Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

doi:10.1006/jmbi.2001.4823 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2001) 310, 1109–1123

JMB

Solution Structure of the Pseudoknot of SRV-1 RNA, Involved in Ribosomal Frameshifting

Paul J. A. Michiels[1], Alexandra A. M. Versleijen[1], Paul W. Verlaan[2] Cornelis W. A. Pleij[2], Cornelis W. Hilbers[1] and Hans A. Heus[1]*

[1]*NSR Center for Molecular Structure, Design and Synthesis, Laboratory of Biophysical Chemistry University of Nijmegen Toernooiveld, 6525 ED Nijmegen, The Netherlands*

[2]*Leiden Institute of Chemistry, Gorlaeus Laboratories University of Leiden Einsteinweg 55, 2333 CC Leiden, The Netherlands*

**Corresponding author*

RNA pseudoknots play important roles in many biological processes. In the simian retrovirus type-1 (SRV-1) a pseudoknot together with a heptanucleotide slippery sequence are responsible for programmed ribosomal frameshifting, a translational recoding mechanism used to control expression of the Gag-Pol polyprotein from overlapping *gag* and *pol* open reading frames. Here we present the three-dimensional structure of the SRV-1 pseudoknot determined by NMR. The structure has a classical *H*-type fold and forms a triple helix by interactions between loop 2 and the minor groove of stem 1 involving base-base and base-sugar interactions and a ribose zipper motif, not identified in pseudoknots so far. Further stabilization is provided by a stack of five adenine bases and a uracil in loop 2, enforcing a cytidine to bulge. The two stems of the pseudoknot stack upon each other, demonstrating that a pseudoknot without an intercalated base at the junction can induce efficient frameshifting. Results of mutagenesis data are explained in context with the present three-dimensional structure. The two base-pairs at the junction of stem 1 and 2 have a helical twist of approximately 49°, allowing proper alignment and close approach of the three different strands at the junction. In addition to the overwound junction the structure is somewhat kinked between stem 1 and 2, assisting the single adenosine in spanning the major groove of stem 2. Geometrical models are presented that reveal the importance of the magnitude of the helical twist at the junction in determining the overall architecture of classical pseudoknots, in particular related to the opening of the minor groove of stem 1 and the orientation of stem 2, which determines the number of loop 1 nucleotides that span its major groove.

Keywords: frameshifting; NMR; pseudoknot; RNA; SRV-1

Introduction

Simian retrovirus type-1[1] uses programmed −1 ribosomal frameshifting as a mechanism for regulating translation of its polycistronic mRNA. The frameshift enables the ribosome to skip the stop codon in the 0 frame and to proceed protein synthesis until the first stop codon in the −1 frame is reached. After synthesis the polyprotein is processed by proteolysis resulting in the functional proteins pro, gag and pol. Programmed ribosomal frameshifting has not only been observed for retroviruses, but also in coronaviruses,[2] plant viruses,[3] yeast,[4,5] and bacterial systems .[6,7]

Two elements in the mRNA are responsible for efficient ribosomal frameshifting, (i) a heptanucleotide sequence and (ii) a secondary structural element six to eight nucleotides downstream of this sequence, which in almost all cases is a pseudoknot. The heptanucleotide or so-called shifty sequence has an X XXY YYZ consensus, where X can be any base, Y is A or U and Z is A, U or C. It is thought that when this sequence enters the A and P site of the ribosome, the pseudoknot structure causes pausing of the ribosome, giving time to

Abbreviations used: SRV-1, simian retrovirus type-1; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy, MMTV, mouse mammary tumour virus; IBV, infections bronchitis virus; HMQC, heteronuclear multiple-quantum correlation spectroscopy; TYMV, turnip yellow mosaic virus; BWYV, beet western yellow virus.

E-mail address of the corresponding author: hans@nmr.kun.nl

shift to the −1 reading frame. Although in most cases a pseudoknot is found to be responsible for ribosomal frameshifting, also a hairpin structure in HIV[8] and *Escherichia coli dnaX*[9,10] and a three-way junction in IS911[11] have been shown to be functionally active. Whether the pseudoknot structure acts as an independent unit, regulating the ribosomal frameshift by its structure or stability, or acts in concert with auxiliary proteins is still not understood. The importance of its structural identity is evident, however, since pseudoknots can for instance not be replaced by a functionally different pseudoknot[12] or replaced by a hairpin with a similar base-pair composition.[13]

The pseudoknot of SRV-1 RNA is of the classical *H*-type,[14–16] with two stacked stems S1 and S2, connected by two loops L1 and L2. Mutational analysis of the pseudoknot in the overlapping region of the *gag* and *pol* genes of SRV-1 showed that the lengths of L1 and L2 and composition of the base-pairs in S1 are important for its function.[17,18] In the SRV-1[18] and other frameshifting pseudoknots[19] changing L2 nucleotides close to the junction can have drastic effects on frameshift efficiency. These results, together with recently solved structures of other *H*-type pseudoknots,[20,21] suggested a prominent structural role for L2 having various tertiary interactions at the junction.

Here we present the high resolution NMR structure of the RNA pseudoknot present in the overlapping region of the *gag* and *pol* genes of SRV-1. The structure is further characterized by a highly organized L2-S1 interface, which forms a triple helix by various tertiary interactions and extensive stacking. Base-pairing at the junction of the two stems is maintained, which shows that a pseudoknot without an intercalated base at the junction can induce a frameshift. The structure explains nearly all results of mutational studies obtained so far, and the architecture of the structure helps in understanding how a pseudoknot promotes frameshifting.

Results

Design of an SRV-1 pseudoknot sequence optimal for NMR

An optimal SRV-1 pseudoknot sequence was designed for the NMR experiments by introducing a few mutations, while maintaining its frameshift efficiency (Figure 1). (i) The third G-C base-pair in S2 was changed into an A-U base-pair. This enabled us to distinguish the two stems, which are both very G-C rich, by NMR spectroscopy, and avoided the possibility of rearrangement of the pseudoknot into an alternative structure with two hairpins. (ii) The three GCU residues in L2 were deleted. Frameshift efficiencies for mutants (i) and (ii)[17,18] were found to be higher compared to the wild-type level (23 %, Figure 1(a)), amounting to 25 % and 30 %, respectively (Figure 1(b) and (c)).

However, pseudoknots with a guanosine at position 20 showed a tendency to form multimers. This was perceived from a concentration-dependent behaviour of the imino proton resonances (between 10 and 300 μM). We believe that the concentration dependence originated from formation of dimer pseudoknot structures in which the G20-A21′ and A21-G20′ form an intermolecular sheared tandem G-A base-pair motif at the dimer interface.[22,23] Mutational studies[18] showed that chan-

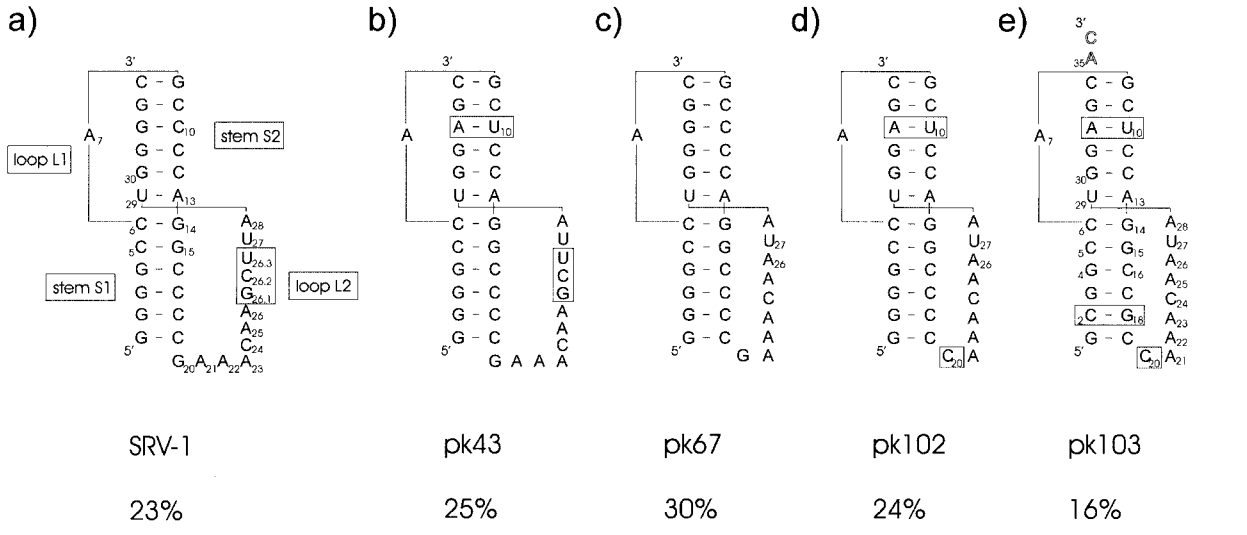

Figure 1. Secondary structure representation of the SRV-1 wild-type (a) and various mutant pseudoknots (b)-(e) with their corresponding frameshifting efficiencies. Mutations as compared to the wild-type are boxed. The two stabilizing residues at the 3′-end of pk103 are drawn as open characters. The three boxed residues in L2 highlighted in (a) and (b) have been deleted in (c)-(e), the boxed base-pairs in S1 and S2 as well as C20 (d), (e) indicate mutations with respect to the wild-type.

ging ten nucleotides at the 5′-side of L2 yields a frameshift efficiency close to the wild-type level (20 %). This suggests that the identity of the bases in the SRV-1 L2 are not important. Therefore, (iii) we changed G20 into a cytidine, thus destabilising the putative pseudoknot-dimer interaction, and indeed for this mutant, pk102, no concentration-dependence behaviour of the imino proton resonances was observed. Frameshift efficiency for the triple mutant pk102 (Figure 1(d)) was again comparable to wild-type level, 24 %.

However, pk102 showed an increased intensity of two broad peaks in the region of the NMR spectrum where the non-hydrogen bonded imino protons tend to resonate (10.5-11.5 ppm), which indicated aggregation of a different kind. This multimerization effect could be attributed to the four consecutive guanosine residues at the 5′-side of the RNA pseudoknot. These base-pairs are probably important for function: in almost all frameshifting pseudoknots, S1 starts with three or four G-C base-pairs, and mutational studies showed that changing the third base-pair decreased frameshift efficiency by approximately 50 %.[17] To be able to study a monomeric RNA pseudoknot (iv) we changed the second base-pair of S1 into a C-G base-pair, thereby interrupting the stretch of four guanosine bases. This mutation resulted in a complete disappearance of the two broad imino-proton peaks around 11 ppm supporting our hypothesis that the multimerization was due to the four consecutive guanines.

