Structure and Function of the Ribosomal Frameshifting Pseudoknot RNA from Beet Western Yellow Virus

by **Martin Egli***[a]), **Sanjay Sarkhel**[a]), **George Minasov**[b]), and **Alexander Rich***[c])

[a]) Vanderbilt University, Department of Biological Sciences, Nashville, Tennessee 37235 USA
(fax: 1-615-343-6707, e-mail: martin.egli@vanderbilt.edu)
[b]) Northwestern University, Department of Molecular Pharmacology and Biological Chemistry, Chicago,
Illinois 60611, USA
[c]) Massachusetts Institute of Technology, Department of Biology, Cambridge, Massachusetts 02139, USA
(fax: 1-617-253-8699, e-mail: cbeckman@mit.edu)

Dedicated to Professor *Jack D. Dunitz* on the occasion of his 80th birthday

Many viruses reprogram ribosomes to produce two different proteins from two different reading frames. So-called − 1 frameshifting often involves pairwise alignment of two adjacent tRNAs at a 'slippery' sequence in the ribosomal A and P sites such that an overlapping codon is shifted upstream by one base relative to the zero frame. In the majority of cases, an RNA pseudoknot located downstream stimulates this type of frameshift. Crystal structures of the frameshifting RNA pseudoknot from Beet Western Yellow Virus (BWYV) have provided a detailed picture of the tertiary interactions stabilizing this folding motif, including a minor-groove triplex and quadruple-base interactions. The structure determined at atomic resolution revealed the locations of several magnesium ions and provided insights into the role of structured water stabilizing the RNA. Systematic *in vitro* and *in vivo* mutational analyses based on the structural results revealed specific tertiary interactions and regions in the pseudoknot that drastically change frameshifting efficiency. Here, we summarize recent advances in our understanding of pseudoknot-mediated ribosomal frameshifting on the basis of the insights gained from structural and functional studies of the RNA pseudoknot from BWYV.

Introduction. – Pseudoknots represent a common folding element of RNA in which nucleotides from single-stranded loops undergo base pairing with nucleotides outside the loop region [1]. The majority of them belong to the H-type (*h*airpin) class of pseudoknots, featuring two stem regions (S1 and S2) and two connecting loops (L1 and L2; *Fig. 1*) [2]. The two stems form a more or less continuous double-helical structure, whereby one strand is continuous and the other is interrupted. The shorter loop 1 crosses the major groove of stem 2, and loop 2 crosses the minor groove of stem 1. The variations in the relative orientations of stems 1 and 2 result in considerable diversity in the global structures of pseudoknots as well as the arrangements of loops and their interactions with the helical segments [3][4]. Thus, the two stems can be stacked more or less seamlessly, giving rise to a continuous A-form duplex. Alternatively, bending, over- and underwinding, and helical displacement or combinations thereof at the stem 1−stem 2 junction produce pseudoknots whose overall conformations can differ considerably [4].

Pseudoknots have been found in practically all classes of RNAs. They include ribosomal RNAs [5], mRNAs [6][7], catalytic and self-splicing RNAs [8−10], ribonucleoprotein complexes [11], viral genomic RNAs [12], and *in vitro*-selected
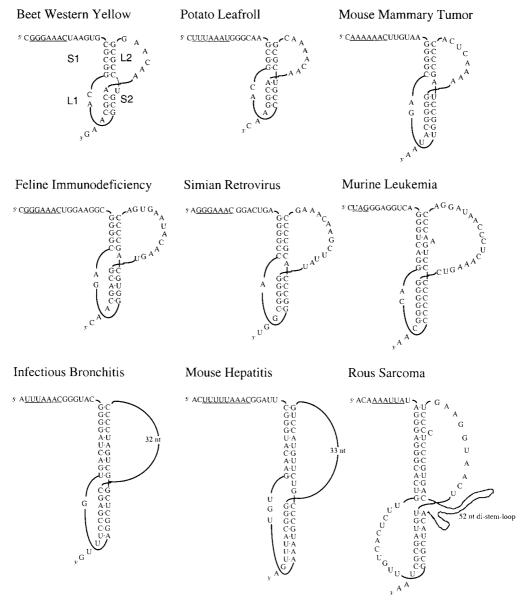
Fig. 1. *Proposed secondary structures of ribosomal frameshifting pseudoknots from a variety of viruses.* The slippery or shifty sequences are underlined. In the individual pseudoknot structures, stem 1 (S1) and loop 2 (L2) are at the top, and stem 2 (S2) and loop 1 (L1) are at the bottom.

RNA aptamers [13]. Beyond their role as structure-stabilizing folding motifs, pseudoknots in mRNAs were found to act as regulators of protein synthesis. For example, pseudoknots sometimes modulate the specific binding of proteins to their

cognate mRNAs in translation initiation, thus regulating or autoregulating expression of the downstream gene [7][14][15]. In these cases, the pseudoknot is positioned in the noncoding leader region or overlaps with the ribosome binding site and/or the mRNA initiation codon. Pseudoknots present in the coding regions of mRNAs can stimulate programmed −1 ribosomal frameshifting, a mechanism used by many viruses, including both tumor viruses and retroviruses, bacteria, yeast, and DNA insertion sequences (reviewed in [16][17]). A completely different form of translational suppression, found in murine leukemia virus, so-called readthrough suppression, also uses a downstream pseudoknot signal [18][19].

As a result of a −1 frameshift in the reading frame, a stop codon is avoided, and a fusion protein is produced. For example, in *Rous* sarcoma retrovirus, a −1 frameshift leads to expression of a *gag-pol* fusion protein between the *pol* gene that encodes integrase, protease, and reverse transcriptase, and the downstream *gag* gene that encodes virus core proteins [20] (*Fig. 2*). This poly-protein precursor is later processed to yield the mature protein products. The frameshifting efficiency varies greatly in different systems and can range from *ca.* 1% to 30%. The level of frameshifting appears to be somewhat dependent on the complexity of the pseudoknot motif, *i.e.*, the lengths of the stems and loop 2 (*Fig. 1*). The pseudoknot is positioned six to eight nucleotides downstream from a 'slippery sequence' (*Fig. 1*) and causes ribosomal pausing and simultaneous slippage of the tRNAs bound at the A and P sites by one base in the 5′-direction [20] (*Fig. 2*). Messenger-RNA shift sites exhibit an X XXY YYN consensus sequence (the spaces indicate the original reading frame and X and Y can be identical).

