Minor groove RNA triplex in the crystal structure of a ribosomal frameshifting viral pseudoknot

Li Su¹, Liqing Chen¹,², Martin Egli³, James M. Berger⁴ and Alexander Rich¹

Many viruses regulate translation of polycistronic mRNA using a -1 ribosomal frameshift induced by an RNA pseudoknot. A pseudoknot has two stems that form a quasi-continuous helix and two connecting loops. A 1.6 Å crystal structure of the beet western yellow virus (BWYV) pseudoknot reveals rotation and a bend at the junction of the two stems. A loop base is inserted in the major groove of one stem with quadruple-base interactions. The second loop forms a new minor-groove triplex motif with the other stem, involving 2'-OH and triple-base interactions, as well as sodium ion coordination. Overall, the number of hydrogen bonds stabilizing the tertiary interactions exceeds the number involved in Watson–Crick base pairs. This structure will aid mechanistic analyses of ribosomal frameshifting.

In many viruses, including both tumor viruses and retroviruses, the programmed -1 ribosomal frameshifting of polycistronic mRNA regulates the relative level of structural and enzymatic proteins important for efficient viral assembly¹,². Because of the -1 shift in reading frames, the stop codon is avoided, and a single fusion protein is produced. For example, in the Rous sarcoma virus, the pol gene that encodes integrase, protease and reverse transcriptase is expressed with the upstream gag gene (encoding virus core proteins) through a gag-pol fusion protein. The mature products are later obtained by processing the poly-protein precursor¹. The -1 frameshifting is found not only in retroviruses³–⁵ but also in coronaviruses⁶, in yeast⁷ and in plant viruses⁸, as well as in bacterial systems¹. Frameshifting levels can range from 1% to over 30% in different systems to produce gene products in a functionally appropriate ratio. However, the mechanism of ribosomal frameshifting is not understood. It is postulated that a complex mRNA structure six to eight nucleotides downstream from the ‘slippery sequence’, in many cases a pseudoknot, leads to ribosomal pausing and minor groove of stem 1, respectively¹⁰. Many mutational studies have been performed on frameshifting pseudoknots to correlate structure and function⁴–⁷, and certain mutations at the junction of the two stems or loop regions can have deleterious effects on frameshifting efficiencies¹¹,¹². These findings suggest that there are special structural features involving tertiary interactions in a frameshifting pseudoknot that are important for function.

Several NMR studies of pseudoknots have been reported during the past few years¹³–¹⁸. Those pseudoknots active in ribosomal frameshifting revealed that the two base-paired stems are stacked upon each other¹⁴,¹⁶,¹⁷, and are frequently in a bent conformation¹⁴,¹⁷. In addition, loop nucleotides may be restrained by base-stacking interactions¹⁴. However, the junction of the two stems in frameshifting pseudoknots is not well defined, and very little is known about loop–stem interactions. Here we present a 1.6 Å resolution crystal structure of a 28-nucleotide pseudoknot from beet western yellow virus (BWYV), a plant luteovirus that is active in ribosomal frameshifting and regulates the production of an RNA-dependent RNA polymerase⁶. The structure reveals an RNA triplex formed by loop 2 in the minor groove of stem 1, and quadruple interactions between a loop 1 base and the major groove of stem 2. In addition, there is a marked rotation and bending at the junction of the two stems.

Overview of the structure

The crystallographic analysis of the structure is summarized in Methods. Data have been collected to 1.6 Å resolution and the structure refined to an R-factor of 20.7%, free R-factor of 25.4% (Table 1). Experimental electron density maps are also included (Fig. 1a,b). The overall structure is very compact (Fig. 2a), with van der Waals dimensions of approximately 32 Å x 36 Å x 22 Å. The first dimension is in the 5’ to 3’ direction.
An extended minor-groove RNA triplex

In several luteovirus frameshifting pseudoknots, the sequence of AACAAA (loop 2 A20–A24 and the continuing junctional base A25) is conserved. Loop 2 forms a stacking ladder with each adenine (or cytosine) rotated in somewhat different orientations to yield a variety of interactions with the minor groove of stem 1 (Fig. 3a). This extended RNA triplex, which has six layers of consecutive interactions in the minor groove, has not been previously seen and differs substantially from the nucleic acid triplexes with the third strand in the major groove. The triplex is RNA specific, as every interaction described here is contributed by an RNA base. A nucleic acid triplex with the third strand in the minor groove has six layers of consecutive interactions in the minor groove. The triple-base interactions are not coplanar in many cases. Several interactions involve the N1 and N6 of loop adenines, and N3 and N2 of guanines in the stem. These interactions with the third strand contribute to the large propeller twists in stem 1, and the gradual bend toward the major groove.