Finally, evaluation of a few additional H5/H6 cross-peaks with minor intensities (approximately 10 %) in the TOCSY experiment pointed out the presence of a species with a local alternative conformation, originating from heterogeneity at the 3′-end. The purified $n+1$ transcript showed less of the undesired minor conformation, which is probably caused by stacking effects of the additional single-stranded nucleotide at the 3′-end. Therefore, two residues, an adenosine and a cytidine, were added to the 3′-end of the pseudoknot, thus favouring the major conformation. The final mutant, pk103, combined all necessary changes for studying the monomeric form of the SRV-1 pseudoknot by NMR, while maintaining an adequate level of frameshift efficiency (16 %, Figure 1(e)).

Resonance assignment and structure determination

The secondary structure of pk103, which turned out to be identical to the published model,[17] was evident from a complete sequential NOE walk, connecting all base-paired imino protons with the exception of G1 (typical for a 5′-fraying residue) and the step between G14 and U29 at the S1-S2 junction. Despite the absence of a NOE between the slightly broadened imino proton resonances of residues G14 and U29, most likely due to dynamic properties of the junction and enhanced chemical exchange of the imino protons, the assignment of the imino proton of U29 and the presence of a U-A base-pair at the junction could be established unambiguously. While a single adenine spans the major groove, still a clear sequential NOE contact was observed between G8 H1 and G33 H1 at the 3′-end.

Standard assignment procedures,[24,25] using homo- and heteronuclear experiments in ^2H_2O led to a near complete assignment of the non-exchangeable protons. The sequential walk from anomeric to aromatic protons was interrupted between residues 6-8, 19-20, in the middle of L2 at residues 23-24 and at residues 28-29, indicative of a change in direction of the backbone or increased flexibility in these regions. Strong intra-residue NOEs between H1′/H2′/H3′ and H6 of cytidine 24 suggested higher flexibility for this residue in comparison with the rest of the loop residues. This was confirmed by initial $T_1\rho$ relaxation experiments, which indicated an increased flexibility for this residue. All residues involved in Watson-Crick base-pairs showed typical *A*-type helical NOE contacts. A structurally important ''pseudo-sequential'' NOE contact was observed between C6 H2′ and U29 H5 and H6 which helped restraining the relative orientation of the two stems.

Sequential NOEs, involving ribose and aromatic C6/C8 protons observed for practically all residues in L2 indicated stacking in an *A*-type fashion. Assignment of the loop 2 adenine H2 resonances was obtained using a 3D HCCH-TOCSY,[26] which connected each H2 with the already assigned C8 protons. This allowed for resonance assignment using the large number of NOEs observed between the H2s and the H1s of different residues in S1 and L2. In Figure 2(a) strips of all H2 resonances in the 3D NOESY-HMQC are shown with NOEs to other H2 and H1s.

The high adenine content of L2, six out of nine residues, and the large number of NOEs turned out to be very important in the structure determination. All L2 adenine H2s showed sequential NOEs to H1s in L2, and a number of long range NOEs, predominantly to the S1 strand running anti-parallel to L2 (schematically displayed in Figure 2(b)). Figure 3 shows a schematic presentation of all 547 NOE-derived distance restraints that were used together with 272 dihedral restraints. For the structure calculations the torsion angle dynamics approach implemented in the X-PLOR program[27] was used as described by Stein *et al*.[28] Mainly due to the large number of distance restraints between L2 and S1, the structure calculations resulted in a very compact and well defined pseudoknot structure as shown by the superimposed structures in Figure 4.

Description of the structure

A single adenine spans the major groove of S2

The S2 stem consists of six base-pairs. Its major groove is spanned by the single residue, A7, with

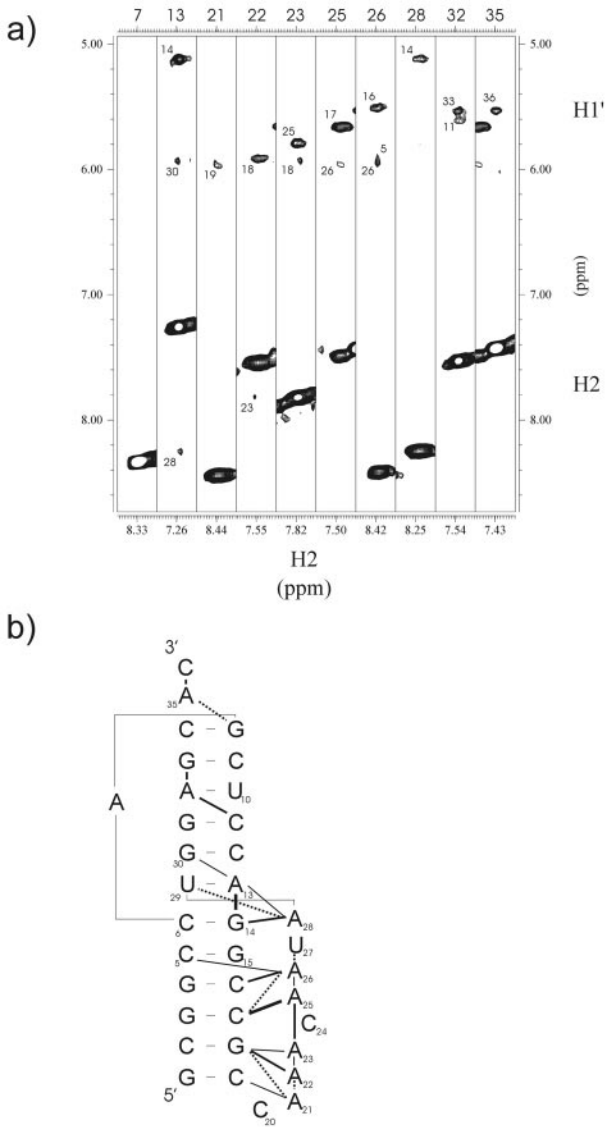

Figure 2. (a). Strips from a 3D NOESY-HMQC experiment belonging to all H2 protons of pk103 showing long range NOEs to other H2 and H1′ protons. (b) Overview of all NOEs from H2 to H1′ protons observed in 2D NOESY and 3D NOESY-HMQC experiments. The thickness of the lines corresponds with the intensity of the NOEs.

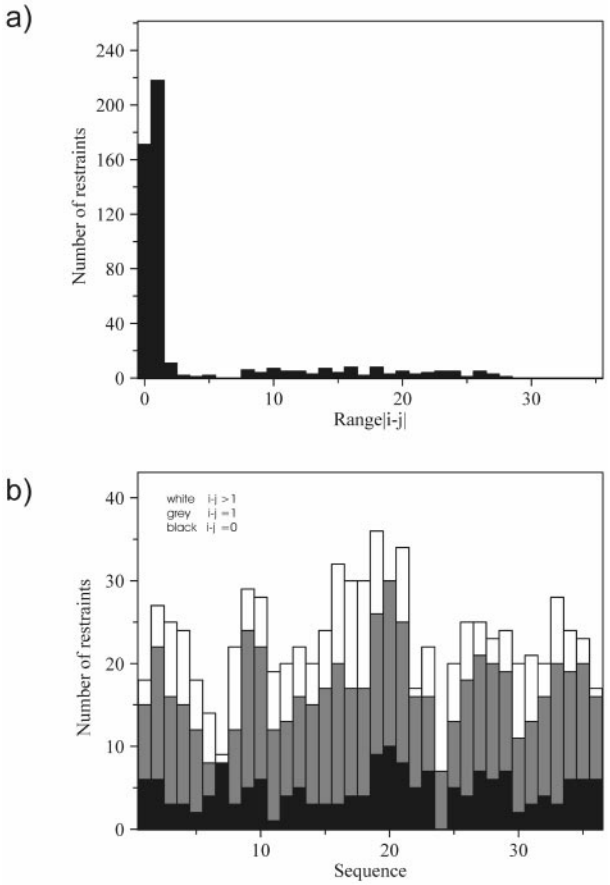

Figure 3. (a) Bar graph of all distance restraints used in the structure calculation. (a) Representation of the number of distance restraints between residues *i* and *j*. (b) Representation of the number of distance restraints and the range for the individual residues. Intraresidual distance restraints are coloured black, sequential grey and everything beyond sequential is coloured white. Bar graphs are generated using DYANA[62] and the number of restraints, as counted by DYANA, are devided by a factor of two resulting in the ''real'' values as used in the structure calculation.

the base buried deeply into the groove (Figure 5(a)). The sugar pucker of A7 is of the S-type and the γ torsion angle is *trans* instead of the normal *gauche*$^+$, resulting in an average elongated phosphate-phosphate distance of 6.8 Å between residues 7 and 8. Similar atypical torsion angles were observed in the TYMV pseudoknot for extending one of the two residues of L1.[29] However, the phosphate-phosphate distance over six base-pairs in an *A*-type helix is approximately 11 Å,[14,30] which is too large for a single residue to span. To compensate for this effect, the distance is shortened by a large helical twist of 49° between the two base-pairs at the junction (*vide infra*).

