshifting bacteriophage T2 pseudoknot (Holland et al., 1999), a frameshifting pseudoknot derived from the SRV sequence (Michiels et al., 2001), and a hybrid of these pseudoknots (in the present work) shows that several significant structural features are shared: Specifically, each pseudoknot has a similar local structure at the junction of stems 1 and 2, a similar overrotation of the base pairs at the junction of the stems, there is little bending between the stems, and in each case loop 1 consists of a single nucleotide. Apparently, the key that confers frameshifting ability does not lie solely among these particular structural features.

In the present work, experiments were designed to test whether the presence of an unpaired (wedged) adenosine at the junction of the pseudoknot stems, or a significantly bent structure, are important determinants of frameshift efficiency, It has previously been shown that the removal of an unpaired adenosine from the junction of the two stems of a pseudoknot derived from the MMTV *gag-pro* shift site reduces frameshift-stimulating ability (Chen et al., 1995), and this change has also been shown to be destabilizing (Theimer & Giedroc, 1999, 2000), Results show that neither the wedged adenosine nor the bent structure are absolute requirements; these conclusions are consistent with, and add further support to, the recent results of Michiels et al. (2001) in their structural and mutational studies of other pseudoknots with sequences derived from the SRV *gag-pro* frameshift site. The identities of the nucleotides near the stem1/stem2 junction (either in base pairs or in loop 2) do, however, have an influence upon frameshifting efficiency,

Why does the bacteriophage T2 pseudoknot fail to stimulate frameshifting? The results of the present work point toward an insufficiently stable stem 1, as increasing the G-C content of stem 1 (as was done in making our hybrid pseudoknot) did create a reasonably efficient stimulator of frameshifting. Comparison of the nucleotide sequences of naturally occurring viral frameshifting pseudoknots reveals that many do have a G-rich 5′ end of stem 1 (Brierley, 1995), Because pseudoknots are believed to have the ability to pause translating ribosomes (Tu et al., 1992; Somogyi et al., 1993; Lopinski et al., 2000), and because many of the retro-

viral pseudoknots have G-C rich stems, it would at first glance appear to be a plausible hypothesis that pseudoknot stability should be the primary determinant for frameshifting efficiency. However, a high G-C content in stem 1 is not by itself sufficient for frameshifting, as examples of such pseudoknots with poor frameshifting abilities have been reported (Kang et al., 1996), and frameshifting efficiency can be quite sensitive to mutations in regions of the structure other than stem 1. Investigations of the SRV-1 *gag-pro* pseudoknot (ten Dam et al., 1995) and the IBV pseudoknot (Liphardt et al., 1999; Napthine et al., 1999) indicate that frameshift-stimulating ability does not appear to be related to stem stability in a simple way. Interestingly. the process of creating a frameshifting pseudoknot from the nonframeshifting bacteriophage T2 pseudoknot did not introduce major structural changes, at least in the stems, and at the junction of the stems.

In the bacteriophage T2 pseudoknot, the imino protons in the base pairs at the junction of the two stems exchange with the solvent relatively slowly, compared to the analogous protons in the frameshifting pseudoknots; these observations suggest that the opening frequencies for the junction base pairs are relatively slow in the bacteriophage pseudoknot. Pseudoknot dynamics at the junction of the stems may therefore play a role in determining frameshifting efficiency, Mutational data show that the identity of the loop 2 nucleotide nearest the junction influences frameshifting; this loop 2 nucleotide is among those most likely to influence pseudoknot dynamics at the junction of the two stems. Interestingly. Liphardt et al. (1999) found a requirement for adenosine as the last nucleotide in loop 2 in a recent mutational analysis of the IBV frameshifting pseudoknot. Recent crystallographic studies of the BYWV frameshift-stimulating pseudoknot (Egli et al., 2002) indicate multiple conformations for several of the nucleotides, suggesting flexibility for this pseudoknot as well. A potential relationship between pseudoknot dynamics at the junction of the stems and frameshifting is intriguing in light of the fact that pseudoknots do not stimulate frameshifting with 100% efficiency; perhaps pseudoknots are in dynamic equilibrium between active and inactive forms. Because RNA structures are intrinsically flexible, it is possible that the signal for a frameshift is hidden within a pseudoknot conformation that is only occasionally sampled by the RNA. Perhaps the relatively rigid (and nonframeshifting) bacteriophage pseudoknot does not sample the specific conformation at the junction of its stems that is required for frameshifting activity, even though its average structure is similar to that of the more dynamic frameshifting pseudoknots.