The first base of loop 2 (G19) projects away from stem 1 (Fig. 2a). However, the adjacent residue (A20) is 30° out of plane from the G4–C17 base pair (Fig. 2a), leading to a stabilizing stacking interaction with the ribose O4' of residue C5, as found in many nucleic acid structures (reviewed in ref. 22). This anchorlike adenosine spans the minor groove and interacts with both strands of stem 1. It is held in this position by an extensive network of seven hydrogen bonds (Fig. 3b), using both base and sugar interactions. The 2'-OH of A20 contributes four of the hydrogen bonds by interacting with two layers of bases (Table 2). Two of the hydrogen bonds are bifurcated, involving O2 atoms of residues C5 and C17, resulting in the tilt of these bases (Fig. 2c). Some of these adenosine interactions have also been seen in the crystal structures of the hammerhead ribozyme and group I intron. In both cases, the adenosine in an RNA tetraloop contacts the minor-groove side of a G–C base pair to form four of the seven hydrogen bonds.
observed here. In contrast to A20, the adenines there are not tilted and are largely in the plane of the base pair.

The next two residues A21 and C22 (Fig. 2a) are somewhat rotated away from the minor groove but the bases are stacked. The backbone conformation is extended and appears to be stabilized by hydrogen bonding between bases to the 2'-OH of G16 (Fig. 3c) and by a sodium ion that mediates base–base interactions (Fig. 5c). The adjacent A23 is stacked on the exocyclic amino group of C22 (Fig. 3a) and utilizes its Watson–Crick face to interact with the neighboring C15 (Fig. 3d). A24, the last base of loop 2, stacks on A23 (Fig. 3a) and interacts with guanosine G7 through three hydrogen bonds (Fig. 3e). This minor groove-specific interaction is unlike other G–A interactions described previously\textsuperscript{21–27}. The junction between A24 and A25 is stabilized by a 2.82 Å hydrogen bond between the 2'-OH of A24 and the O4' of A25 (not shown). With the exception of A20, loop 2 is not involved in lattice contacts.

The minor-groove triplex formed by loop 2 and stem 1 has a total of 16 hydrogen-bonding interactions; 7 to one strand of stem 1 and 9 to the other (Table 2). The extended backbone conformations in loop 2 are related to the interactions with stem 1 that switch from one strand to the other. For instance, P–P distances of 6.8–6.9 Å are found between A20–A21 and A23–A24, where the first base interacts with one strand of stem 1, and the next base interacts with the opposite strand (Figs 2a, 3). In addition, it is interesting that the temperature factors in the backbone of loop 2, in particular A21 through A25, are 60% higher than those in stem 1, suggesting increased mobility.

**The stem 1–stem 2 junction**

The first residue of stem 2, A25, stacks on the edge of the A24 base (Fig. 3a). There, it is tilted and interacts with bases in different layers (Figs 2a, 4). Rather than pairing with U13 as predicted (Fig. 2b), the A25 adenine forms three hydrogen bonds at the junction through its Watson–Crick face (Fig. 4b). The interactions with the O2 and 2'-OH of C14 are similar to those observed between A23 and C15 (Fig. 3d). On the strand opposite to A25 (Fig. 2a,c), the extruded U13 has a rotated and reversed sugar
articles

direction in relation to the adjacent nucleotides. This results in considerable compression of the backbone with a close U13–C14 P–P distance of 4.65 Å, which allows G12 to stack directly upon the base plane and furanose oxygen of C14 (Fig. 2c). Other bulge nucleotide conformations have been observed, although their sugar rotations are less marked28–30. In the crystal structure, U13 makes lattice contacts through stacking and base-pairing interactions. Symmetry-related bases that are bulged or unpaired, namely U13, G19 and G1, form a continuous stacking core of six layers. In addition, the U13 nucleotides in this stacking core use their Watson–Crick faces to form Hoogsteen base pairs with symmetry-related adenosine 20.