Continuous stacking and base triple interactions form a highly structured L2-S1 interface

No less than 36 long range contacts were determined between the C2 protons of all adenines in L2 and the base-pairs in S1. The large number of NOE-derived distance restraints results in a very well defined structure of L2, which is locked in the minor groove of S1 by various interactions. At the 5′-end of L2, at C20, the backbone turns sharply into the opposite direction and is held in position by A21, which forms two hydrogen bonds between its N1 and the amino proton of G18, and between its amino proton to the O2 of C2 (Figure 5(b)). Subsequently, A21 serves as a platform for consecutive stacking of all residues up to U27, with the exception of C24. A23 and A25 are stacked, evidenced by NOEs from the H2 of A23 to the H2 and H1′ of A25, while C24 is bulged. The dynamic properties

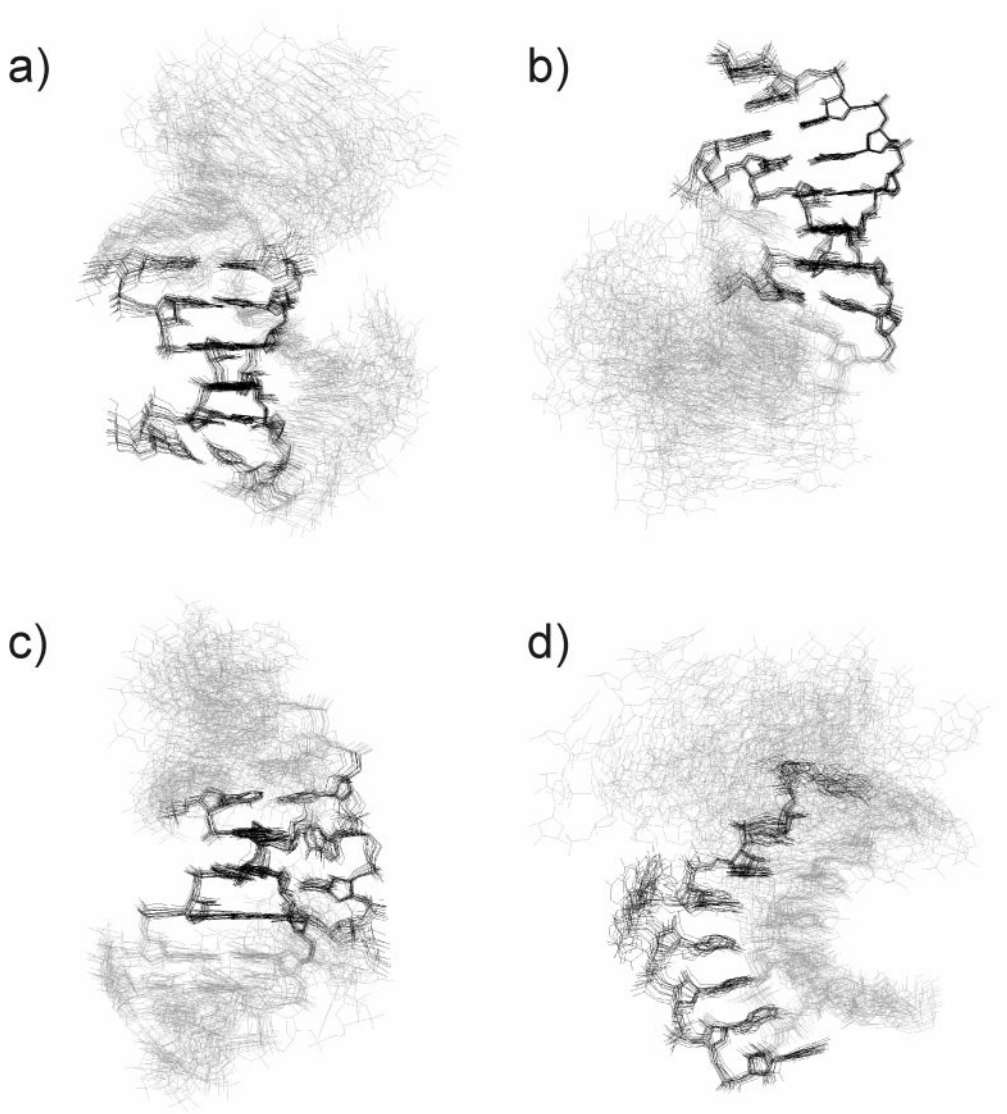

Figure 4. Heavy atom superpositions of the 15 lowest energy structures fitted to: (a) S1 residues 1-6/14-19 (rmsd = 0.5 Å); (b) S2 residues 8-13/29-34 (rmsd = 0.3 Å); (c) junction residues 13-16/4-6,29, 26-28 (rmsd = 0.6 Å); (d) L2 residues 20-28 (rmsd = 1.3 Å).

of C24 (visualized in Figure 4(d)) are transposed to the next residue and cause line broadening of the aromatic proton resonances of A25. However, in the calculated structures A25 is firmly held in position close to the minor groove connected by a strong NOE between its H2 and C17H1′ (Figure 2(a)). No hydrogen bonds can be detected between residues 22-25 and S1, so it appears that the structure of this part of L2 is primarily formed by stacking interactions between the adenine bases.

At the 3′-side of the loop two base triples are present. Adenine 26 is stacked upon residue 25 and forms a base triple with G4-C16 in S1 (Figure 5(c)). Additionally, a single ribose zipper motif[31] is present between A26 and C16, where the 2′-hydroxyl group and N3 of A26 share a hydrogen bond with the 2′-hydroxyl group of residue C16. The other base triple is formed between U27 and G15-C5, stabilized by a hydrogen bond between U27 O2 and the amino group of G15 (Figure 5(d)).

The large helical twist between S1 and S2 at the junction together with an S-turn at the 3′-side of L2 allow the three strands of S1 and L2 to come in close proximity at the junction and position the 3′-side of L2 in the minor groove for forming the base triples (Figure 6(a)). A28 fulfils a key role in maintaining the structural arrangement at the junction. The S-turn is held in place by the rather unusual position of the A28 base (Figure 6(b)), which is fixed perpendicular to the plane of the canonical base-pairs in the stem region, stabilized by hydrogen bonds between the N1 and the amino protons of A28 and the 2′-hydroxyl group of G14 (Figure 6(a)). The ribose moiety of A28 is tilted by

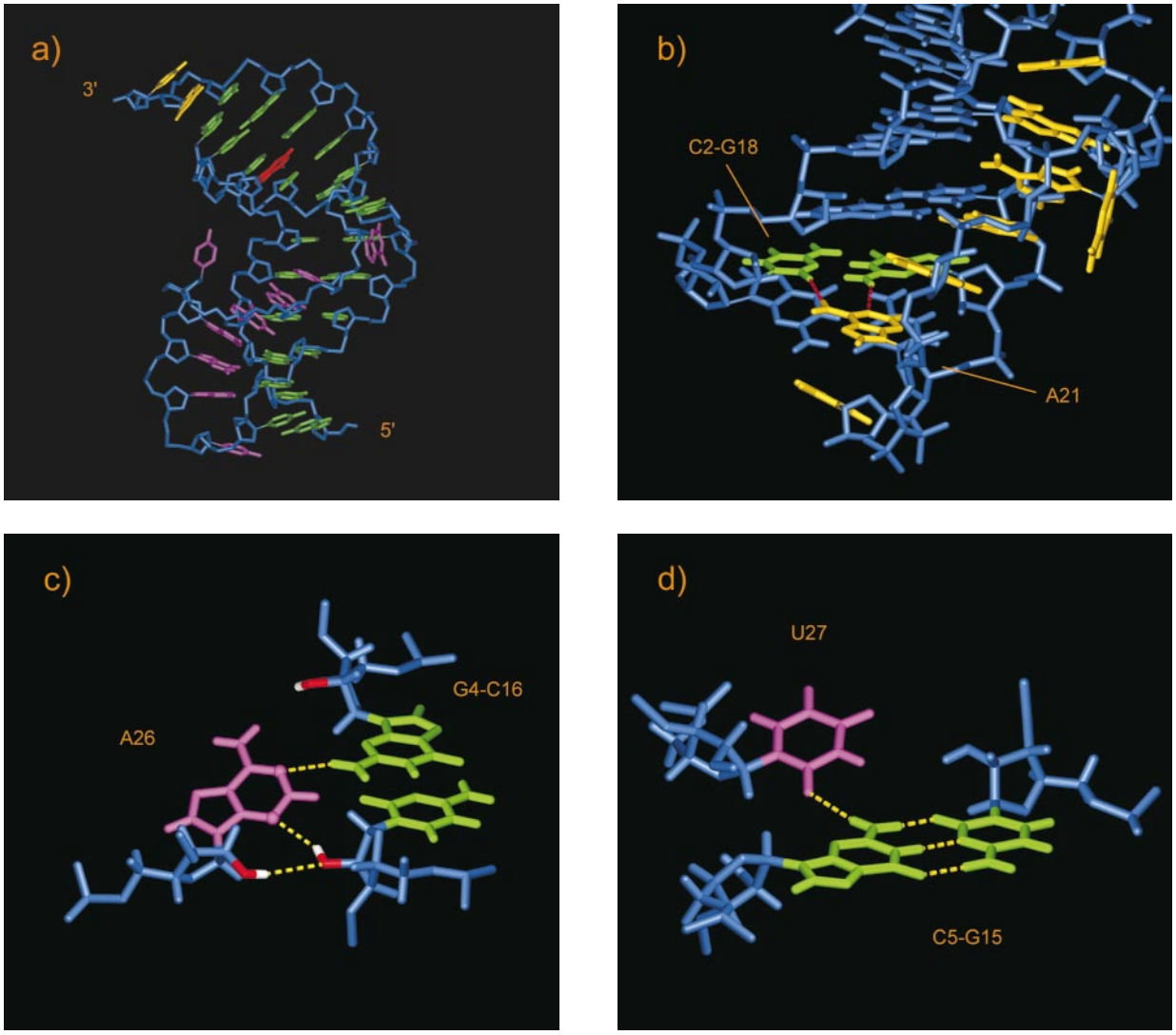

Figure 5. (a) Lowest energy structure of the SRV-1-derived pseudoknot causing ribosomal frameshifting. The bases of the two stems involved in Watson-Crick base-pairs are coloured green, the adenine comprising loop L1 red, the bases of loop L2 magenta and the two 3′-stacking bases are coloured yellow. (b) Detailed view of the 5′-side of the SRV-1 pseudoknot with the bases C20-U27 of loop L2 coloured in yellow. Adenine 21 has two hydrogen bonds to the second base-pair in stem S1 (coloured green), indicated by the broken red lines, and serves as a platform for consecutive stacking of the following adenine residues. (c) Detailed view of the triple interaction between A26 of loop L2 and the G4-C16 base-pair from stem S1. Three hydrogen bonds are present, indicated by the broken yellow lines. A26 is directed into the minor groove and forms a hydrogen bond by its N1 and the amino proton of G4. A ribose zipper motif was observed between the other two hydrogen bonds from the N3 and the hydroxyl group of A26 to the hydroxyl group of C16. (d) Triple base-pair formed by U27.C5-G15. Uridine 27 is held in the minor groove by a hydrogen bond between O2 and the amino proton of guanine 15.

about 90°, bringing the H1′ of A28 and U29 in close proximity (Figure 6(a)). Figure 6(a) also clearly shows that the sharp turn between A28 and U29 and the tilted ribose of A28 positions U29 in such a way that it can base-pair with A13 in a canonical manner without disturbing the base triple below. Finally, the phosphate group between A28 and U29 is in close proximity to the 2′-hydroxyl group of C6 in the opposite strand, which could possibly form a hydrogen bond to stabilize the unusual conformation of the backbone around A28.