Three-dimensional structures of frameshifting pseudoknots have been determined by both solution NMR [21–26] and X-ray crystallographic methods [27][28]. The structures confirmed that the two stems are stacked upon each other [21][23][24][27], and that the helical junction in some pseudoknots is bent [21][24][27]. The structural results based on a variety of studies have recently been reviewed [3][4]. The crystal structure of the frameshifting RNA pseudoknot from Beet Western Yellow Virus (BWYV; *Fig. 3*), a plant luteovirus that regulates the expression of an RNA-dependent RNA polymerase [29], has yielded detailed insights into the interactions between loop 2 and the minor groove of stem 1, and those formed between loop 1 and the major groove of stem 2 [27]. The structure of the BWYV pseudoknot at atomic resolution also revealed binding of several di- and monovalent metal cations, some of which may play a functional role [28]. Combined with the detailed analysis of the frameshifting efficiencies of mutants of the BWYV pseudoknot *in vitro* and *in vivo*, the structure data provide a better understanding of the role of the viral RNA pseudoknot in the −1 frameshifting process.

Results. – *Overall Structure of the BWYV RNA Pseudoknot and Tertiary Structural Interactions.* The crystal structure of a 28-nucleotide construct of the BWYV RNA pseudoknot (*Fig. 3*) in a trigonal crystal form was initially determined by the multiple isomorphous replacement technique and refined to a resolution of 1.6 Å [27]. Subsequently, the resolution of the structure of this crystal form was improved to 1.25 Å and published along with the structure of a cubic crystal form at 2.85-Å resolution [28]. Final coordinates and structure factors for the trigonal and cubic crystal forms are deposited in the *Protein Data Bank* [30] (PDB codes 1L2X and 1L3D, resp.). The two
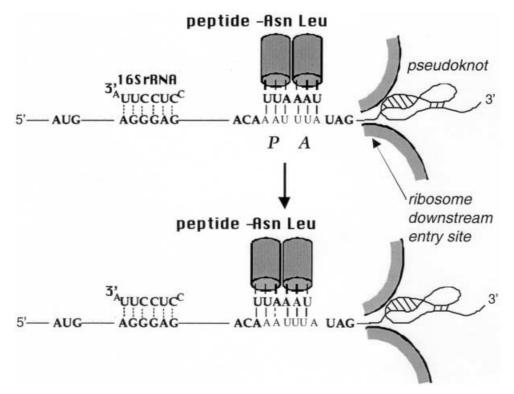
Fig. 2. *Schematic illustration of − 1 translational frameshifting in* Rous *sarcoma virus regulating the relative expression of the* gag-pol *protein, precursor to the enzymes responsible for reverse transcription and integration* [20]. Simultaneous slippage of two adjacent tRNAs by a single nucleotide in the 5′-direction at the shifty sequence (residues in thin-line font) mediated by a downstream stimulator (pseudoknot) suppresses the *gag* terminator (**UAG**). The tRNA[Asn] carrying the nascent peptide and the tRNA[Leu] (gray cylinders) are depicted bound to the *gag*-frame codons, AAU and UUA, at the ribosomal **P** and **A** sites, respectively (*top*). After frameshift (*bottom*), peptidyl transfer and three-nucleotide translocation (not shown) the next pol-frame codon (**AUA**) is placed at the **A** site where it is decoded by tRNA[Ile].

5′-terminal G residues (G1 and G2 in *Fig. 3*) constitute a single-stranded overhang in the *in vitro* transcribed RNA, but are not actually part of the natural BWYV pseudoknot sequence.

The BWYV RNA pseudoknot adopts a compact structure with overall dimensions of *ca.* $32 \times 36 \times 22$ Å (*Fig. 4*). The main features of the structure are the junction between stems 1 and 2, marked by a kink and over-twisting, the minor-groove triplex between loop 2 and stem 1, and a new quadruple-base interaction that comprises the G12-C26 base pair, C8 from loop 1 and A25. The absence of base pair U13-A25 in stem 2 at the helical junction (*Fig. 1*) constitutes a further surprise of the crystal structure (*Figs. 3* and *4*). Instead, U13 is looped out, and A25 stacks onto loop 2 and interacts with both C14 and C8 from stem 1 and loop 1, respectively. Another remarkable finding is that the number of H-bonds mediating tertiary interactions (26) in the pseudoknot exceeds that of *Watson−Crick*-type H-bonds (24). The intricate H-bonding motifs
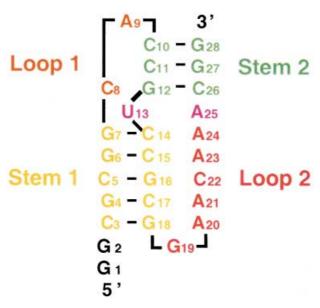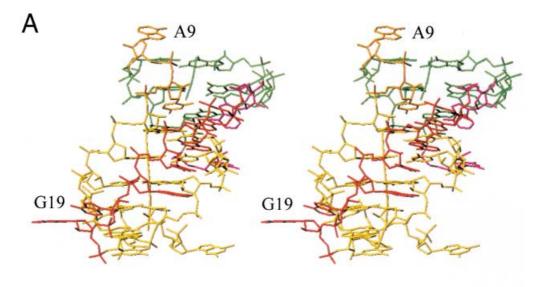
Fig. 3. *Secondary structure diagram of the RNA pseudoknot from BWYV based on the X-ray crystallographic analysis.* Note that the orientation here is inverted relative to that in *Fig. 1.*

involving base functions and ribose 2′-OH groups as well as phosphate O-atoms in the tight turns formed by the junctions are consistent with a stable fold of the pseudoknot that may bear upon its function.

At the helical junction, stem 1 and stem 2 are tilted relative to each other by *ca.* 25° (*Fig. 4,A*). The positions of the helical axes calculated with the program CURVES [31] indicate a *ca.* 5 Å relative displacement of stems 1 and 2. In addition, G7 and A25, the top and bottom bases of stem 1 and stem 2, respectively, are rotated by *ca.* 50°, thus considerably exceeding the canonical A-form twist of 33° between adjacent base pairs. Both duplexes exhibit deviations from standard A-form geometry, and they curve toward their major grooves (*Fig. 4,B*). Compared to the predicted secondary structure in which both stems are presumed to be coaxially stacked, the C26–G28 strand of stem 2 is rotated away. As a result, the helical junction is formed by G12 and C14, which are stacked, and A25 linking C26 to loop 2 (*Fig. 3* and *Fig. 4,B*).