Quadruple interactions of loop 1
Most frameshifting pseudoknots have a short loop 1, usually one to two nucleotides in length4–6,8,11. In the BWYV pseudoknot, a C8–A9 sequence crosses the major groove, and the C residue is conserved in some luteovirus pseudoknots6. In the BWYV structure, C8 is inserted into the major groove of stem 2, forming a quadruple-base interaction with G12, A25 and C26 (Figs 2a,4a). This is the first example of one base hydrogen bonding to three other bases, two of which form a base pair. The G12–C26 base pair has a strong propeller twist, which is needed for C26 to stack upon the tilted A25 (Figs 2c,4a). As noted by the asterisk in Fig. 4a, the distance between N3 of C8 and O6 of G12 is 2.82 Å, and it is likely that C8 is protonated in the N3 position to form a hydrogen bond. Protonation of cytosines is also found in other nucleic acid structures31–33. Two water molecules stabilize this core interaction by mediating the 2'-OH of C8 to base contacts (A25 and C26) (Fig. 5b). Overall, these major groove interactions and the consecutive minor-groove triplex appear to drive the rotation and bend at the stem 1–stem 2 junction.

Sharp turns, ions and water molecule stabilization
In complex RNA structures, changes in backbone direction are achieved by sharp turns such as the U turn seen in tRNA3, the hammerhead ribozyme26,27 and group I intron24, which are stabilized by multiple interactions. In the BWYV structure, we have named the loop 1 to stem 2 transition a ‘C turn’ (Fig. 5a), which has some features in common with the U turn. The ‘C turn’ involves the change in direction between loop 1 and the top of stem 2, through three nucleotides (C8, A9 and C10), the last C being base paired. The turn is less sharp than the 180° U turn, but is stabilized by a hydrogen bond from N4 of C10 to the phosphate oxygen of A9, which is the second phosphate in the 5' direction. In comparison, the U turn uses a uracil N3 hydrogen-bonded to the fourth phosphate group in the 3' direction. In addition, adenine 9 is partially stacked at an angle on cytosine 10, which is associated with a parallel tilting of C10, propeller twisted in relation to its base-paired G28 (Fig. 2c). An organized network of water molecules facilitates the turn by bridging phosphate or base contacts (Fig. 5a). The array of water molecules on the other side of loop 1 stabilizes its conformation in the major groove (Fig. 5b).

A sharp turn is also found at the junction between stem 1 and loop 2 at G18-G19-A20 (Fig. 2a,d). This turn is facilitated by C2'-endo sugar puckers (the remainder of the RNA is C3'-endo, except for U13, which has both conformations) at residues G18
and G19, which lead to an extended backbone conformation. In contrast to the 'C turn', the only interaction at this turn is a self base (N3) to 2'-OH interaction at G19 fixing the base orientation.

RNA pseudoknots are generally stabilized from unfolding by magnesium ions\textsuperscript{34}. However, in many cases, high concentrations of monovalent ions can also stabilize the conformation\textsuperscript{34,35}. The sodium ion found in the minor groove is located at a key position to tie together stem 1 and loop 2, where a base–base contact is not possible because of the gap (Figs 2d,3c). Water molecules are also involved in stabilizing the conformation around the sodium ion (Fig. 5c). An octahedrally coordinated magnesium ion is bound to the triphosphate at the 5' end (Fig. 2d), coordinating to phosphate oxygens of the β- and γ-phosphate, the pro-S, oxygen of the G2 phosphate and three water molecules.