Discussion

Structural comparison

This work presents, after elucidation of the TYMV[20] and BWYV[21] pseudoknots, the third structure of an *H*-type or classical pseudoknot (for a recent review discussing nomenclature of pseudoknots see Hilbers *et al.*[16]) solved at sufficient resolution to provide detailed structural information. The structure of the frameshifting RNA pseudoknot of SRV-1 exhibits the typical topology seen for other previously studied *H*-type pseudoknot

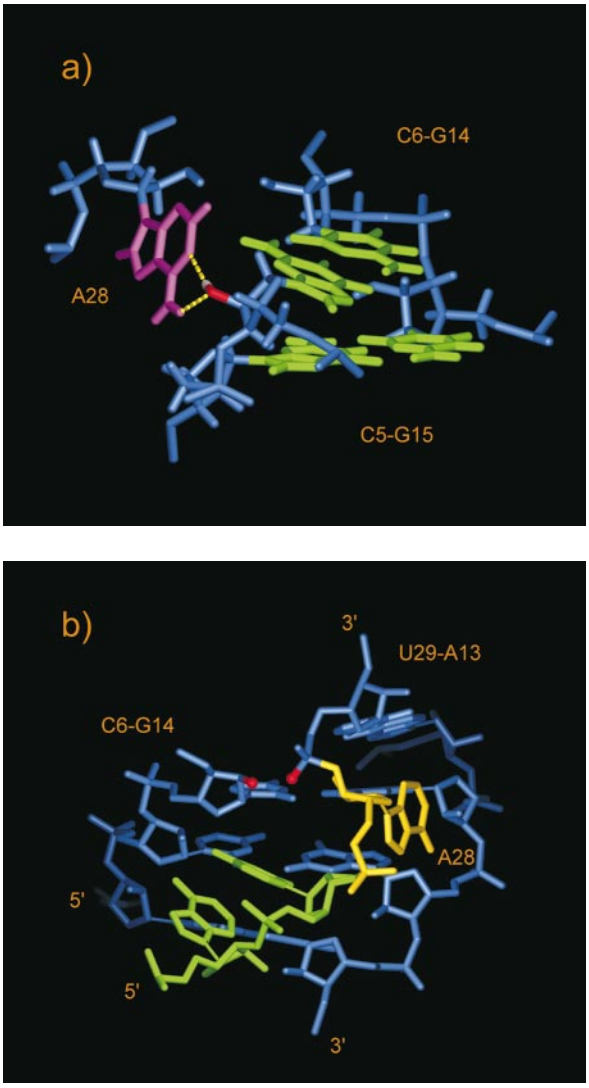

Figure 6. (a) Adenine 28 (coloured magenta) is oriented perpendicular to the plane of base-pairs and is hydrogen bonded by N1 and the amino proton to the hydroxyl group of guanine 14. (b) Representation of the heavy atoms of the junctional region (G4-C6, A13-C16, A26-U29). The tilted ribose of adenine 28 (coloured yellow) changes the direction of the backbone in such a way that uridine 29 can be base-paired with adenine 13. The two oxygen atoms in red might be involved in a water-mediated or direct hydrogen bond.

structures. At the local structural level, however, there are a number of interesting differences and similarities which we will elaborate below.

The present SRV-1 pseudoknot structure exhibits an unexpectedly compact conformation with several tertiary interactions, mainly at the S1-L2 interface (Figure 4) similar to what has been observed for the TYMV and BWYV pseudoknots. Most of the interactions between S1 and L2 consist of base triple formation between adenines in the loop and the bases of the base-pairs in the stem. We want to mention the one exception, namely the A26:G4-C16 base triple, which is stabilized by a single base zipper motif (Figure 5(c)). Intramolecular ribose zippers, unique to RNA since hydroxyl groups are involved, have previously been observed in the P4-P6 domain of group I intron ribozymes,[31] the HDV ribozyme[32] and the L11-binding domain of 23 S rRNA.[33,34] In these structures at least two pairs of riboses interact by hydrogen bonding to form the ribose zippers, which are characterized by shared hydrogen bonds between the 2′-hydroxyl group and a pyrimidine O2 (or purine N3) of one nucleotide and the 2′-hydroxyl of its partner. In the SRV-1 pseudoknot the riboses of the next residues in the stack, U27 and G15, are not participating in the ribose zipper. Therefore, in a more general sense the hydrogen-bonding pattern found here for a single ribose pair might be considered the signature of a ribose zipper motif.

A pertinent finding in this work is that the base-pairs predicted in the secondary structure model of the SRV-1 pseudoknot[35] are formed in the three-dimensional structure. Different views, however, have been articulated for the presence of the U29-

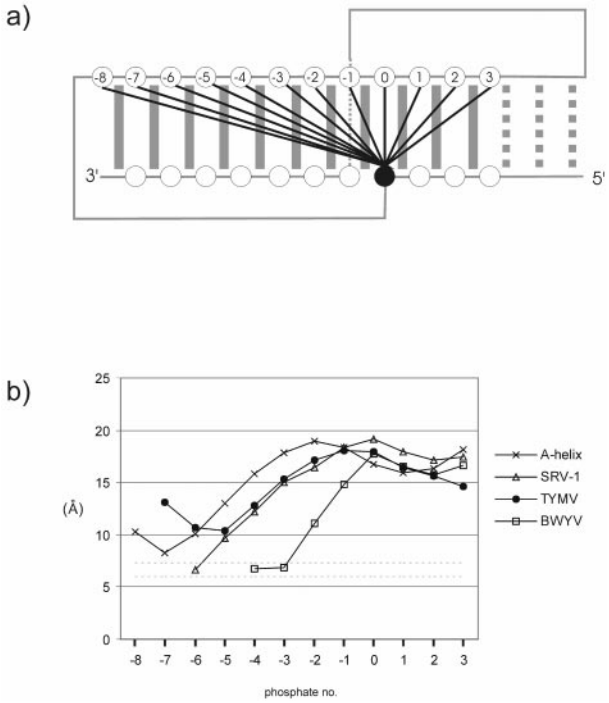

Figure 7. Interstrand phosphorus-phosphorus distances in *A*-form RNA[63] and pseudoknots. The curves (b) represent the distances between the sampling phosphate in one strand (indicated by the black dot in (a)) and the phosphates (numbered from −8 to 3 in (a) and (b)) on the other strand. For the *A*-form helix the choice of the sampling phosphate is arbitrary, for the pseudoknots the sampling phosphate is 5′ to the first L1 nucleoside. The broken lines in (b) indicate 6.0 and 7.4 Å, the sequential phosphate-phosphate distance in a regular *A*-form helix and the maximal distance that can be spanned by one nucleotide, respectively.

A13 base-pair at the junction. In the Mouse Mammary Tumor Virus (MMTV) pseudoknot an unpaired adenine was found at the junction, leading to a bent pseudoknot, which has been proposed to be an essential structural feature for frameshifting.[36,37] The MMTV pseudoknot and variants thereof with an intercalated adenine base at the junction showed frameshift efficiencies of about 20 %. For variants without an intercalated adenine the frameshift efficiencies were much lower (2 %). Functional data[38] led to the proposal that the SRV-1 pseudoknot would, at the dispense of the U29-A13 base-pair, also contain an unpaired intercalating adenine at the junction to create this characteristic bent conformation.[37] Mutational studies[39] of the SRV-1 pseudoknot in which the uridine had been changed into all other three nucleotides without affecting frameshift efficiency further supported this hypothesis. In our work, however, we firmly assigned a hydrogen bonded U29 imino proton and unambiguously established the presence of a closed U29-A13 base-pair at the junction. Earlier work by Hoffman and co-workers, who were able to assign the U29 imino proton, also indicated the presence of this base-pair.[40] Thus, an unpaired, intercalated nucleotide at the junction is not necessarily required for efficient frameshifting of the SRV-1 pseudoknot.

Geometrical aspects of the classical pseudoknot architecture

In the *H*-type pseudoknots only a small number of nucleotides, i.e. one or two, span the major groove of S2. The number of base-pairs that can be bridged in this way has been a matter of debate. For instance, NMR studies of the pseudoknot in gene 32 mRNA of bacteriophage T2 showed that a single nucleotide can span the major groove of a six or seven base-pair stem, but not five or eight.[41] On the other hand, it has been observed that in the BWYV pseudoknot only one nucleotide spans the major groove of S2, which effectively consists of four base-pairs. Using the SRV-1, TYMV and BWYV structures we now further elaborate on these considerations.

To this end the interstrand phosphorous-phosphorous distances from the 3′-end of S1, defined as the sampling phosphate (Figure 7(a)), to the opposite strand in S2 are compared with the distances expected for a continuous *A*-type RNA double helix (Figure 7(b)). This graph shows that for canonical *A*-type RNA the shortest phosphorous-phosphorous distance (~8 Å) occurs after a build-up with seven base-pairs, thus leaving a gap that can be bridged by one (or two) nucleotides. The distances measured in the TYMV and SRV-1 pseudoknots nearly coincide. In the SRV-1 pseudoknot a distance of 6.8 Å is reached when S2 is six base-pairs long. This distance can easily be bridged by one nucleotide. A similar situation obtains in the TYMV pseudoknot. Here S2 is five base-pairs long and the remaining gap, 11 Å, can be spanned by two nucleotides, the effective length of L1. In the BWYV pseudoknot the situation is significantly different. Here the phosphorous atoms of the third and fourth base-pair of S2 are already at a distance of 6.5 Å from the sampling phosphate. Again this distance can be, and in fact is spanned by one nucleotide. Below it will be discussed that the twist angle between the two base-pairs at the pseudoknot junction is an essential element in imposing these differences.

