
A GCUA tetranucleotide loop found in the poliovirus *oriL* by in vivo SELEX (un)expectedly forms a YNMG-like structure: Extending the YNMG family with GYYA

WILLEM J.G. MELCHERS,¹ JAN ZOLL,¹ MARCO TESSARI,² DENIS V. BAKHMUTOV,^{3,5}
ANATOLY P. GMYL,³ VADIM I. AGOL,^{3,4} and HANS A. HEUS²

¹NCMLS, Department of Medical Microbiology, Radboud University Nijmegen, Nijmegen, The Netherlands

²IMM, Department of Biophysical Chemistry, Radboud University Nijmegen, Nijmegen, The Netherlands

³M.P. Chumakov Institute of Poliomyelitis & Viral Encephalitis, Russian Academy of Medical Sciences, Moscow, Russia

⁴M.V. Lomonosov Moscow State University, Moscow, Russia

ABSTRACT

The cloverleaf structure in the 5'-untranslated region of enterovirus RNA that regulates viral RNA replication contains an evolutionarily conserved YNMG tetraloop closed by a Y-G base pair. This loop is believed to interact specifically with the viral protease 3C. To further characterize the specificity of this interaction, the tetraloop and two flanking base pairs of the poliovirus RNA were randomized, and viable viral clones were obtained using in vivo SELEX. Among many different mutants with the canonical YNMG sequences to be described elsewhere, a large-plaque-forming clone contained a deviating uGCUAg sequence. The NMR structure of a small hairpin capped with uGCUAg that we present here shows that the GCUA tetraloop adopts a novel fold, which is highly similar to that of the YNMG tetraloop with common stacking properties and hydrogen-bond interactions including an unusual *syn* conformation of the adenosine. Thermodynamic studies show moderate stabilities of hairpins with canonical YNMG and the novel GCUA loops, which, together with the similarity of spatial structures, illustrates that the tetraloop structure itself is crucial for the RNA-protein interaction required for the viral replication. A re-evaluation of the ribosomal secondary structure database reveals a hairpin containing a GCUA loop, which covaries with YNMG and is involved in a tertiary interaction, and in the 50S ribosomal subunit from *Haloarcula marismortui* the structurally comparable apex of stem-loop 35a is a recognition site for protein L2. These observations show a more general occurrence and importance of the so-far unrecognized GYYA hairpin loops.

Keywords: RNA structure; RNA tetraloop; poliovirus; in vivo SELEX; RNA-protein recognition; NMR

INTRODUCTION

RNA tetranucleotide loops, also referred to as “tetraloops,” are the most common loops in biologically active RNAs and have been the most extensively studied by theoretical, molecular biology, biochemical, and biophysical approaches over the past two decades. The pioneering quintessential phylogenetic studies by Gutell, Woese, and collaborators (Gutell et al. 1985; Woese et al. 1990) have revealed that out of all loops of 16S rRNA, >50% were tetraloops, with GNRA and UNCG being the most abundant followed by

CUUG (N corresponds to any nucleotide, and R is either A or G). Since then, tetraloops were recognized as frequent and important elements in many other RNAs as well (Poritz et al. 1988; Pace et al. 1989; Michel and Westhof 1990). Thermodynamic studies revealed that hairpins with these tetraloops are also very stable (Antao et al. 1991), and because the majority of the UNCG and GNRA loops are closed by a C-G base pair, which also contributes to a higher stability (reversing the closing base pair of a UUCG loop to G-C destabilizes tetraloop formation by a $\Delta(\Delta G^{37^\circ})$ of 2.3 kcal/mol [Antao et al. 1991; Hilbers et al. 1994]), the sequences of the two families were subsequently refined to cUNCGg and cGNRAg (the lowercase letters represent the base pair closing the loop).

The most frequently occurring loops in 16S rRNA, i.e., cUUCGg, cGAAAg, and cGCAAg, were the first RNA structures solved at atomic resolution by NMR and revealed

⁵Deceased (August 20, 2004).

Reprint requests to: Hans A. Heus, IMM, Department of Biophysical Chemistry, Radboud University Nijmegen, Nijmegen, The Netherlands; e-mail: H.Heus@science.ru.nl; fax: 31-24-362112.

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.113106>.