In summary, the results of the present study provide significant new data regarding the relationships between pseudoknot primary sequence, structure, dynamics, and frameshifting efficiency. Results show that these relationships are complex. An explanation for this apparent complexity could be provided by the presence of specific RNA–RNA or RNA–protein interactions between the pseudoknot and the translational apparatus, although the existence of such interactions remains hypothetical at this time. As further biochemical, structural, and thermodynamic data on frameshift-stimulating sites within messenger RNAs become available, the mechanisms and requirements for an efficient frameshift event will likely become more fully understood.

MATERIALS AND METHODS

In vitro frameshift assays

Complementary oligonucleotides with protruding *Bst*EII and *Kpn*I ends were cloned into vector pGB01 (Matsufuji et al., 1995) such that the downstream rabbit β-globin coding region was in the -1 reading frame relative to the upstream GST coding region. Clones with the correct inserts were identified by DNA sequencing and plasmids were purified by cesium chloride ultracentrifugation. Plasmid DNAs were linearized with restriction endonuclease, *Sal*I, and used as templates for T7 transcription. Transcripts were translated in reticulocyte lysates (Promega) with ^{35}S-methionine. Protein products were separated by SDS-PAGE and visualized and quantified on a Molecular Dynamics PhosphorImager using ImageQuant software. Frameshifting efficiencies were determined after normalization for methionine content in the termination and frameshift products. The frameshifting efficiency of the DH40 hybrid pseudoknot (Fig. 3B) was measured in four independent assays, giving an average value of 18.5% $\pm$ 0.6%. The error for the relative efficiencies of frameshifting was therefore $\pm$3%. All frameshifting percentages were rounded to the nearest whole number. The dual-luciferase reporter system was used as described (Grentzmann et al., 1998). Oligonucleotides containing the HERV shift site plus the pseudoknot were cloned into the *Sal*I and *Bam*HI sites of vector p2luc. Control constructs (containing an extra nucleotide at the shift site to place the two luciferase coding sequences in the same reading frame) were also constructed. Transcripts were generated, as above, after plasmids were linearized with *Pml*I, and translated in reticulocyte lysates supplemented with a full complement of amino acids. Firefly and renilla luciferase activities were sequentially assayed on a Dynatech MLX microtiter plate luminometer using the Dual-Luciferase TM reporter assay reagents (Promega). The -1 frameshifting efficiency was determined by comparing the ratios of firefly luciferase to renilla luciferase activities of the test versus the control constructs.

NMR experiments

The RNA molecules used in NMR experiments, with the sequences shown in Figure 3, were transcribed 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. NMR spectra of the RNA pseudoknots were collected at 500 MHz using a Varian Unity-Inova spectrometer. Samples typically contained 6 to 10 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 to 6.8. Two-dimensional NOE spectra in 90% H_2O/10% D_2O were acquired using the jump and return method for solvent suppression (Plateau & Guéron, 1982) and mixing times of 60 to 200 ms. ^{15}N-^{1}H correlated spectra were acquired using either a *z*-axis pulsed field gradient or the jump and return method for solvent suppression. Two-dimensional NOE spectra in 99.9% D_2O solvent were acquired using presaturation for solvent suppression, and several mixing times between 60 and 300 ms. TOCSY (60 ms mixing time) and DQF-COSY spectra were obtained in 99.9% D_2O solvent using presaturation for solvent suppression.