The helical twist between the base-pairs in an idealized *A*-type RNA helix is 32.7°. Compared to this value large helical twists at the junction between S1 and S2 seem typical for the formation of classic pseudoknots. Thus, for the SRV-1 pseudoknot a helical twist of 49° is observed between the two base-pairs at the junction. In the TYMV-pseudoknot[16] this twist amounts to 52° and for the BWYV pseudoknot[21] a twist of about 90° has been deduced.[16] The large helical twist in these pseudoknots serves two purposes. First, it facilitates bridging of the major groove of S2 by a small number (one or two) of nucleotides. Secondly, it opens the minor groove at the junction to accommodate base triple formation in the TYMV[20,16] and the SRV-1 pseudoknots (Figure 4) or base quadruple formation in the BWYV pseudoknot.[21] Formation of the base quadruple in the latter molecule also accounts for the large difference with the proposed secondary structure model.

Using a very simple model it can be demonstrated that overwinding the helical twist of one of the base-pairs can induce a helical bend. This is shown in Figure 8(a), where a cylinder serving as a model for an RNA helix is intersected by a plane at a certain angle representing the tilt of the base-pair. Subsequently, the upper part of the helix (cylinder) is rotated by 180° around an axis perpendicular to the plane through the mentioned base-pair, taking the point where the helix axis crosses the plane as the pivot point. If one takes the base-pair tilt equal to 45° it can be envisaged easily that the combined large twist and tilt angles lead to a bend of 90° between the upper and lower part of the helix. It is also clear from this example that smaller tilts will result in smaller bends.

At this point we like to mention that, interestingly, superposition of the pseudoknot structures on that of a regular *A*-type helix shows that in the pseudoknots the deviations of the local as well as the global helix axes from those in the regular *A*-type helix are not as outspoken as suggested in earlier publications. For this one has to realize that the number of base-pairs in the helical stems, S1 and S2, in the pseudoknots do not entail a complete helical turn, suggesting that the reported bends refer to local instead of global bends. To appreciate the importance of this remark a comparison with the situation in an *A*-type helix is enlightening. If one takes a perfect *A*-form helix of 12 base-pairs the apparent bend between the helix axis of the upper and lower 6 bp stems, generated by either programme Curves[42,43] or NEWHEL93

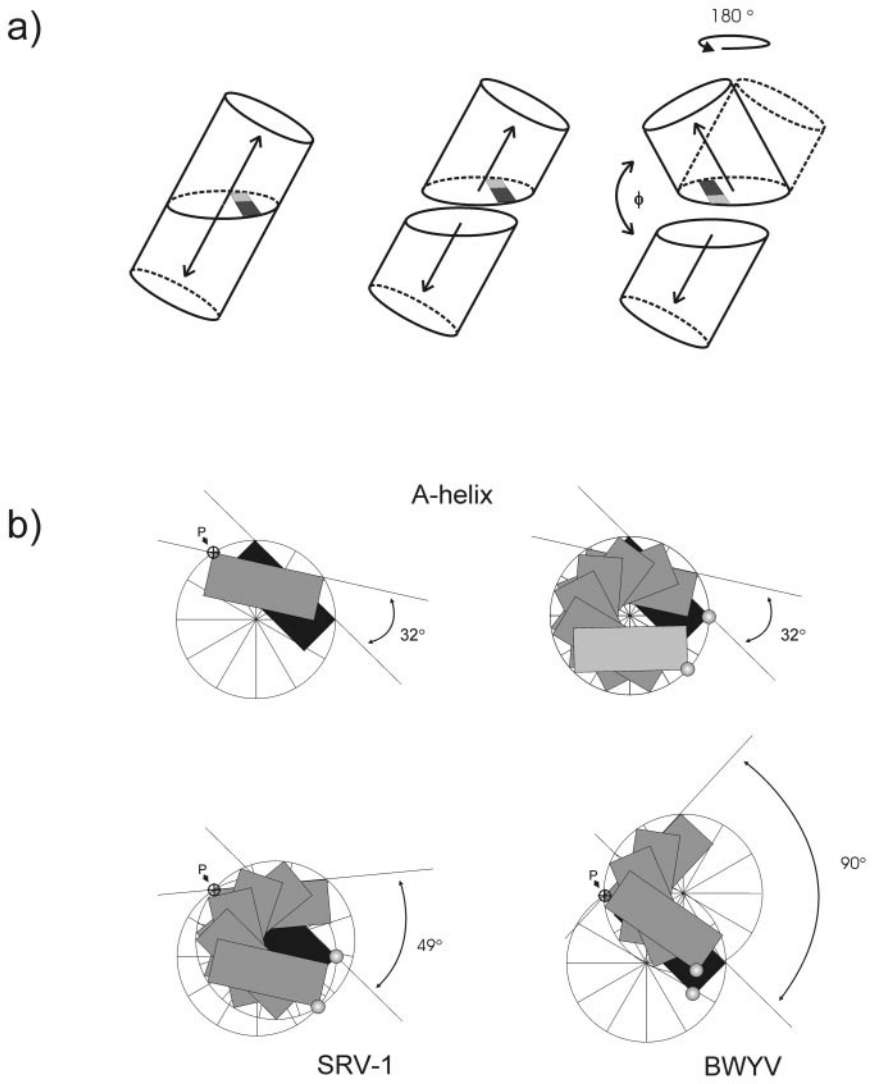

Figure 8. (a) Schematic representation of creating a helical bend (ϕ) by overwinding the helical twist of one base-pair step to 180°. (b) Schematics of base stacking in helical *A*-form RNA and SRV-1 and BWYV pseudoknots. Base-pairs are represented by rectangles with the C1′ of standard Watson-Crick base-pairs at the corners touching the sugar-phosphate backbone, indicated by the circle. Helical twists are indicated as the angle between adjacent C1′-C1′ vectors across base-pairs.

(R.E. Dickerson, University of California, Los Angeles), amounts to ~22°. Thus, examination of two stacked helical *A*-type stems smaller than one helical turn by either these programmes or by eye inspection may lead to the conclusion that the molecule is bent while actually the global helix is straight. This does not mean that in pseudoknot structures no bending of the global helix axis occurs at all. It is, however, in our opinion rather modest.

The overwinding with the pivot point located at the centre of the helix is, however, not possible in practice. It would require a lengthening of the phosphorous-phosphorous distance, which is prohibited. In fact, variation of the distance between two consecutive phosphorous atoms in the same strand is only possible within small margins. In nucleic acid structures this is reflected in the presence of almost constant phosphorous-phosphorous distances.[44] This means that in reality the rotation axis has to be moved close to the cylinder wall. Then rotation of the upper part of the helix will leave the orientation of its helix axis unaltered with respect to the same rotation in the situation above. However, this rotation will now result in a shift of the helical centre of the upper with respect to that of the lower helix in the intersecting plane. This is illustrated in Figure 8(b), where the effect of overwinding to a helical twist of 49° and 90° in the SRV-1 and BWYV pseudoknot, respectively, is compared with the situation in a normal *A*-type helix. In Figure 8(b) the black rectangle represents the last base-pair of the lower helix, and the grey shaded rectangles represent base-pairs in the upper helix stacked upon the last base-pair of the lower helix. Build up of the upper helix in the normal

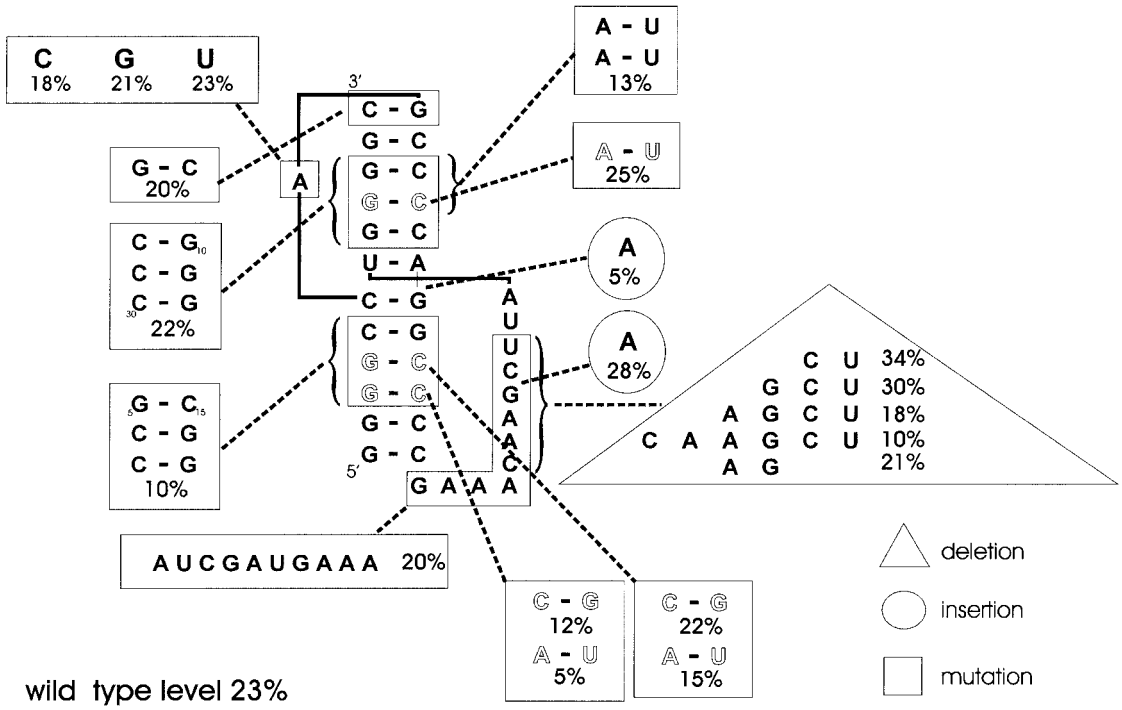

Figure 9. Frameshift efficiencies of selected SRV-1 pseudoknot mutants.[17,18]

A-type helix by seven base-pairs brings, in this two-dimensional projection, the 5′ terminus of the upper stem, indicated by the little sphere attached to the upper base-pair, close to the 3′ terminus of the lower stem, indicated by the little sphere attached to the black base-pair. This results in a distance of ~8 Å between these terminal phosphates (cf Figure 8(b)). In the SRV-1 pseudoknot the twist angle between the upper base-pair of the lower helix (stem S1) and the lowest base-pair in the upper helix (stem S2) amounts to 49°, giving rise to a shift of the helix centre of the upper helix, as can be seen from the shift of the upper helix with respect to that of the lower helix. Note that the rotation is around the pivot point P located at the crossing of the two circumferences, leaving the distance of the consecutive phosphorous atoms of the upper base-pair of stem S1 and the lowest base-pair of stem S2 virtually unaltered. A build up of the upper helix by six base-pairs brings its 5′ terminus as close to the 3′ terminus of the lower stem as the similar build up by seven base-pairs in the *A*-type helix. This clearly underscores the proposition that increasing the twist angle shortens the distance between the two termini in nice correspondence with the results depicted in Figure 7(b). The effect is even more dramatic in the BWYV pseudoknot. A twist angle of 90° between the base-pairs at the pseudoknot junction leads to a much larger shift between the centres of the helices, leading to a short distance of the free termini of the upper and lower stems after a build up with three base-pairs and the stacking adenine. Again this nicely conforms to the experimental results in Figure 7(b).