Relevance to frameshifting and RNA packing

The BWYV crystal structure provides a first detailed look at a frameshifting pseudoknot with ions and water molecules, and completely defines the interactions of loops. Pseudoknots are usually in some type of equilibrium between the folded and the unfolded form\textsuperscript{14}, and it is likely that frameshifting occurs when the pseudoknot does not unravel during translocation. Therefore, features that stabilize the compact form of the pseudoknot may be important for function. It is notable in this regard that the BWYV pseudoknot is stabilized by more hydrogen bonds involved in tertiary interactions (26 total, Table 2) than those engaged in Watson–Crick base pairs (24 total). As the pseudoknot barrier induces ribosomal pausing, it is likely that the ribosome contacts both the slippery site and the pseudoknot simultaneously\textsuperscript{16}. Continued activity of the mRNA translocation machinery probably induces deformation of the pseudoknot, transmitting the "tug" through stem 1 into loop 1. Understanding how the tertiary stabilized pseudoknot dynamically responds to deformation in this process may shed light on the -1 slippage mechanism.

As the ribosome travels along the messenger RNA in translation, the RNA triplex formed at the 5' end of the molecule is likely to be the first feature that the ribosome encounters\textsuperscript{16}, and the third strand may prevent the unraveling of stem 1 by cross-strand interactions. This feature would generally be absent in large stem–loop structures, which are known to delay ribosomal movement but are unable to induce frameshifting in pseudoknot systems\textsuperscript{61,11}. The triple-base contacts in the structure show a preference for adenosine residues in loop 2, and appear to be sequence specific. Mutations in other pseudoknots that invert base pairs in stem 1\textsuperscript{4,5,7} or eliminate the high adenosine content in loop 2 (ref. 12) are reported to have adverse effects on frameshifting. These mutations may disrupt triplex interactions and cause the pseudoknot to unravel.

A second stabilizing feature in the pseudoknot is the array of interactions at the stem 1–stem 2 junction. C8 seems to play an organizing role in assembling the junction between the stems. The junctional A25 base not only provides a continuous stacking platform into stem 2 but also plays an important role in stabilizing the C8 base-specific quadruple interactions. Probable unpairing of a junctional base pair in frameshifting pseudoknots as observed here with U13 has also been reported in several biochemical\textsuperscript{11,37} and NMR studies\textsuperscript{14}. The ribosomal frameshifting efficiency of the luteovirus pseudoknot is relatively low compared to those that have different junctional sequences\textsuperscript{4,6}. While the stabilization in loop 2 appears to be a general feature utilized by frameshifting pseudoknots, the junctional interactions may be different in high-efficiency frameshifting systems\textsuperscript{6,11,12}. In those systems, the identity of the loop 1 base analogous to C8 is not crucial for frameshifting activity. Nevertheless, it has been proposed that a certain bent conformation induced by an adenosine wedged base between the two stems may also be important for high-level frameshifting\textsuperscript{11}. Recently, a 4.8 Å crystal structure of a synthetic SELEX pseudoknot complexed with HIV-1 reverse transcriptase was analyzed. Although not all of the pseudoknot was observed, the structure showed a pronounced kink between the two stems, thereby optimizing its contacts with subunits of the heterodimer\textsuperscript{38}.