R_1, R_2, and ^{15}N-^{1}H NOEs for the base-paired guanosine N1 and uridine N3 nitrogens were detected via the 90-Hz (single-bond) coupling to the imino protons, using pulse sequences described by Farrow et al. (1994), which include gradient selection, sensitivity enhancement, and pulses for minimizing the saturation of the water. R_1 relaxation times were calculated from peak heights in seven data sets acquired with relaxation delays of 10 to 1,000 ms (obtained at 293 K). Similarly, five data sets were acquired for the R_2 relaxation measurements, with relaxation delays of 29 to 145 ms. ^{15}N-^{1}H NOEs were measured using data sets acquired with and without ^{1}H-presaturation for 4 s, with a 6-s recycle time between the acquisition of each free induction decay, Presaturation of the ^{1}H spectrum was accomplished using a 4-s-long series of 120 degree nonselective ^{1}H pulses separated by 5-ms delays. The R_1, R_2, and NOE data sets were each processed identically. Relaxation times were found by fitting to a single exponential decay equation. Parameters describing the motions of the pseudoknot were derived from the R_1, R_2, and NOE data (Mandel et al., 1995) with the assistance of the Modelfree 4.0 and pdb_inertia programs (obtained from the web site of Dr. Arthur G. Palmer, see http://cpmcnet.columbia.edu/dept/gsas/biochem/labs/palmer), with the relaxation analyses being performed as previously described (Lillemoen & Hoffman, 1998). Throughout the analysis, a distance of 1.01 Å was assumed for the imino ^{15}N-^{1}H bond lengths, and a chemical shift anisotropy of -130 and -100 ppm was assumed for the guanosine N1 and uridine N3 nuclei, respectively (Akke et al., 1997).

Saturation transfer was used to estimate of the rates at which the observed imino protons within the Watson–Crick base pairs exchange with solvent (cf. Lillemoen et al., 1997). A series of spectra were acquired in 90% H_2O/10% D_2O solvent using a jump-return selective excitation pulse sequence for solvent suppression (Plateau & Guéron, 1982; Spera et al., 1991), with selective saturation of the solvent H_2O resonance for a range of times from 0.2 to 3 s prior to the jump-return sequence. When the solvent resonance is saturated, the imino proton resonance intensities are reduced by an amount that is related to the fraction that have exchanged with the solvent, with additional intensity changes occurring due to the nuclear Overhauser effect and T_1 relaxation (Forsen & Hoffman, 1963). The observed imino resonance intensities were fit to the equation

$$I(t)/I(0) = (1 + T_1)/(1 - T_1)\exp(-K_{HX}t) + \eta(t),$$

where I(t) is the imino resonance intensity after the solvent has been saturated for a time t, K_{HX} is a rate constant for proton exchange with the solvent. η is used to account for the NOE, and was assumed to be zero in the absence of saturation (at $t = 0$), and to be a negative constant at times $t >$ 0.2 s. It is noted that the time dependence of the NOE is actually more complex than this assumption. However, the decay of resonance intensity is dominated by the saturation transfer effects, so that approximate yet comparatively useful values for the solvent exchange rates can be obtained in this manner. The imino proton exchange rates of nucleotides G16 and U34 of the SRV mutant frameshifting pseudoknot were estimated from the resonance linewidth, because the exchange was too rapid to estimate by saturation transfer.