To further characterize residue specificity of the D-loop–3CD interaction, we used an *in vivo* SELEX procedure, which, in principle, can sample the entire functional loop sequence space. The poliovirus tetraloop D and two flanking base pairs were randomized, yielding $4^8 = 65,536$ variants, and inserted into the infectious viral genome. The viable clones were isolated and investigated. The results of this study will be presented elsewhere. It should be noted, however, that along with numerous YNMG variants, a uGCUAg sequence was unexpectedly observed. This was surprising, indeed, because one could expect that this loop should share structural properties with the GNRA loop, which markedly differs from the YNMG structure, in particular by the presence of a sheared G–A base pair (Heus and Pardi 1991). Despite many biochemical, biophysical, and structural studies on RNA tetraloops, GNYA loops have escaped attention so far, probably because they are much less frequent in the ribosomal RNA and group I/II intron secondary structure models, which have typically been used for the phylogenetic studies.

Since there was no reference structure of a GNYA tetraloop at the start of these studies, we decided to determine its structure and thermodynamic properties. We find that the GCUA loop adopts a novel fold, highly similar to the YNMG motif, and exhibits a moderate stability. These results strongly suggest that it is the structure rather than stability that is primarily responsible for the selection of this loop in the viable viral clones. Importantly, the phylogenetic and structural comparison shows precedence for a more general role for GYYA loops in RNA folding and protein recognition.

RESULTS

Resonance assignment and structure determination

For the NMR studies, we used the non- $(^{13}\text{C}/^{15}\text{N})$ -labeled, chemically synthesized 14-nt RNA hairpin, composed of the selected poliovirus mutant apical D-loop, GCUA, and five flanking base pairs as depicted in Figure 1B. The small-size molecule gave excellent spectral dispersion, providing sufficient classic NOE distance and torsion angle information, which supplemented with modern implementation of RDC data allowed for a high-resolution structure determination without the need of a $^{13}\text{C}/^{15}\text{N}$ -labeling. Figure 2 gives some examples of the quality of the spectra. Proton resonance assignments were achieved using standard protocols, starting from sequential imino to imino and aromatic to H1'/H2' sequential walks. A strong mutual NOE between the U5 and G10 imino protons, resonating at 11.8 and 10.85 ppm, provided evidence for the presence of a U–G wobble pair. No imino proton resonance was found resonating at ~ 10.8 ppm, the typical position for a G imino proton in a sheared G–A base pair of the GNRA-like loop structure. A very large intraresidual H1'–H8 cross-peak was

observed for A9 (Fig. 2A), indicating a *syn* conformation. Other structural data consistent with a *syn* conformation were the presence of A9H2–A9H3', and U8H1', and the absence of A9H8–U8H1' NOE cross-peaks. These data indicated that the loop structure was different from that of the GNRA and was closed by a U–G wobble base pair. Besides the common twin amino proton resonances of cytidines in the 6.8–7.2-ppm region of the 2D NOESY recorded in H₂O, also broad resonances were found around 6.5 ppm, which could be attributed to the G6 and A9 amino protons. NOE cross-peaks involving these resonances, notably a G6H1'–A9N6H contact, and implementation of RDCs further supported a G–A(*syn*) noncanonical base pair between the first and the last loop nucleotides.

Analysis of the DQF-COSY and TOCSY spectra revealed that all residues adopt N-type sugar conformations, except for the loop C7 and U8 residues, where the intermediate–large 8–10-Hz H_{1'-2'} couplings indicated equilibria with predominant S-type sugars. C1' and aromatic carbon resonances were assigned using a natural abundance HSQC spectrum. With the excellent spectral dispersion and narrow linewidths (as a result of using a nonlabeled sample), RDCs were easily collected from a natural abundance ^{13}C -coupled 2D HSQC (Fig. 2B), using cryoprobe technology at 800 MHz. Statistics for the structure calculations are summarized in Table 1.