The BWYV structure presents the first detailed view of an extended minor-groove RNA triplex, a new motif for mediating intramolecular and possibly intermolecular interactions. Interactions of this type have been predicted to form between the P1 substrate helix and the single-stranded 18/7 segment in the active site of the Tetrahymena intron (Szewczak, A.A., Orteleva-Donnelly, L., Ryder, S.P., Moncoeur, E. & Strobel, S.A., pers. comm.). That study suggested that four consecutive base-triples and 2'-OH mediated tertiary interactions form from the P1 helix. In BWYV, 10 of the 16 hydrogen-bonding contacts in the minor groove are mediated by 2'-OH groups. The 2'-OH groups extend from both edges of the minor groove and are involved in every
loop 2 nucleotide interaction. Therefore, it is not surprising that
the minor groove is narrowed to clamp the third strand into close
contact. The loop 2–stem 1 interactions may also be a general fea-
ture in other pseudoknots that are not involved in frameshifting. For
example, three hydrogen-bonding contacts between loop 2 and stem 1 have been reported in the NMR study of a pseudoknot
formed by the turnip yellow mosaic virus genomic RNA.
In the BWYV loop 2 conformation, the 2'-OH and certain
hydrogen bond donor/acceptor groups of the adenine bases are
positioned on the outside surface where they are available for
higher order contact (Fig. 3a). It is possible that this interface is
recognized by proteins or can have dynamic interactions with the
ribosome. An interface of this kind could also be involved in
domain organization for packing of large RNA molecules. Other
regions of loop 2 that are not involved in tertiary interactions may
be fully accessible for ribosomal contact, for example, near G19.
Insight into the requirements for ribosomal frameshifting
has been obtained using a reticulocyte lysate translation system
to measure the activity of the BWYV pseudoknot and its
mutants (Kim, Y.G., Su, L., Maas, S. & Rich, A., pers. comm.). Changing the base pair C3–G18 at the bottom of stem 1 has lit-
tle effect, but substitution of A25 or C8 abolishes activity. The
studies also show that frameshifting is dependent upon loop 2
adenosines. However, the extruded U13 can be deleted, and
most of the wild-type frameshifting activity is retained. These
findings suggest the importance of the RNA triplex and quater-
nary interactions for ribosomal frameshifting.
The BWYV pseudoknot structure at atomic resolution iden-
tifies several features that appear to be important for -1 riboso-
mal frameshifting, and will facilitate further analysis regarding
its mechanism. Future work that focuses on the importance
and nature of the ribosome–pseudoknot contact as well as the
possible participation of proteins or other factors will be
required to provide a full picture in this process. In addition,
this structure may help design drugs for antiviral therapy that
interfere with frameshifting, an event critical for the viral life
cycle.

Methods
Crystalization and data collection. The 28-nucleotide RNA was
prepared by in vitro T7 transcription and purified to homogeneity
by polycrylamide gel electrophoresis. Brominated RNA was pre-
pared by substituting UTP with 5-Br-UTP, leading to the incorpora-
tion of one bromo-uracil in the RNA sequence. The RNA was
extracted from the gel by the crush-and-soak method and concen-
trated with Centriplus (10 kDa) (Amicon) units. After extensive dialy-
sis against double-distilled water, the RNA was neutralized with 300
mM sodium cacodylate (pH 6.5), incubated at 60 °C for 10 min in the
presence of 15 mM MgCl2 and then cooled to 25 °C and

![Fig. 5 Stereo view of stabilizing interactions at sharp turns, ions and water molecules.](image)
**Table 1 Data collection, phase determination, and refinement statistics**

| Crystallographic data | Data set | Resolution (Å) | Completeness (%) | Rmerge (%) | R cryst (%) | Rfree (%) | Redundancy |
|-----------------------|----------|----------------|------------------|------------|-------------|-----------|------------|
| Native I              | 30–1.6   | 93.0/69.9      | 7.0/16.5         | 18.7/3.5   | 7.1         |           |            |
| Native II             | 25–2.1   | 91.8/67.2      | 6.8/18.4         | 13.1/3.4   | 5.8         |           |            |
| Os (λ = 1.1395 Å)     | 25–2.2   | 86.8/54.0      | 5.7/18.9         | 18.5/5.4   | 5.4         |           |            |
| Br (λ = 0.8550 Å)     | 25–2.0   | 89.2/88.5      | 5.3/32.1         | 16.9/2.3   | 5.8         |           |            |
| Heavy-atom sites (fractional) | Br | x = 0.7536 | y = 0.0683 | z = 0.0990 |            |           |            |
| Os (major)            |         | x = 0.2045     | y = 0.9397       | z = 0.0607 |            |           |            |

**MIRAS phase determination (native I as native)**

| Resolution range (Å) | Rmerge (sites) | Rmerge (iso/calc) | Rmerge (ano) | Phasing power iso/calc, ano |
|----------------------|---------------|-------------------|--------------|-----------------------------|
| Os                   | 25–2.5        | 22.7%             | 0.82/0.84    | 0.96/1.3, 2.5               |
| Br                   | 25–2.5        | 13% (1)           | 0.87/0.85    | 1.10/1.3, 1.2               |

Overall figure of merit after solvent flattening (43% solvent content) 0.896 (1.9 Å)

**Refinement (native I, 766 nonhydrogen atoms, 124 water molecules, one magnesium ion, and one sodium ion)**

| Resolution (Å) | Reflections/free set | Rmerge, Rfree (%) | R.m.s.d. | Phases (°) | Angle (°) | B-factor (Å²) | RNA |
|----------------|----------------------|------------------|----------|------------|-----------|---------------|-----|
| 8–1.6          | 8,793/916            | 20.7/25.4        | 0.017    | 1.87       | 7.0       | 24.6          |     |

1All data last shell. The last shell corresponds to 1.66–1.6 Å, 2.2–2.1 Å, 2.2–2.2 Å and 2.06–2.00 Å for native I, native II, Os and Br data sets, respectively.