The triple-stranded arrangement between loop 2 (residues A20 to A25) and the minor groove of stem 1 differs significantly from the familiar triplexes formed by a B-form DNA duplex and a third strand in the major groove. The interactions between loops and stem 1 are RNA-specific, as every interaction between a loop-2 nucleotide and stem 1 in the six layers of the triplex involves a ribose 2′-OH group. The orientation of loop-2 residues relative to the stem-1 minor groove is relatively irregular in the sense that the loop bases switch between the two strands of stem 1 in different layers of the triplex motif [27]. Thus, A20 contacts two layers of base pairs through a base triplet. A21 and C22 contact G16, A23 interacts with C15, and A24 is engaged in H-bonds to both the base and the ribose moieties of G7. A20 acts as an anchor for loop 2 in the minor groove and forms no fewer than seven H-bonds to stem-1 residues (*Fig. 4,A*).
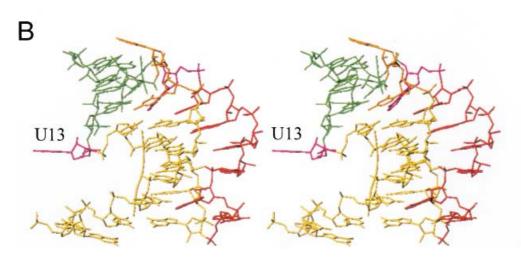
A

B

Fig. 4. *Stereo illustrations of the overall structure of the BWYV RNA pseudoknot. A*) View onto loop 1 (orange) and loop 2 (red). *B*) Rotated by 90° around the vertical. The color scheme is identical to that in *Fig. 3*, and helical axes of stem 1 and stem 2 are shown in yellow and green, respectively. Helical parameters were calculated with the program CURVES [30].
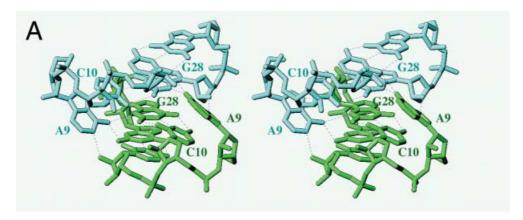
Loop 2 and loop 1 get to lie on the same face of the pseudoknot structure, the latter crossing the major groove of stem 2 (*Fig. 4,A*). In fact, only C8 is positioned in the major groove. Residue A9 caps stem 2 and is part of the 'C-turn' (C8, A9, and C10) that brings about a sharp change of direction at the loop 1 to stem 2 junction. C8 engages in

a quadruple-base interaction with G12, A25, and C26 [27]. Its *Watson–Crick* face is directed toward the major-groove edge of base pair G12-C26, leading to formation of three H-bonds (N(4)$-$H[C8]$\cdots$N7[G12], N(3)$^+$$-$H[C8]$\cdots$O6[G12], and O(2)[C8] $\cdots$H$-$N(4)[C26]; *Fig. 4*). The close spacing of N(3) of C8 and O(6) of G12, both are normally H-bond acceptors, is consistent with protonation of C8. Indeed, the thermodynamic stability of the BWYV RNA pseudoknot is increased at lower pH values (*i.e.*, pH 6 *vs.* pH 8) [32]. An additional H-bond is then formed between O(2) of C8 and N(6) of A25. This quadruple-base interaction involving a protonated cytosine was recently also found in the solution NMR structure of the frameshifting RNA pseudoknot from pea enation mosaic virus RNA1 (PEMV-1), another plant luteovirus [26]. Even more remarkable, this structural motif was present in almost identical form in the structure of hepatitis delta virus (HDV) ribozyme [10], despite the completely different sequence and structural contexts [3][26].

The compact fold of the BWYV RNA pseudoknot necessitates tight turns at the stem–loop junctions. The above C-turn encompassing C8, A9, and C10 at the transition between loop 1 and stem 2 is one example (see *Fig. 4,A* and *B*, top). Another sharp turn is located at the junction between stem 1 and loop 2 (see *Fig. 4,A* and *B*, bottom). This turn involves residues G18, G19, and A20. At both junctions, the turns are stabilized by H-bonds between base functions and phosphate or 2′-OH groups [27].

Two Crystal Forms. The BWYV RNA pseudoknot crystallizes in two forms, a trigonal high-resolution form and a cubic low-resolution form. Crystals of the first form can be grown from 2-methylpentane-2,4-diol (MPD), and crystals of the second form were grown from high gradients of ammonium sulfate. The three-dimensional structures of the pseudoknot display only minor deviations as a consequence of the different environments in the two crystal lattices [28]. Most of these deviations concern unpaired residues in either the stem-loop junctions or the 5′-end as well as residue U13. The trigonal crystal form is characterized by extensive stacking interactions involving the looped-out residues G1, U13, and G19. In the less densely packed cubic crystal form, G19 at the stem1–loop2 turn participates in a junction formed by four symmetry-related molecules (*Fig. 5*). *Fig. 5* also illustrates the different lattice interactions for residue A9. In the cubic crystal form, A9 is part of the above four-way junction. In the trigonal crystal form, A9 forms H-bonds to C11 and G28 in the minor groove of a symmetry-related molecule *via* its N(6) amino group (*Fig. 5,A*). The similar overall structures of the RNA in the two crystal forms along with conserved features, such as the minor-groove triplex and the quadruple-base interaction in the major groove of stem 2, indicate that the pseudoknot conformation is practically unaffected by lattice forces.