Comparison with frameshift efficiencies

The structural features of the SRV-1 pseudoknot nicely fit the results of mutational studies on frameshift efficiencies and provide leads for further investigations. The absence of interactions between the single L1 residue A7 and the major groove of S2 strongly suggests that neither the identity of the crossing nucleotide nor the base-pair composition of S2 is important for frameshifting. This fully agrees with the various SRV-1 pseudoknots mutated in this region, that only show slight changes in frameshift efficiencies (Figure 9).[18] The only known exception, i.e. a drop of 13% upon changing two G-C pairs into A-U pairs, supports the notion that base-pair composition of S2 is not important provided that stability is maintained. The intricate interactions at the S1-S2 junction and at the L2-S1 interface suggest that in this part of the pseudoknot the nucleotide composition is much more important. Again this concurs with the various mutational studies performed. The wild-type sequence at the 3′-side of L2, CUUA, is apparently suboptimal in terms of efficiency, since deletion of two (CU) or three (GCU) residues from L2 was shown to increase frameshift efficiency to a level of 34% and 30%, respectively.[18] This suggests that the interactions at the L2-S1 interface, i.e. base stacking, base-triples and the ribose zipper motif, are probably important elements in the mechanism of ribosomal frameshifting. In other words, absence

of the well-defined structure in the wild-type pseudoknot at the junction might explain the decrease in frameshift efficiency. When AG, which does not show direct contacts with S1, was deleted from the loop, frameshift efficiency was slightly less (21 %) than the wild-type level (23 %). In the case of deletion of CU, a guanosine is left between A26 and U27. Here the higher frameshift efficiency might be due to additional stacking interactions of the guanosine between A26 and U27 and a triple interaction with one of the base-pairs in S1. Further shortening of L2 progressively diminishes frameshift efficiency, which can be the result of disruption of the prime interactions at the junction by tension of a too short L2. When a block of three base-pairs in S1 is inverted, the frameshifting efficiency drops to 10 %,[17] which can be attributed to the loss of two triple interactions. For our NMR studies we introduced a single base-pair inversion at G2-C18. In the NMR structure the sharp turn of L2 at the base of S1 is held by a hydrogen bond between the N1 and the amino proton of A21 and the amino proton of G18 and the O2 of C2, respectively. Although it can be envisaged that in the wild-type pseudoknot a similar hydrogen bond can be formed by shifting A21 slightly over in the L2 minor groove, this minor change might decrease the overall frameshift efficiency to the respectable value of 16 %.

In summary, nearly all the results of the mutational studies can be explained by our structure. However, the single base-pair inversion of G3-C17 decreases the frameshift efficiency to 12 %,[18] while this base-pair is not involved in any interaction in our structure. This clearly shows the need for further structural and functional studies to fully comprehend the importance of the structural details of the SRV-1 pseudoknot. In particular, it will be very important to see whether the tertiary interactions at the junction as found in the pseudoknot structure of the SRV-1 derivative presented here are also present in the wild-type pseudoknot. Also, to date the base triples and other interactions at the junction have not been put to a functional test yet.

The importance of adenine bases

It is clear from the structures of the SRV-1 pseudoknot, presented here, and the BWYV pseudoknot,[21] that adenine bases can be ascribed a special role in ribosomal frameshifting. Insertion of an adenine between G26.1 and C26.2 in L2 increases frameshift efficiency to 28 %, a result for which no explanation could be given before.[18] Following earlier arguments, this particular adenine might also be involved in stacking and base triple interactions, giving rise to the remarkable high frameshift frequency. Alternatively, the changes in frameshift efficiency upon deletion or insertion of nucleotides in L2 could possibly arise from an optimization of the loop length as well as of the adenine content. A similar result has been observed in BWYV, where the insertion of an adenine residue after the first residue in L2 (G19A19a) increased frameshift efficiency by 70 %.[45]

It would be interesting to investigate the effect on the L2 backbone conformation and frameshift efficiency following replacement of adenine 28 by a pyrimidine. A purine at position 28 (Figure 6) seems important, since the single aromatic ring of a pyrimidine will probably be too far away to form a hydrogen bond with the 2′-hydroxyl group of G14. In line with this observation Brierley and coworkers[19] found for an infectious bronchitis virus (IBV) derived pseudoknot an adenine at the 3′-end of an 8 nt L2 essential for efficient ribosomal frameshifting. By increasing the size of the loop to 14 nt, the specific requirement of this adenine is lost. However, the IBV pseudoknot is structurally similar to the MMTV frameshifting pseudoknot with two nucleotides in L1 and an intercalating adenine at the junction. Since our structure of the SRV-1 pseudoknot is distinctly different from the MMTV pseudoknot[36] it is not evident that when the three bases (GCU), deleted from loop L2 in SRV-1, are reinserted, the typical structural features and requirement of adenine 28 is lost in SRV-1 as well.

Various mutational studies conducted on MMTV[46,38] and IBV[19] pseudoknots, compared with the SRV-1 mutation studies [17,18] have led to the impression that the structure of the SRV-1 pseudoknot would differ from the minimal IBV-derived pseudoknot, TYMV or BWYV structures.[19,47] In particular the absence of an effect of the available L2 mutations on SRV-1 frameshift efficiency led to the suggestion that the SRV-1 pseudoknot would not contain L2-S1 interactions as shown for the TYMV and BWYV pseudoknots.[19,47] However, at the time of these studies frameshift data of critical mutants involving tertiary interactions at the stable junction were not available. Based upon our structural study we propose that the IBV-derived pkA-A and minimal IBV pseudoknot[19] have structures similar to the SRV-1 pseudoknot with similar L2-S1 interactions.

Implications for the ribosomal frameshift mechanism

The solution structure of the SRV-1 pseudoknot involved in ribosomal frameshifting presented here, has several special structural properties that seem very important for its function. Frameshifting appears to be a delicate mechanism that can be fine-tuned for any particular species in order to obtain the proper ratio between functional proteins needed in the life cycle of the virus. Base-pairing as obtained by S2, stacking of the bases in L2 and tertiary interactions between S1 and L2, seem to be used as a tuning device in order to regulate the ratio between the translational products of the overlapping reading frames. Beyond the *A*-type double helix, our poor knowledge on the structural properties of loops and bulges leaves us unable to explain the effect certain mutations have on the fra-

meshift efficiency. Therefore it is important to understand the structural effect of subtle changes in the loops, such as the change of a single adenine in the IBV-based pseudoknot.[19] Further structural analyses by either NMR or crystallography not only of the wild-type sequences, but also of mutant pseudoknots are called for.

Despite the large amount of structural and functional data the actual role of the pseudoknot structure in the mechanism of ribosomal frameshifting is still unclear (for a recent review see also Giedroc *et al.*).[47] Most confusing is the large number of pseudoknots differing in size, loop sizes and nucleotide composition, not to mention the hairpin structures and three-way junctions that have been found to induce ribosomal frameshifting as well. Therefore, the discovery of a common feature present in all types of pseudoknots might provide a basis for a better understanding. In fact the only general feature that has been recognized so far for the frameshifting pseudoknots is the definition of formation of a classic pseudoknot itself: a single-stranded region which folds back and forms base-pairs with residues in a hairpin loop. Therefore, as a consequence of the pseudoknot structure, the translating ribosome proceeding along the mRNA will approach a hairpin structure and a third strand connected to the loop upstream. Unwinding of the pseudoknot might be difficult because of the stability of S2 and the loop-helix interactions, thereby leaving time for the ribosome to pause and shift to the −1 frame. Alternatively the specific interactions between L2 and the minor groove of S1 as seen in the SRV-1 pseudoknot that lead to the formation of a triple helix raises the intriguing possibility that a triple helix connected by a loop might be the recognition signal for a translating ribosome to frameshift.

Materials and Methods

In vitro transcription and translation assay

Mutations made in the SRV-1 *gag-pro* frameshift signal discussed in this report were made starting from the pSF43 mutant previously described by ten Dam and coworkers.[18] Plasmids were purified with the Wizard plasmid DNA purification system and used in the Promega TNT SP6 quick coupled transcription/translation system according to the manufacturer's protocol. The products of translation of the ''in frame'' and −1 reading frame, are a 19 kDa and 22 kDa fragment in size and contain ten and 11 methionine residues, respectively. Synthesised proteins were separated on a 15% (w/v) SDS-polyacrylamide gel. The relative amounts of the proteins from both reading frames were quantified by determining the incorporation of [^{35}S]methionine by scanning the gel with a Biorad phospho-imager correcting for the background and the differential methionine content of the products.

Sample preparation

An unlabelled and four different residue-type 36-mer RNA molecules specifically labelled with $^{13}C/^{15}N$ were prepared enzymatically by *in vitro* transcription using T7 RNA polymerase[48] with a DNA template. The products were purified using preparative polyacrylamide gel electrophoresis and recovered by electroelution. Purified RNA was extensively dialysed against 100 mM NaCl, 10 mM $KHPO_4/KH_2PO_4$ buffer (pH 6.4) and concentrated using Centricon microconcentrators. The final concentration of the unlabelled sample was 3 mM and 1-2 mM for the residue-specific $^{13}C/^{15}N$-labelled RNA samples. Addition of 6 mM magnesium did not alter the chemical shifts or NOE patterns, indicating that the presence of magnesium does not influence the pseudoknot structure.