2Rmerge = Σl<|FI|>–|<F|>/Σl|FI|, where l is the observed intensity, <I> is the statistically weighted average intensity of multiple observations of symmetry-related reflections.

3I/o = Average intensity/error.

4Rmerge = Σl<|FI|>–|<F|>/Σl|FI|, where l is the nucleic acid structure factor amplitude and |l|F|> is the heavy-atom derivative structure factor amplitude. 

5RI = c<phase-integrated lack of closure>/<|F|>–|<F|> <I> is centric data and ac is acentric data; iso is isomorphous and ano is anomalous.

6Phasing power = c<|F|> / phase-integrated lack of closure>, where |F| is the heavy-atom structure factor amplitude.

7Rcryst = c<|FH|–|<FPH|>/c<|FH|>, where |F| is the heavy-atom derivative structure factor amplitude.

8Rsms = c<|FH|–|<FPH|>/c<|FH|>, where |FH| is the heavy-atom structure factor amplitude.

9Overall R value = 0.15.`

10Values are calculated for the remaining 90% of reflections used for structure refinement.

11Rms.d. is the root mean square deviation from ideal geometry.

4 °C. Native I crystals (0.6 × 0.15 × 0.15 mm³) were grown at 25 °C from a sitting drop containing 4 mg ml⁻¹ RNA, 5 mM MgCl₂, 5% sec-butanol, 2 mM spermidine, 100 mM potassium MOPS buffer (pH 7.0) over a reservoir of 18% sec-butanol. The crystals diffracted to 1.6 Å, and crystallized in the space group P₃₂₁ (a = b = 30.08 Å, c = 140.08 Å). Brominated crystals could only be grown from solutions containing 5% 2-methyl-2,4-pentanediol (MPD), 5 mM MgCl₂, 1 mM spermidine, 100 mM potassium PIPES buffer (pH 6.5). These brominated crystals were slightly non-isomorphous with the native I crystals grown out of sec-butanol. However, native crystals (native II) later grew from the MPD conditions and were fully isomorphous with the brominated crystals. Osium derivatives were prepared by soaking the native II crystals in 10 mM Osium (iii) hexamine trifluoromethanesulfonate for 15 h and backsoaking them in mother liquor overnight.

Data collection on native II crystals was carried out on a Rigaku R-AXIS IIC imaging plate detector with a Rigaku RU200 copper rotating-anode generator, equipped with double-focusing mirrors. Native I data were collected at Brookhaven beamline X4A. Multiple anomalous dispersion (MAD) data for the bromine and osmium derivative crystals were collected at the Brookhaven X4A and CHESS F2 beamlines, respectively. All data were collected at 100 K, and crystals were soaked in mother liquor containing 25% MPD as a cryoprotectant for 1 min before freezing. Data were processed using the programs DENZO, SCALEPACK⁴, and the CCP4 suite⁴⁴.

**Phase determination and structure refinement.** There is one RNA pseudoknot per asymmetric unit. The single bromine site was located from the 2.1 Å isomorphous-difference Patterson map of native II and the bromine derivative data. Five Os sites were located in the Os derivative data collected at the absorption edge by difference Fourier techniques, using the bromine phases. These sites were also confirmed in the Os anomalous-difference Patterson maps. We used a combination of multiple isomorph-replacement and anomalous-scattering (MIRAS) methods for phasing with the program SHARP⁴⁵. The Residual maps from SHARP revealed other characteristics of the heavy-atom sites, and the space group enantiomorph was determined. The electron density maps were greatly improved through the addition of an alternative position for the major osmium-binding site along with refinement of the brominated crystals. Osium derivatives were prepared by soaking the native II crystals in 10 mM Osium (iii) hexamine trifluoromethanesulfonate for 15 h and backsoaking them in mother liquor overnight.