Hydration. Water molecules form an intricate part of the RNA pseudoknot structure. Ordered solvent molecules help stabilize junction regions and mediate stem-loop interactions. For example, loop 2 dissects the minor groove of stem 1. As a result, phosphate groups from loop 2 and the C5–G7 strand portion from stem 1 face each other across one side of the dissected minor groove (*Fig. 4* and *Fig. 6,A*). Phosphate groups and 2′-OH groups serve as a scaffold for an extended spine of H_2O molecules and a putative sodium ion that runs along the entire loop 2 and comprises W128, W178, W108, Na63, W165, W181, W111, W153, W187, W137, and W176 (*Fig. 6,B*). In addition, a string of H_2O molecules links the ribose 2′-OH group of residue G18 to the looped-
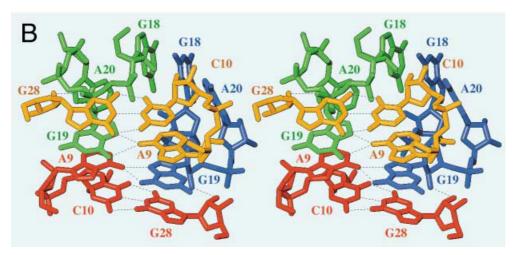
Fig. 5. *Stereo illustrations of the different packing motifs involving the capping residue A9 in the* A) *trigonal and* B) *cubic crystal forms of the BWYV RNA pseudoknot.* Individual RNA molecules are colored differently and residues are labelled.

out residue G19 at the stem 1–loop 2 junction: [G18]O(2′)···W99···W224···W37···W231···W239···W235···O(6)/N(7)[G19]. A further H_2O molecule (W73) bridges W231 to the C8 position of G19. Other noteworthy hydration patterns between stem 1 and loop 2 include a five-membered ring (W125, W118, W92, Na(61) and O(1P) [A21]). W125 connects the O(1P) of phosphate of A21 to both the 2′-OH group and N(3) of residue G6 in stem 1. W118 is engaged in direct contacts to residue G7 of stem 1 and also interacts with A23 of loop 2 *via* the O(4′) and N(7) positions, respectively.

As part of the interactions between loop 1 and stem 2, H_2O molecules stabilize the insertion of C8 into the major groove and the formation of the quadruple-base motif (*Fig. 7*) [27]. Thus, W97 bridges the ribose 2′-OH of C(8) and C(5) of residue C26 (*Fig. 7,B*). In addition to a direct interaction between the phosphate group of A9 and
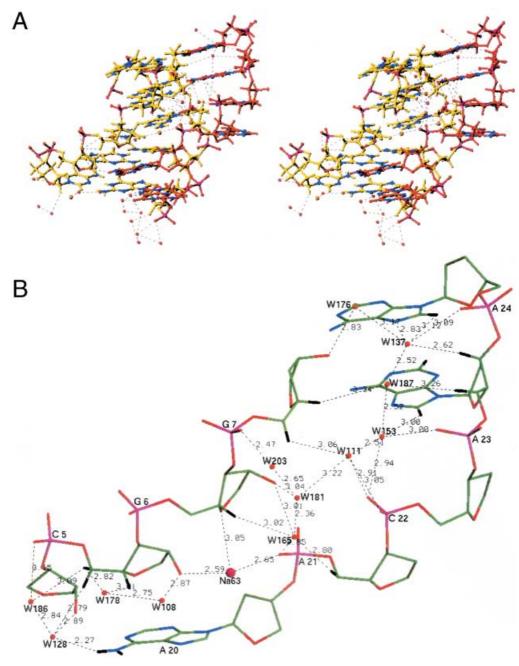
Fig. 6. A) *Stereo diagram of the water structure between strand C3-G7 and loop 2 residues in the minor groove of stem 1.* Water molecules and metal ions are depicted as red and orange spheres, respectively, and H-bonds (C−H donors included) are dashed lines (cutoff 3.4 Å). The color scheme for C-atoms is identical to that in *Fig. 3*, and RNA atoms are colored red, blue, and magenta for phosphate O-atoms O1P and O2P and N and P, respectively. B) *Close-up view of the water structure in a portion of the minor groove of stem 1.* Water molecules are red spheres, and H-bonds are dashed lines with distances given in Å.

the exocyclic amino group of C10 in the C-turn [27], O(2P) of A9 is also involved in a H_2O-mediated (W71) interaction to N(4) of C11. The same phosphate O-atom is also linked to phosphates from stem-2 residues *via* a tandem of H_2O molecules: [A9]O(2P) $\cdots$ W71 $\cdots$ W68 $\cdots$ O(1P)[C10,C11]. A series of H_2O molecules (W97, W139, W223, W75, W44, and W39) are crucial for connecting loop 1 to stem-2 residues. Thus, both the O(2P)-atom of phosphate of A9 and the 2′-OH group of C8 are bridged to stem-1 residues by the above hydration motif (*Fig. 7,B*). The following combination of H_2O
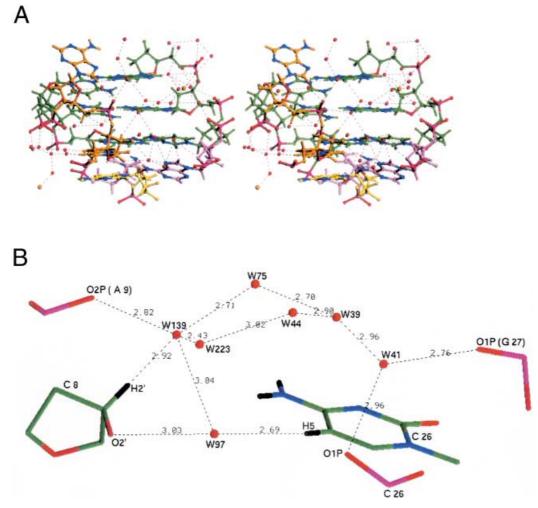

Fig. 7. A) *Stereo diagram of the water structure around loop-1 residues C8 and A9 in the major groove of stem 2.* Water molecules and metal ions are depicted as red and orange spheres, respectively, and H-bonds (C−H donors included) are dashed lines (cutoff 3.4 Å). The color scheme for C-atoms is identical to that in *Fig. 3*, and RNA atoms are colored red, blue, and magenta for phosphate O-atoms O1P and O2P, and N, and P, respectively. B) *Close-up view of the water network on one side of loop 1 in the major groove of stem 2.* Water molecules are red spheres, and H-bonds are dashed lines with distances given in Å.

molecules connects 2′-OH of C8 to the O(6) positions of G27 and G28: W97, W139, and W75. Similarly, the following H_2O molecules connect the above 2′-OH to N(7) of G28: W97, W139, W223, W44, and W40. The following combination of H_2O molecules connects the O(2P)-atom of phosphate of A9 to the O(6) positions of G27 and G28: W139 and W75. Similarly, the following H_2O molecules connect O(2P) of A9 to N(7) of G28: W139, W223, W44, and W40. Two more H_2O molecules, W39 and W41, help extend the above H_2O network to the phosphate backbone of the second strand in stem 2 (*Fig. 7,B*). Thus, [C8]O(2′)···W97···W139···W75···W39···W41··· O(1P)[C26,G27] and [C8]O(2′)··· W97···W139···W223···W44···W39···W41··· O(1P)[C26,G27]. Similarly, the ordered water structure that links the phosphate group of A9 and stem-2 residues includes W139···W75···W39···W41···O(1P)[C26,G27] and W139···W223···W44···W39···W41···O(1P)[C26,G27].