NMR methods

All spectra were recorded on Varian Inova 500 MHz and 750 MHz and Bruker DMX 600 MHz spectrometers, at 10° C for samples in H_2O and at 30 °C for those in ^2H_2O. Exchangeable protons were assigned from 2D NOESY experiments[49] which were recorded using a jump return pulse sequence[50] for water suppression or watergate pulse sequence with a water flip-back pulse.[51] Non-exchangable protons were assigned using 2D NOESY spectra employing mixing times of 50, 100, 200 and 300 ms, and from DQF-COSY[52] and TOCSY[53] experiments. Phosphorus chemical shifts were assigned by a 2D ^1H-^{31}P Hetero-TOCSY-NOESY[54] and a ^1H-^{31}P HETCOR experiment.[55]

Four samples with the different residue types separately $^{13}C/^{15}N$ labelled were used to separate overlapping resonances and ^1H- ^{13}C HMQC,[56] ^1H- ^{13}C 3D NOESY-HMQC[57] and HCCH-TOCSY spectra[58,26] were recorded.

Structural restraints

Proton distances involving non-exchangeable protons were predominantly estimated from 100, 200, and 300 ms 2D NOESYs as well as a 300 ms 3D NOESY-HMQC. NOE cross-peak volumes involving non-exchangeable protons from the 100 ms NOESY were converted into distance restraints using the approach as described by Barsukov and Lian,[59] which qualitatively takes care of spin diffusion, by using the H5-H6 NOEs (2.45 Å) for short and intraresidue H1′-H6/H8 NOEs of helical residues for longer distances (3.65 Å on average) for those cross-peaks that were only visible at longer mixing times. By this method, the distances that are derived from the 100 ms spectrum agree very well with those from longer mixing times, and the possibility of large errors exceeding the error bandwidth can be virtually excluded. Initially we made use of ±20% error bounds for distances derived from the 100 ms NOESY experiment, because in our experience this gives a higher convergence rate, while the number of violations is kept to a minimum. Additional tests, using larger (−30%/+40%) error bounds resulted in structures with virtually identical overall and local folds, albeit with slightly larger overall rmsds (1.84 Å) and number of violations. Small NOEs taken from the 300 ms 3D NOESY-HMQC were subdivided in four classes strong (1.8-3.0 Å), medium (2.0-4.0 Å), and weak (3.0-5.0 Å). Small NOEs from the 200 ms NOESY were classified as very weak (4.0-6.0 Å). As spin diffusion might become important for a molecule of this size, the upper bound for very weak NOEs was set to 6 Å, which is a rather conservative estimate. Distance bounds for partially overlapping peaks were set to 1.8-6.0 Å. Non-NOEs (i.e. lower bounds of

Table 1. Structural statistics of the final ensemble of 15 structures

Distance restraints	
Intranucleotide	294
Internucleotide	171
Hydrogen bonds	34
Non-NOE	48
Rms deviations (Å, °)	
Distance restraints (547)	0.075 ± 0.002
Dihedral restraints (272)	1.203 ± 0.050
Rms deviation from idealized geometry	
Bonds (Å)	0.0125 ± 0.0004
Angles (°)	1.998 ± 0.092
Impropers (°)	0.471 ± 0.023
Restraint violations	
Number of distance violations[a] >0.4 Å	1 ± 1
Number of dihedral violations[b] >4°	8 ± 2
Atomic rms deviations (Å)[c]	1.72 ± 0.41

[a] None larger than 0.5 Å.
[b] None larger than 6°.
[c] With respect to the average structure.

4.5-5 Å between a given proton pair) were put in at later stages of the structure calculations when the distance between a given proton pair came out smaller than 5 Å, while there was clearly no cross-peak visible in any of the NOE spectra. The given lower bound is very conservative, and had no influence on the calculated structures, but merely helps defining the allowed conformational space during the molecular dynamics runs. Non-NOEs were only included for residues for which we established separately that they are not involved in intermediate exchange. Scalar couplings between H1′ and H2′ were determined from a DQF-COSY and a NOESY experiment, and were converted to N- or S-type sugar puckers (using dihedral restraints) in case of small (<3 Hz) or high (>7 Hz) values, respectively. The sugar puckers of A21, A22, A23, C24 and C36 were left unrestrained because intermediate scalar couplings of 4-6 Hz were measured for these residues. All residues had intra residual H1′/H2′/H3′ to H6/H8 NOE intensities that are indicative of an (-angle in the *anti* range, except for C24 and were therefore restrained 202(±30)° for the stem residues and 202(±60)° for the others.

Both stem regions adopt an *A*-helical conformation, evidenced by chemical shifts, sugar puckers and NOE connectivities. All residues within the stem regions were therefore assigned glycosidic and backbone torsion angles in accordance with canonical *A*-type RNA with error bounds of ±20° for those residues. Distances relating to imino protons were conservatively estimated from a 300 ms NOESY spectrum. Hydrogen bond restraints, according to Saenger[60] with error bounds of 0.1 Å, and planarity restraints were imposed for residues involved in Watson-Crick base-pairing, as evidenced by the imino proton spectra and all observed NOEs and chemical shifts. Hydrogen bond restraints for the two junctional base-pairs were loosened (i.e. error bounds were increased to 0.5 Å) in order obtain a more unbiased result in this region of the RNA molecule. No standard restraints were imposed for any of the backbone torsion angles between the sugar of C6 to G8, and that of C19 to U29, to ensure an unbiased course of the polynucleotide chain. For these domains, the β and γ torsion angles were left unrestrained, except for A29, for which we found evidence in the ^{1}H-^{31}P HETCOR experiment[55] to restrain β to 0-120°. The angle ε was set to 225(±60)°, in order to exclude the stereochemically forbidden *gauche*$^{+}$ region. No downfield shifted phosphorous chemical shifts were found for any of the pseudoknot residues and therefore the angles α and ζ were conservatively restrained to 0(±120)° to exclude the *trans* region.

Structure calculations

A set of 100 structures was calculated using the torsion angle dynamics (TAD) protocol[28] in X-PLOR.[27] The TAD cooling step in this protocol was increased to 90 ps so as to obtain a higher convergence rate. 15% of the obtained pool of structures had no distance and dihedral restraint violations larger than 0.5 Å and 6°, respectively. All structures of this ensemble, best fitting the experimental data, were selected for presentation and their statistics are listed in Table 1. Colour Figures were generated using MOLMOL.[61]

Coordinates

Coordinates for the set of 15 final structures as well as a full list of restraints used in X-PLOR have been deposited in the protein data bank, accession code 1E95.

Acknowledgements

The NMR experiments were performed at the SON Large Scale NMR facility (Nijmegen and Utrecht, The Netherlands). This research was supported by NWO-CW and NWO-ALW.

References

1. Power, M. D., Marx, P. A., Bryant, M. L., Gardner, M. B., Barr, P. J. & Luciw, P. A. (1986). Nucleotide sequence of SRV-1, a type D simian acquired Immune deficiency syndrome retrovirus. *Science,* **231**, 1567-1572.
2. Brierley, I., Digard, P. & Inglis, S. C. (1989). Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. *Cell,* **57**, 537-547.
3. Garcia, A., van Duin, J. & Pleij, C. W. A. (1993). Differential response to frameshift signals in eukaryotic and prokaryotic translational systems. *Nucl. Acids Res.* **21**, 401-406.
4. Tzeng, T.-H., Tu, C.-L. & Bruenn, J. A. (1992). Ribosomal frameshifting requires a pseudoknot in the *Saccharomyces cerevisiae* double-stranded RNA virus. *J. Virol.* **66**, 999-1006.
5. Dinman, J. D. & Wickner, R. B. (1992). Ribosomal frameshifting efficiency and gag/gag-pol ratio are critical for yeast M1 double stranded RNA virus propagation. *J. Virol.* **66**, 3669-3676.
6. Larsen, B., Wills, N. M., Gesteland, R. F. & Atkins, J. F. (1994). rRNA-mRNA base-pairing stimulates a programmed −1 ribosomal frameshift. *J. Bacteriol.* **176**, 6842-6851.
7. Gesteland, R. F. & Atkins, J. F. (1996). Recoding: dynamic reprogramming of translation. *Annu. Rev. Biochem.* **65**, 741-768.
8. Parkin, N. T., Chamorro, M. & Varmus, H. E. (1992). Human immunodeficiency virus type 1 gag-pol frameshifting is dependent on downstream