| Residue | Distance (Å) | Residue | Distance (Å) |
|---------|--------------|---------|--------------|
| Loop 2  |
| A20 Z'-OH¹ | C5 O2 | 2.83 | A20 N1 | G16 Z'-OH | 3.05 |
| A20 Z'-OH² | C17 O2 | 2.81 | A20 N1 | G16 Z'-OH | 2.90 |
| A20 Z'-OH³ | G16 N2 | 3.17 | A20 N1 | G16 Z'-OH | 2.89 |
| A20 Z'-OH⁴ | C17 Z'-OH | 2.57 | A20 N1 | G16 Z'-OH | 3.30 |
| A20 N1 | G4 Z'-OH | 2.70 | A20 N1 | G15 Z'-OH | 2.69 |
| A20 N3 | G4 N2 | 3.02 | A20 N3 | G15 O2 | 2.98 |
| A20 N7 | C5 Z'-OH | 3.30 | A20 N7 | G15 N | 3.30 |
| A24 N1 | G7 N2 | 3.16 | A24 N1 | G7 N3 | 3.01 |
| A24 N6 | G7 Z'-OH | 2.91 | A24 N6 | A25 O4 | 2.82 |

¹Bifurcated hydrogen bonds.

²Probably protonated.

³Ifurcated hydrogen bonds.
anisotropic temperature factors for the bromine site. Using the program O (ref. 46), a poly(C) model with appropriate sugar puckers was built into the solvent-flattened electron density maps generated from native I and the bromine/oxygen derivative data. Using SHARP, an excellent 1.9 Å electron density map was calculated from native I and the bromine/oxygen derivative data, confirming the trace, clarifying the identity of bases and revealing the position of the 5'-end -phosphate. Initial refinement cycles were performed with X-PLOR (ref. 66); early observation of unaccounted densities in the MRAS maps and the later appearance of connected density in the simulated-annealing omit maps suggested two conformations of nucleotide U13 and part of the phosphate backbone of residue C14. The two conformations were included in refinement and relative occupancies calculated (C2'-endo/C3'-endo 60% : 40%). Final rounds of refinement to 1.6 Å were performed using CNS (ref. 67), which rapidly decreased both the R-factor and Rfree. Bulk solvent and anisotropic temperature factor corrections were carried out in the final process. Statistics are presented in Table 1. A high-density peak that appeared in the solvent-flattened maps was tentatively identified as a sodium ion, with the coordination distance criteria to water molecules or donors of the RNA as 2.4–2.8 Å. The magnesium ion was identified by its strict octahedral coordination with ligand distances between 2.1 and 2.3 Å. Helical parameters for the structure were generated by the program CURVES (ref. 77). Coordinates. Coordinates have been deposited in the Protein Data Bank (accession code 437D).

Acknowledgments

We thank E. de La Fortelle for extensive help on the program SHARP, G. Minasov and I. Berger for help in data processing and refinement, Y.-G. Kim for frameshifting discussions, C. Ogata (X4A-NSLS), M. Sofits (SSRL) and the staff at CHESS F2 for synchrotron assistance. We acknowledge help from T. Schwartz, D. Chan, M. Schade and K.-J. Wu on data collection; V. Tershko, A. and H. Sioke for technical assistance; J. Cate, D. Bartel, M. Roul, U. RajBhandary, A. Herbert and B. Brown for helpful discussions. We are indebted to H. Tauze and S. Holbrook for providing us with the osmium hexammine trflate compound; J. McCloseley and P. Crain for bromine mass spectroscopy analysis. We also thank L. Gerlhe, K. Lowenhaupt and F. Houser-Scott for assistance in the T7 transcription system; D. Laufenerburger and S. Zhang at the MIT Center for Biomedical Engineering for providing the computer facility. This research was supported by grants from the National Science Foundation, the National Institutes of Health, the National Foundation for Cancer Research and the National Aeronautics and Space Administration.

Received 24 September, 1998; accepted 7 December, 1998.