Metal Ion Coordination. Metal ions, in particular Mg^{2+}, can make significant contributions to the thermodynamic stability of RNA pseudoknots (*e.g.*, [32][33]). The crystal structure of the trigonal form of the BWYV RNA pseudoknot refined to 1.25-Å resolution revealed several mono- and divalent metal cations [28]. Six ordered Mg^{2+} ions were observed per crystallographic asymmetric unit. Thus, each RNA molecule exhibits contacts to 15 Mg^{2+} ions. The most-interesting cations with regard to a stabilizing influence are Mg45 and Mg52, bound in the major groove of stem 1 (*Fig. 8*). Both are hexahydrates and are coordinated to O(6) and N(7) of guanine(s). The two ions are also the only ones among the six that have contacts to a single RNA molecule rather than linking neighboring pseudoknots in the crystal lattice. Mg52 occupies a position roughly in the plane defined by the guanine base of residue 4, and its contribution to RNA stability is most likely of an electrostatic nature (*Fig. 8*). Mg45 occupies a key position in the pseudoknot as it is located at the interface between the two stems [28]. It contacts O(6) and N(7) of both G6 and G7 *via* its hydration shell and makes an additional H-bond to the phosphate group of residue G12 from stem 2 across the groove (*Fig. 8*). Therefore, it can stabilize the particular conformation of the helical junction, specifically the overwinding, kinking, and displacement at the stem 1-stem 2 transition.

Compared with Mg^{2+}, the number and nature of coordinated monovalent metal ions appear to be less-significant with regard to the overall stability of the BWYV pseudoknot [32]. In the structure of the trigonal crystal form, four putative alkali metal ion binding sites were identified [28]. Initially, all solvent peaks in the electron-density maps were treated as H_2O molecules. The above four peaks displayed unusually low temperature factors and engaged in multiple interactions to RNA atoms, one or more of which were phosphates. To further address the question of whether these H_2O molecules were actually alkali metal ions, anomalous diffraction data were collected at the absorption edges of K^+ [34] and Tl^+ (with crystals soaked in thallium acetate) [35]. Significant buildup of anomalous difference electron density was observed only around one of the four peaks [28]. However, this K^+ ion is bound near G1 and G2, and thus not of relevance for RNA folding and stability. The three remaining sites were treated as putative Na^+ ions. Two of them are located in the stem-1 minor groove, adjacent to the stem 1−loop 2 junction [28]. There, they link the phosphate group of nucleotide A21 to the ribose 2′-OH groups from residues on opposite strands and to base atoms at the floor of the groove. It is plausible that such a coordination mode can draw together
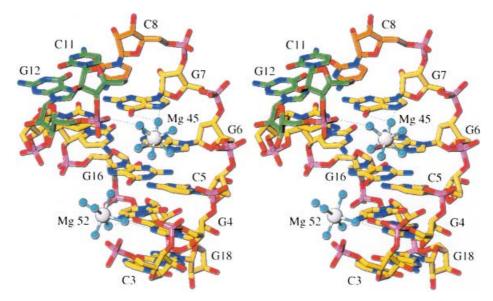
Fig. 8. *Stereo illustration of the coordination modes of two magnesium hexahydrates in the major groove: Mg45 bound at the junction between stem 1* (bottom) *and stem 2* (top)*; Mg52 bound in stem 1*. RNA Atoms are colored yellow (stem 2 in green and loop 1 in orange), red, blue, and magenta for C-, O-, N-, and P-atoms, respectively. Water molecules are cyan spheres, H-bonds are dashed lines, and selected residues are labeled.

stem 1 and loop 2 by potentially relieving an energetically unfavorable interaction between the phosphate group and negatively polarized regions in the minor groove.

Structure and Function. The detailed structural insights for the BWYV RNA pseudoknot served as a guide in a systematic mutational analysis of its frameshifting efficiency *in vitro* and *in vivo* [36]. To assess the efficiencies as a function of the RNA pseudoknot sequence *in vitro*, a cassette consisting of slippery sequence, linker, and pseudoknot was inserted in a vector between the glutathione S-transferase gene and the green fluorescent gene (*Fig. 9*). For the *in vivo* frameshifting assays, luciferase was used as the downstream reporter gene. Endonuclease restriction sites placed immediately upstream of the slippery sequence, within the linker region, as well as downstream of the pseudoknot allowed separate assessments of frameshifting efficiency as a result of mutations in the slippery sequence and the sequence of the pseudoknot (*Fig. 9*) [36]. Moreover, the wild type (wt) pseudoknot and its mutants could be combined with a slippery sequence from another virus to enhance the levels of wt frameshifting. Frameshifting was then measured in two different systems: *in vitro* with rabbit reticulocyte lysate and *in vivo* with human embryonic kidney cells [36].

Mutations, deletions, insertions, and combinations thereof in all regions of the BWYV pseudoknot and their effects on − 1 frameshifting were then tested. The majority of sequence alterations led to a diminished frameshifting efficiency (*Fig. 10*) [36]. The causes of the various effects can be broadly classified into two groups: disruptions of RNA tertiary structure, and interference with the specific interactions between pseudoknot and the downstream ribosome entry site. For example, all base
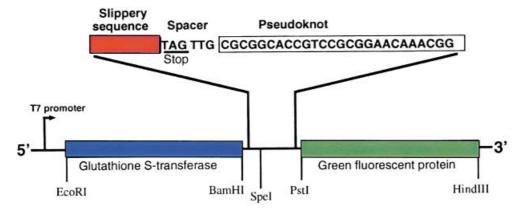
Fig. 9. *Schematic representation of the reporter construct and resulting protein products used for* in vitro *ribosomal frameshifting measurements*

changes in the two loops and the helical junction resulted in significantly reduced efficiencies of frameshifting (*Fig. 10*). This is not surprising given the important role of specific residues, such as C8, A20, or A25, in tertiary-structure formation and stabilization. In addition, shorter loops as a consequence of deletions will lead to disruption of folding or may require significant adaptations at the tertiary-structural level. Similarly, base inversions in the stems had a mostly detrimental effect. This can be explained by the observation in the three-dimensional structure that loop 2 – stem 1 and loop 1 – stem 2 interactions are sequence-specific, consistent with the conservation of adenosine-rich loops 2 and a cytidine in loop 1 with frameshifting pseudoknots from luteoviruses (*Fig. 1*).
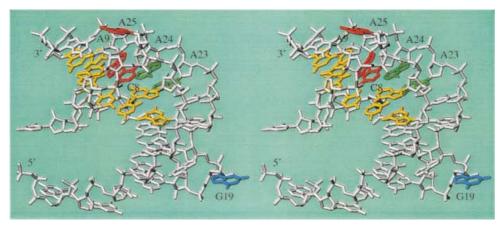

Fig. 10. *Effects of mutations* (green), *mutations/deletions* (red), *insertions* (cyan) *or base-pair inversions* (yellow; G12:C26, C11:G27, G7:C14, G6:C15) *on frameshifting activity.* These changes virtually eliminate frameshifting. The mutation/insertion G19 → U19C19a leads to a 300% increase in frameshifting [36].