mRNA secondary structure: demonstration by expression *in vivo*. *J. Virol.* **66**, 5147-5151.
9. Tsuchihashi, Z. (1991). Translational frameshifting in the *Escherichia coli* dnaX gene *in vitro*. *Nucl. Acids Res.* **19**, 2457-2462.
10. Larsen, B., Gesteland, R. F. & Atkins, J. F. (1997). Structural probing and mutagenic analysis of the stem-loop required for *Escherichia coli dnaX* ribosomal frameshifting: programmed frameshifting of 50 %. *J. Mol. Biol.* **271**, 47-60.
11. Rettberg, C. C., Prère, M. F., Gesteland, R. F., Atkins, J. F. & Fayet, O. (1999). A three-way junction and constituent stem-loops as the stimulator for −1 frameshifting in bacterial insertion sequence IS911. *J. Mol. Biol.* **286**, 1365-1378.
12. Kang, H., Hines, J. V. & Tinoco, I., Jr (1996). Conformation of a non-frameshifting RNA pseudoknot from mouse mammary tumor virus. *J. Mol. Biol.* **259**, 135-147.
13. Somogyi, P., Jenner, A. J., Brierley, I. & Inglis, S. C. (1993). Ribosomal pausing during translation of an RNA pseudoknot. *Mol. Cell. Biol.* **13**, 6931-6940.
14. Pleij, C. W. A., Rietveld, K. & Bosch, L. (1985). A new principle of RNA folding based on pseudoknotting. *Nucl. Acids. Res.* **13**, 1717-1731.
15. ten Dam, E. B. & Pleij, C. W. A. (1992). Structural and functional aspects of RNA pseudoknots. *Biochemistry,* **31**, 11665-11676.
16. Hilbers, C. W., Michiels, P. J. A. & Heus, H. A. (1998). New developments in structure determination of pseudoknots. *Biopolymers,* **48**, 137-153.
17. ten Dam, E. B., Brierley, I., Ingles, S. & Pleij, C. W. A. (1994). Identification and analysis of the pseudoknot containing gag-pro ribosomal frameshift signal of simian retrovirus-1. *Nucl. Acids Res.* **22**, 2304-2310.
18. ten Dam, E. B., Verlaan, P. W. & Pleij, C. W. A. (1995). Analysis of the role of the pseudoknot in the SRV-1 gag-pro ribosomal frameshift signal: loop lengths and the stability of the stem regions. *RNA,* **1**, 146-154.
19. Liphardt, J., Napthine, S., Kontos, H. & Brierley, I. (1999). Evidence for an RNA pseudoknot loop-helix interaction essential for efficient −1 ribosomal frameshifting. *J. Mol. Biol.* **288**, 321-335.
20. Kolk, M. H., Van der Graaf, M., Wijmenga, S. S., Pleij, C. W. A., Heus, H. A. & Hilbers, C. W. (1998). NMR structure of a classical RNA pseudoknot: interplay of single- and double-stranded RNA. *Science,* **280**, 434-438.
21. Su, L., Chen, L., Egli, M., Berger, J. M. & Rich, A. (1999). Minor groove RNA triplex in the crystal structure of a ribosomal frameshifting viral pseudoknot. *Nature Struct. Biol.* **6**, 285-292.
22. Santa Lucia, J., Jr & Turner, D. H. (1993). Structure of $(rGGCGAGCC)_2$ in solution from NMR and restrained molecular dynamics. *Biochemistry,* **32**, 6026-6036.
23. Heus, H. A., Wijmenga, S. S., Hoppe, H. & Hilbers, C. W. (1997). The detailed structure of tandem G.A mismatched base-pair motifs in RNA duplexes is context dependent. *J. Mol. Biol.* **271**, 147-158.
24. Wijmenga, S. S., Mooren, M. W. & Hilbers, C. W. (1993). NMR of nucleic acids; from spectrum to structure. In *NMR of Macromolecules. A Practical Approach* (Roberts, G. C. K., ed.), pp. 217-288, Oxford University Press, New York.
25. Wijmenga, S. S. & van Buuren, B. N. M. (1998). The use of NMR methods for conformational studies of nucleic acids. *Prog. Nucl. Magn. Res. Spec.* **32**, 287-387.
26. Legault, P., Farmer, B. T., II, Mueller, L. & Pardi, A. (1994). Through-bond correlation of adenine protons in a ^{13}C-labeled ribozyme. *J. Am. Chem. Soc.* **117**, 11043-11048.
27. Brünger, A. T. (1992). *X-PLOR. A System for X-ray Crystallography and NMR*, Yale Univ. Press, New Haven.
28. Stein, E. G., Rice, L. M. & Brünger, A. T. (1997). Torsion-angle dynamics as a new efficient tool for NMR structure calculation. *J. Magn. Reson.* **124**, 154-164.
29. Kolk, M. H., Wijmenga Heus, H. A. & Hilbers, C. W. (1998). On the NMR structure determination of a 44n RNA pseudoknot: assignment strategies and derivation of torsion angle restraints. *J. Biomol. NMR,* **12**, 423-433.
30. Haasnoot, C. A. G., Hilbers, C. W., van der Marel, G. A., van Boom, J. H., Singh, U. C., Pattabiraman, N. & Kollman, P. A. (1986). On loopfolding in nucleic acid hairpin-type structures. *J. Biomol. Struct. Dynam.* **3**, 843-857.
31. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L. & Kundrot, C. E., *et al*. (1996). Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science,* **273**, 1678-1685.
32. Ferré-D'Amaré, A. R., Zhou, K. & Doudna, J. A. (1998). Crystal structure of a hepatitis delta virus ribozyme. *Nature,* **395**, 567-574.
33. Conn, G. L., Draper, D. E., Lattmann, E. E. & Gittis, A. G. (1999). Crystal structure of a conserved ribosomal protein-RNA complex. *Science,* **284**, 1171-1174.
34. Wimberly, B. T., Guymon, R., McCutcheon, J. P., White, S. W. & Ramakrishnan, V. (1999). A detailed view of a ribosomal active site: the structure of the L11-RNA complex. *Cell,* **97**, 491-502.
35. ten Dam, E. B., Pleij, C. W. A. & Bosch, L. (1990). RNA pseudoknots: translational frameshifting and readthrough on viral RNAs. *Virus Genes,* **4**, 121-136.
36. Shen, L. X. & Tinoco, I., Jr (1995). The structure of an RNA pseudoknot that causes efficient frameshifting in mouse mammary tumor virus. *J. Mol. Biol.* **247**, 963-978.
37. Chen, X., Kang, H., Shen, L. X., Chamorro, M., Varmus, H. E. & Tinoco, I., Jr (1996). A characteristic bent conformation of RNA pseudoknots promotes −1 frameshifting during translation of retrovial RNA. *J. Mol. Biol.* **260**, 479-483.
38. Chen, X., Chamorro, M., Lee, S. I., Shen, L. X., Hines, J. V., Tinoco, I., Jr & Varmus, H. E. (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.
39. Sung, D. & Kang, H. (1998). Mutational analysis of the RNA pseudoknot involved in efficient ribosomal frameshifting in simian retrovirus-1. *Nucl. Acids Res.* **26**, 1369-1372.
40. Du, Z., Holland, J. A., Hansen, M. R., Giedroc, D. P. & Hoffman, D. W. (1997). Base-pairings within the RNA pseudoknot associated with the simian retrovirus-1 gag-pro frameshift site. *J. Mol. Biol.* **270**, 464-470.
41. Du, Z. & Hoffman, D. W. (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. *Nucl. Acids Res.* **25**, 1130-1135.
42. Lavery, R. & Sklenar, H. (1988). The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dynam.* **6**, 63-91.
43. Lavery, R. & Sklenar, H. (1989). Defining the structure of irregular nucleic acids: conventions and principles. *J. Biomol. Struct. Dynam.* **6**, 655-667.
44. Hilbers, C. W., Heus, H. A., van Dongen, M. J. P. & Wijmenga, S. S. (1994). The hairpin elements of nucleic acid structure: DNA and RNA folding. In *Nucleic Acids and Molecular Biology* (Eckstein, F. & Lilley, D. M. J., eds), vol. 8, p. 56, Springer-Verlag, Heidelberg.
45. Kim, Y.-G., Su, L., O'Neill, A. & Rich, A. (1999). Specific mutations in a viral RNA pseudoknot drastically change ribosomal frameshifting efficiency. *Proc. Natl Acad. Sci. USA,* **96**, 14234-14239.
46. Chamorro, M., Parkin, N. & Varmus, H. E. (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.
47. Giedroc, D. P., Theimer, C. A. & Nixon, P. L. (2000). Structure, stability and function of RNA pseudoknots involved in stimulating ribosomal frameshifting. *J. Mol. Biol.* **298**, 167-185.
48. Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. (1987). Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucl. Acids Res.* **15**, 8783-8798.
49. Jeener, J., Meier, B. H., Bachmann, P. & Ernst, R. R. (1979). Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys.* **71**, 4546-4553.
50. Plateau, P. & Guéron, M. (1982). Exchangeable proton NMR without base-line distortion, using new strong-pulse sequences. *J. Am. Chem. Soc.* **104**, 7310-7311.
51. Lippens, G., Dhalluin, C. & Wieruszeski, J.-M. (1995). Use of a water flip-back pulse in the homonuclear NOESY experiment. *J. Biomol. NMR,* **5**, 327-331.
52. Shaka, A. J. & Freeman, R. (1983). Simplification of NMR spectra by filtration through multiple-quantum coherence. *J. Magn. Reson.* **51**, 169-173.
53. Griesinger, C., Otting, G., Wüthrich, K. & Ernst, R. R. (1988). Clean TOCSY for ^{1}H spin system identification in macromolecules. *J. Am. Chem. Soc.* **110**, 7870-7872.
54. Kellog, G. W. & Schweizer, B. I. (1993). Two- and three-dimensional ^{31}P-driven NMR procedures for complete assignment of backbone resonances in oligodeozynucleotides. *J. Biomol. NMR,* **3**, 577-595.
55. Sklenár, V., Miyashiro, H., Zon, G., Miles, H. T. & Bax, A. (1986). Assignment of the ^{31}P and ^{1}H resonances in oligonucleotides by two-dimensional NMR spectroscopy. *FEBS Letters,* **208**, 94-98.
56. Bax, A., Griffith, R. & Hawkins, B. L. (1983). Correlation of proton and nitrogen-15 chemical shifts by multiple quantum NMR. *J. Magn. Reson.* **55**, 301-315.
57. Ikura, M., Kay, L. E., Tschudin, R. & Bax, A. (1990). Three-dimensional NOESY-HMQC spectroscopy of a ^{13}C labeld protein. *J. Magn. Reson.* **86**, 204-209.
58. Bax, A., Clore, G. M. & Gronenborn, A. M. (1990). ^{1}H-^{1}H correlation *via* isotropic mixing of ^{13}C magnetization, a new three-dimensional approach for assigning ^{1}H and ^{13}C spectra of ^{13}C-enriched proteins. *J. Magn. Reson.* **88**, 425-431.
59. Barsukov, I. L. & Lian, L. Y. (1993). NMR of nucleic acids; from spectrum to structure. In *NMR of Macromolecules. A Practical Approach* (Roberts, G. C. K., ed.), pp. 315-357, Oxford University Press, New York.
60. Saenger, W. (1984). *Principles of Nucleic Acid Structure*, Springer-Verlag, New-York.
61. Koradi, R., Billeter, M. & Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51-55.
62. Güntert, P., Mumenthaler, C. & Wüthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* **273**, 283-298.
63. Schindelin, H., Zhang, M., Bald, R., Furste, J. P., Erdmann, V. A. & Heinemann, U. (1995). Crystal structure of an RNA dodecamer containing the *Escherichia coli* Shine-Dalgarno sequence. *J. Mol. Biol.* **249**, 595-603.

Edited by I. Tinoco

(Received 21 March 2001; received in revised form 24 May 2001; accepted 4 June 2001)