Structure of the uGCUAg tetraloop

The structure of the model hairpin consists of an A-form helix capped by a well-defined apical uGCUAg loop structure with pairwise RMSD values of 0.5 and 0.7 Å for the ensembles fitted to all residues, or only the loop and closing base-pair residues, respectively (Fig. 3). U5 and G10 clearly form a wobble base pair that stacks upon the lower A–U base pair with common helical parameters (Heus and Hilbers 2003). Enhanced stacking on the lower A–U base pair due to underwinding of the U–G base pair (25° vs. normal 33°) slightly widens the major groove. The first and last nucleotides of the tetraloop form an unusual G6–A9(*syn*) base pair, where the adenine amino proton is hydrogen-bonded to the guanosine N3 and the amino proton of the guanosine can possibly form a hydrogen bond with the N1 of adenine (Fig. 3C). This unusual G–A *trans* sugar edge/Watson–Crick base-pair conformation has not been described in the isostericity matrices of non-canonical base pairs yet (Leontis et al. 2002). Although no assumption or explicit hydrogen-bond restraints were introduced to force the hydrogen bonds in the G–A(*syn*) base pair, the GN3–AN6 heavy atom distances are within hydrogen-bond range (<2.8 Å) in all calculated structures and the GN2–AN1 distances (<3.8 Å) in 25% of the structures. The G6 2'-hydroxyl oxygen and U8 pro-R-phosphoryl oxygen are also within hydrogen-bond distance (<2.6 Å heavy atom distance in all structures), suggesting