The NMR data provided information from which the structural features of the hybrid pseudoknot were derived, with a particular emphasis on the local structure at the junction of the two stems. The C6-G14 and A13-U29 base pairs at the junction of the helical stems were defined by 27 NOE-derived distance constraints. Nucleotides G8, C34, A35, and C36 at the 3′ end of the molecule were defined by 33 NOE-derived distance constraints. Nucleotides of the central region of each stem (nt 2–5, 9–12, 15–18, and 30–33) were restricted to near A-form helical geometry (Arnott et al., 1972) during the simulated annealing process, as all of the NOE data arising from these base pairs were consistent with an A-form helix. The loop nucleotides were unrestrained in the structure analysis. Depending on its intensity, each NOE cross peak was placed into one of four categories ranging from strong to very 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). The majority of distance 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; several additional NOEs were detected in a spectrum with a mixing time of 300 ms, and these were assigned to distance ranges of 4.5 to 6.5 Å. These NOE/distance assignments were internally verified by comparing with NOE cross peaks arising from the regular A-form helical stem regions of the pseudoknot, where the interproton distances were assumed to be near to the ideal A-form helical values. Torsion angle constraints were used to restrict the glycosidic torsion angle χ of nt 6, 8, 13, 14, 29, 34, 35, and 36 to *anti* (-170 ± 20 degrees) conformation. The torsion angles within the riboses of nt 2–6, 8–18, and 29–34 were restricted C3′ *endo*; torsion angles in other riboses were not restricted. Hydrogen bonds within the 12 Watson–Crick base pairs were introduced as distance constraints between the atoms participating in the hydrogen bond. Hydrogen bond constraints were considered justified for the A13-U29 and C6-G14, as the observed NOEs for nonexchangeable protons were consistent with these nucleotides being Watson–Crick base paired. Hydrogen bond constraints were introduced as a range of distances, with two distances used for each hydrogen bond, chosen so that the

hydrogen bonds were restricted to be within 17 degrees of linear. No additional angle constraints were used to force the bases within each pair to be within the same plane. The structure calculation for the hybrid pseudoknot was carried out using the simulated annealing and energy minimization protocols within the X-PLOR version 3.1 program suite (Brünger, 1993). 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 200 diverse starting structures were used, to insure that the full range of possible conformations were sampled during the simulated annealing process. Starting structures included those with a wide range of 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 constraints exceeded 8 degrees, (4) the angle formed by the three atoms comprising each hydrogen bond involving imino protons was less than 17 degrees, and (5) the angle formed by the three atoms comprising each hydrogen bond involving amino protons was less than 23 degrees. 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 10 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 42 accepted models for the pseudoknot structure. 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 stem 1 and the line defining the helix axis of stem 2, calculated using the coordinates of the central base pairs of each stem.

ACKNOWLEDGMENTS

This work was supported by American Cancer Society grant GMC-89306, National Institutes of Health (NIH) grants R01-AI-40187 and K02-AI01497, and Welch Foundation grant F-1353 to D.W.H., U.S. Department of Energy grant DE-FG03-01ER63132 to R.F.G., and NIH grant GM14852 to J.F.A.

Received March 5, 2002; returned for revision
May 2, 2002; revised manuscript received
May 21, 2002

REFERENCES

Akke M, Fiala R, Jiang F, Patel D, Palmer AG. 1997. Base dynamics in a UUCG tetraloop RNA hairpin characterized by ^{15}N spin relaxation: Correlations with structure and stability. *RNA 3*:702–709.

Arnott S, Hukins DWL, Dover SD. 1972. Optimized parameters for RNA double-helices. *Biochem Biophys Res Commun 48*:1392–1398.

Brierley I. 1995. Ribosomal frameshifting on viral RNAs. *J Gen Virol 76*:1885–1892.

Brierley I, Digard P, Inglis SC. 1989. Characterization of an efficient coronavirus ribosomal frameshifting signal: Requirement for an RNA pseudoknot. *Cell 57*:537–547.

Brünger AT. 1993. *X-PLOR Version 3.1: A System for X-ray Crystallography and NMR*. New Haven, CT: Yale University Press.

Chamorro M, Parkin N, Varmus HE. 1992. An RNA pseudoknot and an optimal heptameric shift site are required for highly efficient ribosomal frameshifting on a retroviral messenger RNA. *Proc Natl Acad Sci USA 89*:713–717.

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.

Choi BS, Redfield AG. 1995. Proton exchange and basepair kinetics of yeast tRNA(Phe) and tRNA(Asp1). *J. Biochem (Tokyo) 117*: 515–520.

Dhavan GM, Lapham J, Yang S, Crothers DM. 1999. Decreased imino proton exchange and base-pair opening in the IHF-DNA complex measured by NMR. *J Mol Biol 288*:659–671.

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 possible family of structurally related RNA pseudoknots. *Biochemistry 35*:4187–4198.