The most-remarkable effect is the strongly increased frameshifting efficiency in mutant BWYV pseudoknots that feature insertions or mutations and insertions at the stem 1−loop 2 linker (*Fig. 10*) [36]. For example, a pseudoknot with a G19 → U19C19a mutation/insertion exhibited a 300% gain in frameshifting efficiency (highlighted in cyan in *Fig. 10*). It is unlikely that this is solely due to an altered three-dimensional structure of the RNA pseudoknot or a structure-stabilizing effect because of the looser loop 2. Rather, significant gains or losses in the level of frameshifting following changes of pseudoknot nucleotides that are extruded may be due to altered interactions between pseudoknot and ribosome.

The stem 1−loop 2 junction region and loop 2 are areas of the folded pseudoknot that are being contacted early on by the ribosome as it moves along the mRNA (*Figs. 3, 4,* and *10*) [36]. Upon encountering the pseudoknot, the ribosome pauses and, following the −1 frameshift, unravels the pseudoknot, and continues to translate the message in the new reading frame [20]. When the pseudoknot is replaced by a simple hairpin motif, pausing still occurs but no frameshift takes place [37]. There are, however, viruses that use a simple stem-loop structure to stimulate −1 type frameshifting (*i.e.*, human immunodeficiency virus type 1, HIV-1, and human T cell leukemia virus type 2, HTLV-2; *cf.* [38] and refs. cit. therein). The observed increase in frameshifting as a result of the above sequence alterations in the stem 1−loop 2 junction could directly affect the interactions between RNA pseudoknot and proteins S3, S4, and S5 surrounding the downstream entry site of the ribsosome (*Fig. 11*) [39]. There, the pseudoknot will initially block entry of the mRNA into the tunnel (pausing; *Fig. 2*), followed by an attempt to translocate the message (−1 frameshifting), unwinding of the pseudoknot, and translation in the new reading frame. Thus, the frameshift favored at the slippery site would result from a backlash of the mRNA, triggered by the pseudoknot that is stuck at the tunnel entry site, resisting the translocation force [39]. Certain pseudoknots with altered stem 1−loop 2 junctions are somehow capable of inducing frameshifts more efficiently, evidently by manipulating the ribosome at the pausing or translocation stages or both.

Discussion. – The work presented here describes in some detail the hydration and organization of ions in a ribosomal frameshifting viral pseudoknot. Determination of the crystal structure of the RNA pseudoknot from BWYV has provided us with the most detailed view of a secondary-structural element responsible for ribosomal frameshifting. The molecule is held together tightly with tertiary H-bonding interactions more numerous than the secondary structure *Watson−Crick* H-bonds, and it underlines the extent to which detailed folding of the pseudoknot is important for its biological activity. Functional experiments have been carried out to assess the importance of various elements in the pseudoknot to biological activity [36]. That study revealed, for example, the importance of the triplex interactions between loop 2 and stem 1, as well as the crucial role played by the H-bonding interactions near the junction of stem 1 and stem 2. All of these interactions were interpreted as important in responding to the stress induced by the tug at the 5′-end of the pseudoknot following its engagement with the ribosome. Most startling of all was the discovery that mutations at G19 near the base of stem 1, close to the 5′-end of the pseudoknot, play a crucial role in determining the level of ribosomal frameshifting.
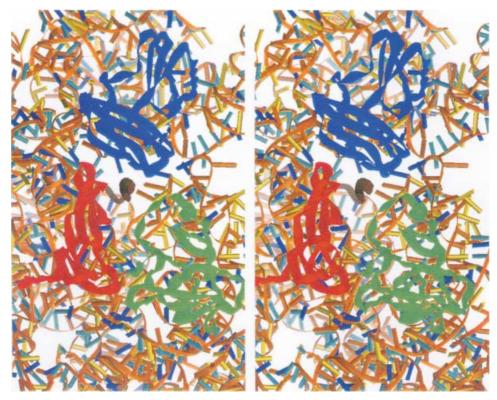
Fig. 11. *Solvent side stereo view of the ribosomal downstream entry tunnel showing the formation of the protein layer surrounding positions +11 to +15 of the mRNA* (center, colored in black) *by proteins S3* (red), *S4* (blue), *and S5* (green), *along with a portion of 16S rRNA* (beige) [39]. The *ca.* 45° bend between the A and P site codons is visible in the gray portion of the mRNA winding through the tunnel.

Our ability to interpret events associated with frameshifting has been transformed by the remarkably clear picture of the ribosome that has emerged from recent X-ray crystallographic studies. We now have detailed views of 50S [40][41] and 30S ribosomal subunits [42][43], as well as a picture of the intact 70S ribosome [44], and a detailed discussion of the path of mRNA as it passes through the ribosome [39]. What is important for ribosomal frameshifting is the fact that the dimensions of the pseudoknot are clearly larger than the dimensions of the ribosomal tunnel through which the mRNA threads on its way to interacting with the decoding A and P sites in the ribosome. The entrance to the ribosomal subunit is flanked by three proteins S3, S4, and S5 as shown in *Fig. 11*. Using the dimensions of the BWYV pseudoknot, it is possible to dock the pseudoknot in the entrance, so that the 5′-end of the pseudoknot is protruding into the tunnel. This allows us to obtain a figure for the length of the polynucleotide chain that is found between residue C3 at the base of the pseudoknot (*Figs. 3* and *4*) and the nucleotide at the 3′-side of the ribosomal A site. From the published work on the path of the mRNA in the ribosome [39], we can see that the

contour distance from the phosphate group of residue C3 of the docked BWYV pseudoknot is 39 Å from the 3′-phosphate of the A site codon (*Fig. 11*). The spacer region between the slippery site and residue C3 of the pseudoknot has six nucleotides (*Fig. 1*). Thus, there are six P⋯P distances in that region.