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*:464–470.

Egli M, Minasov G, Li S, Rich A. 2002. Metal ions and flexibility in a viral RNA pseudoknot at atomic resolution. *Proc Natl Acad Sci USA 99*:4302–4307.

Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, Forman-Kay JD, Kay LE. 1994. Backbone dynamics of a free and phospho-peptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry 33*:5984–6003.

Forsen S, Hoffman RA. 1963. A new method for the study of moderately rapid chemical exchange rates employing nuclear magnetic resonance double resonance. *Acta Chem Scand 17*:178–193.

Garcia A, van Duin J, Pleij CWA. 1993. Differential response to frameshift signals in eukaryotic and prokaryotic translational systems. *Nucl Acids Res 21*:401–406.

Giedroc DP, Theimer CA, Nixon PL. 2000. Structure, stability and function of RNA pseudoknots involved in stimulating ribosomal frameshifting. *J Mol Biol 298*:167–185.

Grentzmann G, Ingram J, Kelly PJ, Gesteland RF, Atkins JF. 1998. A dual-luciferase reporting system for studying recoding signals. *RNA 4*:479–486.

Guéron M, Leroy JL. 1995. Studies of base pair kinetics by NMR measurement of proton exchange. *Methods Enzymol 261*:383–413.

Holland JA, Hansen MR, Du Z, Hoffman DW. 1999. Coaxial stacking of helical stems in a pseudoknot motif: An NMR investigation of the gene *32* messenger RNA pseudoknot of bacteriophage T2. *RNA 5*:257–271.

Hurd RE, Reid BR. 1980. Helix-coil dynamics in RNA: The amino acid acceptor helix of *Escherichia coli* phenylalanine transfer RNA. *J Mol Biol 142*:181–193.

Jacks T, Townsley K, Varmus HE, Majors J. 1987. Two efficient ribosomal frameshifting events are required for synthesis of mouse mammary tumor virus *gag*-related polyproteins. *Proc Natl Acad Sci USA 84*:4298–4302.

Johnston PD, Redfield AG. 1981. Study of transfer ribonucleic acid unfolding by dynamic nuclear magnetic resonance. *Biochemistry 20*:3996–4006.

Kang H, Hines JV, Tonoco I. 1996. Conformation of a non-frameshifting RNA pseudoknot from mouse mammary tumor virus. *J Mol Biol 259*:135–147.