The distance between phosphate groups on a polynucleotide chain is determined largely by the pucker of the ribose ring. In B-DNA, the normal pucker is *C(2′)-endo*, and, in that conformation, the polynucleotide chain is extended with a distance of *ca.* 7 Å between adjacent phosphate residues [45]. In RNA molecules, the presence of an OH group in the 2′-position introduces a close *Van der Waals* contact when the ribose is in the C(2′)-*endo* pucker conformation. Accordingly, ribonucleotides characteristically have a C(3′)-*endo* pucker which avoids the *Van der Waals* clash between the 2′- and 3′-O-atoms. At the same time the P⋯P distance is shortened to *ca.* 5.8 Å. In RNA molecules, unlike DNA, there is a significant energy barrier between these two sugar conformations. Thus, RNA molecules retain the C(3′)-*endo* pucker conformation with a shortened distance between adjacent phosphate residues, unless the chain is under elongation stress. The C(2′)-*endo* pucker conformation is not found in the normal double-helical RNA segments.

Using the structure of the ribosome and the data on the path of mRNA through the ribosome, we can describe ribosomal frameshifting in terms of a sequence of three different stages.

Stage A. In this situation, the ribosome is translating the message, but the two codons of the slippery sequence have not yet arrived at the A site and P site of the ribosome. Instead, the 5′ or first of the slippery codon sites is engaged at the A site codon, but the second is not yet engaged. At that stage, there are a total of nine nucleotides found between the 3′-phosphate in the A site and the phosphate of residue C3 in the pseudoknot. These nine nucleotides are probably all in the C(3′)-*endo* conformation with an inter-phosphate distance of 5.8 Å. Thus, the nine nucleotides would have a total contour distance of slightly over 52 Å. At this stage, the pseudoknot has not yet wedged between the three proteins S3, S4, and S5 that flank the entrance of the tunnel through which the mRNA must thread (*Figs. 2* and *11*).

Stage B. The next state is one in which the mRNA has advanced by one codon, so that now both the A site and the P site of the ribosome are engaging the two codons of the slippery site. The pseudoknot is now fully engaged into the wedge-shaped entrance of the ribosomal tunnel, and the distance between the phosphate of residue C3 in the pseudoknot and the 5′-phosphate of the A site codon is 39 Å, as estimated from the structure of the ribosome. The six nucleotides in the spacer region now have an average P⋯P distance of 6.5 Å, which suggests that some of the nucleotides have already changed their conformation to the more-extended C(2′)-*endo* conformation with a 7-Å spacing between phosphate groups, while other nucleotides retain the more-stable C(3′)-*endo* conformation with a shorter distance between the nucleotides.

Stage C. At this point, the translocational machinery starts to move the 3′ slippery codon from the A site to the P site. The codon consists of three stacked nucleotides, each 3.4-Å apart, so the length of the codon is slightly over 10 Å. However, a kink (*ca.* 45°) is observed in the mRNA between codons in the A site and the P site [39] (*Fig. 11*, center), so the movement of the A-site codon to the P-site codon results in a net extension of 9 Å, rather than slightly over 10 Å. The bend in the mRNA is a

consequence of the fact that the tRNA anticodons are found at the end of double-stranded RNA anti-codon stems, which are themselves 20-Å in diameter, so the kink is necessary for steric reasons [46]. When this 9-Å extension starts to take place, the nucleotides of the linker region become fully extended with six nucleotides that can span a maximum distance of 42 Å. However, the movement of the translocational machinery must pull the message an additional 7 Å to complete the translocation. This is the defining moment for the role of the pseudoknot and frameshifting. The usual consequence is that the pseudoknot unravels when confronted with this considerable extension. That is, frameshifting efficiencies are typically from 1 to 30% in natural systems, so the most-common event is that the pseudoknot unravels. However, a minority of the pseudoknots slip the mRNA by one nucleotide, so that frameshifting occurs as a direct consequence of the attempted hyperextension of the linker region by the ribosomal translocation. The stress is reflected in the frameshifting event. Associated with frameshifting is the unraveling of the pseudoknot, but the sequence of these two events is unclear.

It is likely that a crucial element in the development of stress and its relaxation either by pseudoknot unraveling or frameshifting is related to the exact mode of docking of the pseudoknot in the ribosomal entry site. At present, we have no experimental information about how that is carried out, nor do we know, for example, that the mode of docking stays constant. It is entirely possible that a variety of different docking modes may be used by the same pseudoknot, and that some are more productive for frameshifting than others. This interpretation is strengthened by the extensive studies that were carried out on frameshifting efficiencies, both *in vitro* and *in vivo* by the BWYV pseudoknot and its mutants [36]. In the structure of the pseudoknot, the residue G19 projects in a direction close to that of the 5′-extension of the pseudoknot stem 1 (*Figs. 4* and *10*). G19, located at the junction between stem 1 and loop 2, is in a position where it will interact with the ribosome during this docking in the ribosomal tunnel. In those experiments the unmutated frameshifting efficiency was almost 11%. Changing G19 to U19 increased the frameshifting by *ca.* 30%, and inserting a cytosine after U19 increased the frameshifting almost 300%. Thus, slight changes of residues in the region of the pseudoknot that are involved in docking to the ribosome have a profound effect on frameshifting.

The above model describing the various stages of frameshifting uses the coordinates of the thermophilic ribosomal subunits, while the actual experiments cited were carried out with eukaryotic ribosomes. Although eukaryotic ribosomes have proteins that are comparable to the small unit proteins flanking the downstream entry site of thermophilic ribosomes, it is likely that the geometry may differ somewhat from that seen in the thermophilic ribosomes. However, this discussion shows the extent to which knowledge of the three-dimensional structure of both frameshifting viral pseudoknots and the ribosome will transform our thinking in the future as we try to understand the dynamics and mechanics of this important recoding system.

REFERENCES

[1] C. W. A. Pleij, L. Bosch, *Methods Enzymol.* **1989**, *180*, 289.
[2] C. W. A. Pleij, K. Rietveld, L. Bosch, *Nucleic Acids Res.* **1985**, *13*, 1717.
[3] C. W. Hilbers, P. J. A. Michiels, H. A. Heus, *Biopolymers* **1998**, *48*, 137.
[4] D. P. Giedroc, C. A. Theimer, P. L. Nixon, *J. Mol. Biol.* **2000**, *298*, 167.
[5] T. Powers, H. F. Noller, *EMBO J.* **1991**, *10*, 2203.
[6] T. Jacks, K. Townsley, H. E. Varmus, J. Majors, *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *94*, 4298.
[7] D. S. McPheeters, G. D. Stormo, L. Gold, *J. Mol. Biol.* **1988**, *201*, 517.
[8] E. S. Hass, J. W. Brown, C. Pitulle, N. R. Pace, *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 2527.
[9] E. Jabri, S. Aigner, T. R. Cech, *Biochemistry* **1997**, *36*, 16345.
[10] A. R. Ferré-D'Amaré, K. Zhou, J. A. Doudna, *Nature* **1998**, *395*, 567.
[11] D. Gilley, E. H. Blackburn, *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6621.
[12] E. B. ten Dam, K. Pleij, D. Draper, *Biochemistry* **1992**, *31*, 11665.
[13] C. Wilson, J. Nix, J. Szostak, *Biochemistry* **1998**, *37*, 14410.
[14] C. K. Tang, D. E. Draper, *Cell* **1989**, *57*, 531.
[15] C. Ehresmann, C. Philippe, E. Westhof, L. Bénard, C. Portier, B. Ehresmann, *Biochem. Cell. Biol.* **1995**, *73*, 1131.
[16] P. J. Farabaugh, *Microbiol. Rev.* **1996**, *60*, 103.
[17] R. F. Gesteland, J. F. Atkins, *Annu. Rev. Biochem.* **1996**, *65*, 741.
[18] Y.-X. Feng, T. D. Copeland, S. Oroszlan, A. Rein, J. G. Levin, *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 8860.
[19] N. M. Wills, R. F. Gesteland, J. F. Atkins, *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 6991.
[20] T. Jacks, H. D. Madhani, F. R. Masiarz, H. E. Varmus, *Cell* **1988**, *55*, 447.
[21] L. X. Shen, I. Tinoco Jr., *J. Mol. Biol.* **1995**, *247*, 963.
[22] Z. Du, D. P. Giedroc, D. W. Hoffman, *Biochemistry* **1996**, *35*, 4187.
[23] Z. Du, J. A. Holland, M. R. Hansen, D. P. Giedroc, D. W. Hoffman, *J. Mol. Biol.* **1997**, *270*, 464.
[24] H. Kang, I. Tinoco Jr., *Nucleic Acids Res.* **1997**, *25*, 1943.
[25] M. H. Kolk, M. van der Graaf, S. S. Wijmenga, C. W. A. Pleij, H. A. Heus, C. W. Hilbers, *Science* **1998**, *280*, 434.
[26] P. L. Nixon, A. Rangan, Y.-G. Kim, A. Rich, D. W. Hoffman, M. Hennig, D. P. Giedroc, *J. Mol. Biol.* **2002**, *322*, 621.
[27] L. Su, L. Chen, M. Egli, J. Berger, A. Rich, *Nat. Struct. Biol.* **1999**, *6*, 285.
[28] M. Egli, G. Minasov, L Su, A. Rich, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 4302.
[29] W. A. Miller, S. P. Dinesh-Kumar, C. P. Paul, *Crit. Rev. Plant Sci.* **1995**, *14*, 179.
[30] Web address: http://www.rcsb.org.
[31] a) R. Lavery, J. Sklenar, *J. Biomol. Struct. Dyn.* **1988**, 6, 63; b) R. Lavery, H. Sklenar, *J. Biomol. Struct. Dyn.* **1989**, *7*, 655.
[32] P. L. Nixon, D. P. Giedroc, *J. Mol. Biol.* **2000**, *296*, 659.
[33] P. L. Nixon, D. P. Giedroc, *Biochemistry* **1998**, *37*, 16116.
[34] V. Tereshko, C. J. Wilds, G. Minasov, M. A. Maier, T. Prakash, A. Howard, Z. Wawrzak, M. Manoharan, M. Egli, *Nucleic Acids Res.* **2001**, *29*, 1208.
[35] S. Basu, R. P. Rambo, J. Strauss-Soukup, J. H. Cate, A. Ferré-D'Amaré, S. A. Strobel, J. A. Doudna, *Nat. Struct. Biol.* **1998**, *5*, 986.
[36] Y.-G. Kim, L. Su, S. Maas, A. O'Neill, A. Rich, *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 14234.
[37] P. Somogyi, A. J. Jenner, I. Brierley, S. C. Inglis, *Mol. Cell. Biol.* **1993**, *13*, 6931.
[38] Y.-G. Kim, S. Maas, A. Rich, *Nucleic Acid Res.* **2001**, *29*, 1125.
[39] G. Z. Yusupova, M. M. Yusupov, J. H. D. Cate, H. F. Noller, *Cell* **2001**, *106*, 233.
[40] N. Ban, P. Nissen, J. Hansen, P. B. Moore, T. A. Steitz, *Science* **2000**, *289*, 902.
[41] J. Harms, F. Schluenzen, R. Zarivach, A. Bashan, S. Gat, I. Agmon, H. Bartels, F. Franceschi, A. Yonath, *Cell* **2001**, *107*, 679.
[42] B. T. Wimberly, D. E. Brodersen, W. M. Clemons Jr., R. J. Morgan-Warren, A. P. Carter, C. Vonrhein, T. Hartsch, V. Ramakrishnan, *Nature* **2000**, *407*, 327.
[43] F. Schluenzen, A. Tocilj, R. Zarivach, J. Harms, M. Gluehmann, D. Janell, A. Bashan, H. Bartels, I. Agmon, F. Franceschi, A. Yonath, *Cell* **2000**, *102*, 615.
[44] M. M. Yusupov, G. Z. Yusupova, A. Baucom, K. Lieberman, T. N. Earnest, J. H. D. Cate, H. F. Noller, *Science* **2001**, *292*, 883.

[45] W. Saenger, 'Principles of Nucleic Acid Structure', Springer Verlag, Berlin, 1984.
[46] A. Rich, in 'Ribosomes', Eds. M. Nomura, A. Tissieres, P. Lengyel, Cold Spring Harbor Laboratory, New York, 1974, pp. 871 – 884.