Kontos H, Napthine S, Brierley I. 2001. Ribosomal pausing at a frame-

shifter RNA pseudoknot is sensitive to reading phase but shows little correlation with frameshift efficiency. *Mol Cell Biol 21*:8657–8670.
Lee LK, Rance M, Chazin WJ, Palmer AG. 1997. Rotational diffusion anisotropy of proteins from simultaneous analysis of ^{15}N and $^{13}C^{\alpha}$ nuclear spin relaxation. *J Biomol NMR 9*:287–298.
Lillemoen J, Cameron CS, Hoffman DW. 1997. Stability and dynamics of ribosomal protein L9: An investigation of a molecular strut by amide proton exchange and circular dichroism. *J Mol Biol 268*:482–493.
Lillemoen J, Hoffman DW. 1998. An investigation of the dynamics of ribosomal protein L9 using heteronuclear NMR relaxation measurements. *J Mol Biol 281*:539–551.
Liphardt J, Napthine S, Kontos H, Brierley I. 1999. Evidence for an RNA pseudoknot loop-helix interaction essential for efficient −1 frameshifting. *J Mol Biol 288*:321–335.
Lopinski JD, Dinman JD, Bruenn JA. 2000. Kinetics of ribosomal pausing during programmed −1 translational frameshifting. *Mol Cell Biol 20*:1095–1103.
Luginbühl P, Pervushin KV, Iwai H, Wüthrich K. 1997. Anisotropic molecular rotational diffusion in ^{15}N spin relaxation studies of protein mobility. *Biochemistry 36*:7305–7312.
Mandel AM, Akke M, Palmer AG. 1995. Backbone dynamics of *Escherichia coli* ribonuclease HI: Correlations with structure and function in an active enzyme. *J Mol Biol 246*:144–163.
Matsufuji S, Matsufuji T, Miyazaki Y, Murakami Y, Atkins JF, Gesteland RF, Hayashi S. 1995. Autoregulatory frameshifting in decoding the mammalian ornithine decarboxylase antizyme. *Cell 80*:51–60.
Michiels PJA, Versleijen AAM, Verlaan PW, Pleij CWA, Hilbers CW, Heus HA. 2001. Solution structure of the pseudoknot of SRV-1 RNA, involved in ribosomal frameshifting. *J Mol Biol 310*:1109–1123.
Moore R, Dixon M, Smith R, Peters G, Dickson C. 1987. Complete nucleotide sequence of a milk-transmitted mouse mammary tumor virus. Two frameshift events are required for translation of *gag* and *pol*. *J Virol 61*:480–490.
Napthine S, Liphardt J, Bloys A, Routledge S, Brierley I. 1999. The role of the RNA pseudoknot stem 1 length in the promotion of efficient −1 ribosomal frameshifting. *J Mol Biol 288*:305–320.
Nixon PL, Giedroc DP 2000. Energetics of a strongly pH dependent RNA tertiary structure in a frameshifting pseudoknot. *J Mol Biol 296*:659–671.
Ono M, Yasunaga T, Miyata T, Ushkubo H. 1986. Nucleotide sequence of human endogenous retrovirus genome related to the mouse mammary tumor virus genome. *J Virol 60*:589–598.
Orekhov VY, Nolde DE, Golovanov AP, Korzhne DM, Arseniev AS. 1995. Processing of heteronuclear NMR relaxation data with the new software DASHA. *Appl Magn Reson 9*:581–588.
Plateau P, Guéron M. 1982. Exchangeable proton NMR without line distortion. *J Am Chem Soc 104*:7310–7311.
Saenger W. 1984. *Principles of nucleic acid structure.* New York: Springer-Verlag.
Schommer S, Sauter M, Krausslich HG, Best B, Müeller-Lantzsch N. 1996. Characterization of the human endogenous retrovirus K proteinase. *J Gen Virol 77*:375–379.
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.
Somogyi P, Jenner AJ, Brierley I, Inglis SC. 1993. Ribosomal pausing during translation of an RNA pseudoknot. *Mol Cell Biol 13*:6931–6940.
Spera S, Ikura M, Bax A. 1991. Measurement of the exchange rates of rapidly exchanging amide protons: Application to the study of calmodulin and its complex with the myosin light chain kinase. *J Biomol NMR 1*:155–165.
Su L, Chen L, Egli M, Berger JM, Rich A. 1999. A minor groove RNA triplex in the crystal structure of a viral pseudoknot involved in ribosomal frameshifting. *Nat Struct Biol 6*:285–92.
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. *Nuclec Acids Res 22*:2304–2310.
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*:146–154.
Theimer CA, Giedroc DP. 1999. Equilibrium unfolding pathway of an H-type pseudoknot which promotes programmed −1 ribosomal frameshifting. *J Mol Biol 289*:1283–1299.
Theimer CA, Giedroc DP. 2000. Contribution of the intercalated adenosine at the helical junction to the stability of the *gag-pro* frameshifting pseudoknot from mouse mammary tumor virus. *RNA 6*:409–421.
Tu C, Tzeng TH, Bruenn JA. 1992. Ribosomal movement impeded at a pseudoknot required for frameshifting. *Proc Natl Acad Sci USA 89*:8636–8640.
Voet D, Rich A. 1970. The crystal structures of purines, pyrimidines and their intermolecular complexes. *Prog Nucleic Acid Res Mol Biol 10*:183–265.
Yoshinaka Y, Katoh I, Copeland TD, Oroszlan SJ. 1985. Murine leukemia virus protease is encoded by the *gag-pol* gene and is synthesized through suppression of an amber termination codon. *Proc Natl Acad Sci USA 82*:1618–1622.