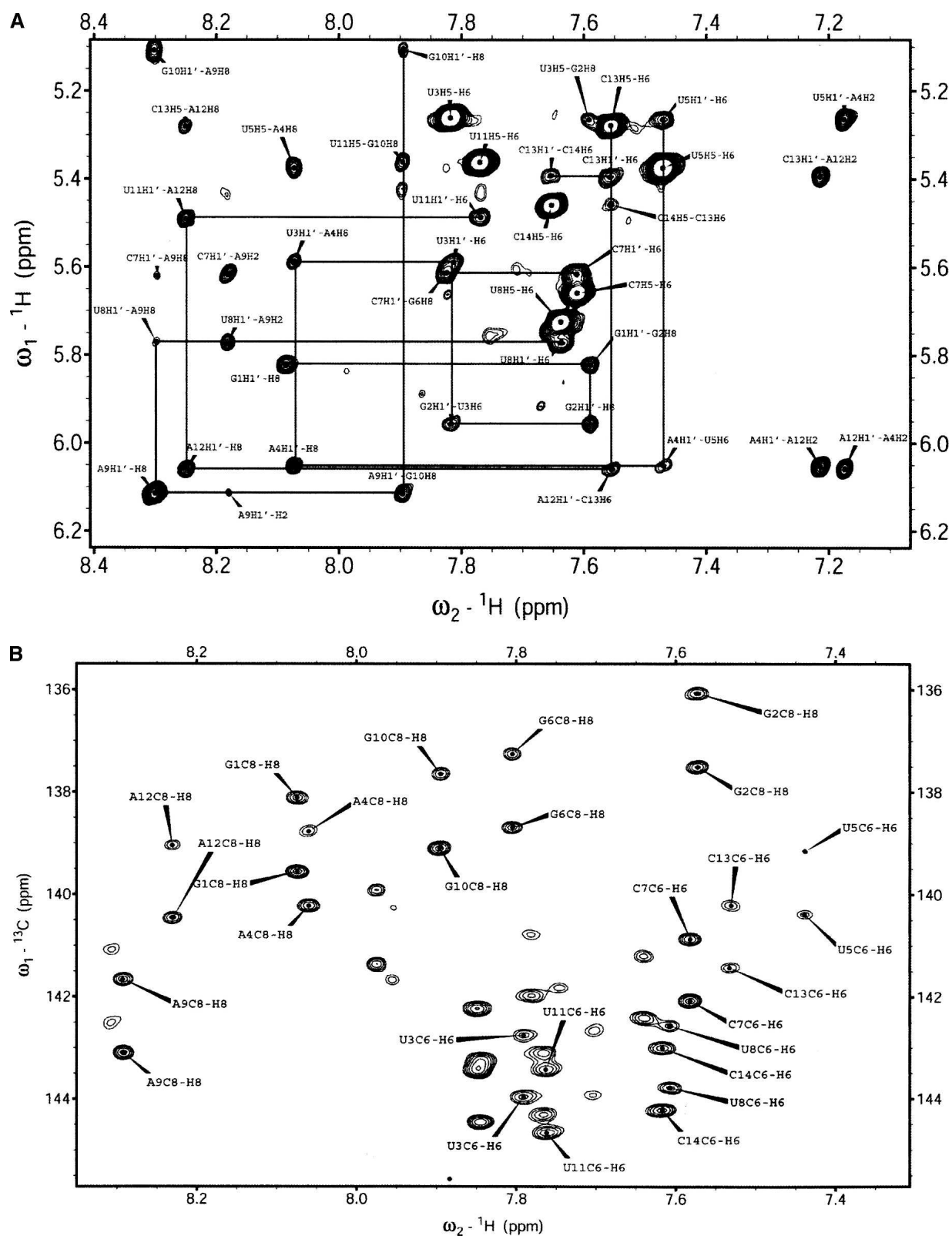


FIGURE 2. (A) Portion of the 2D NOESY (800 MHz, 300 msec mixing time) recorded in D₂O, showing the aromatic proton to H1' region. Sequential H1' to aromatic protons are traced out. (B) Portion of the ¹³C-coupled natural abundance 2D ¹H-¹³C HSQC showing the aromatic C8-H8/C6-H6 region used for collection of RCDs.

TABLE 1. Structural statistics of the final ensemble of 20 lowest-energy structures

Restrains	
Intranucleotide	32
Internucleotide	28
Hydrogen bonds	12
Dihedral angle restraints	103
Residual dipolar couplings	17
RMS deviations (Å, °)	
Distance restraints	0.101 ± 0.007
Dihedral restraints	1.0867 ± 0.1308
RDC	0.3509 ± 0.0298
RMS deviations from idealized	
Bonds (Å)	0.0096 ± 0.0002
Angles (°)	1.151 ± 0.021
Improper (°)	0.858 ± 0.011
Restraint violations	
No. of distance violations ^a >0.2 Å	0
No. of dihedral violations ^b >2°	0
Atomic RMS deviations (Å)	0.17 ± 0.09

^aNone larger than 0.5 Å.^bNone larger than 5°.

an additional hydrogen bond that further stabilizes the tetraloop structure. Reversal of the backbone direction is between the third and the fourth loop residues, as defined by the opposite orientations of the sugars. The second nucleotide of the loop, C7, swings its base into the minor groove with the H5–H6 edge facing the sugar moiety of G6 and the Watson-Crick edge exposed to solvent. The third nucleotide of the loop, U8, stacks upon G6 in the major groove with the H5–H6 pointing into the direction of the U8–5′-phosphate and the Watson-Crick edge exposed to solvent. Stacking interactions and observed hydrogen

bonds satisfactorily explain the many loop NOEs, absence of U7 and U8 imino proton resonances, and partial resolution of the G6 and A9 amino protons.

Thermodynamic studies

Thermodynamic parameters were obtained from the optical melting data for several minihairpins containing the poliovirus D-loop sequence, uUGCGg, and the loop sequence that was selected by in vivo SELEX, uGCUAg, flanked by the 5-bp stem used for the NMR construct. As references, we also determined thermodynamic data for the UUCG loop with C-G or U-G closing base pairs and the GCUA loop flanked by a different 4-bp stem, for which thermodynamic data have been reported earlier by Tinoco and coworkers (Antao and Tinoco 1991; Antao et al. 1991). All melting curves were obtained in 10 mM Na phosphate, 1 M NaCl, and 0.1 mM Na₂EDTA. For these conditions, the nearest-neighbor interactions have been tabulated by Turner and coworkers (Serra and Turner 1995), allowing calculation of stem stabilities. Thus, the ΔG^{37° values for loop formation shown in Table 2 were determined by subtracting the estimated stem stabilities from the hairpin values. Minihairpins with a UUCG loop closed by a C-G base pair, considered to be the classical unusually stable RNA loop, show high T_m 's of 75.2°C and 76.2°C, and the UUCG loop destabilizes the hairpin by only 1–1.2 kcal/mol. Hairpins with down to 7°C lower T_m 's and destabilizing ΔG^{37° up to 2 kcal/mol have been classified as unusually stable as well (Antao and Tinoco 1991; Antao et al. 1991; Hilbers et al. 1994; Proctor et al. 2002). The poliovirus UGCG and novel GCUA loops, closed by a U-G base pair, clearly fall outside this range, exhibiting much lower T_m 's and higher destabilizing ΔG^{37° . As can be inferred from

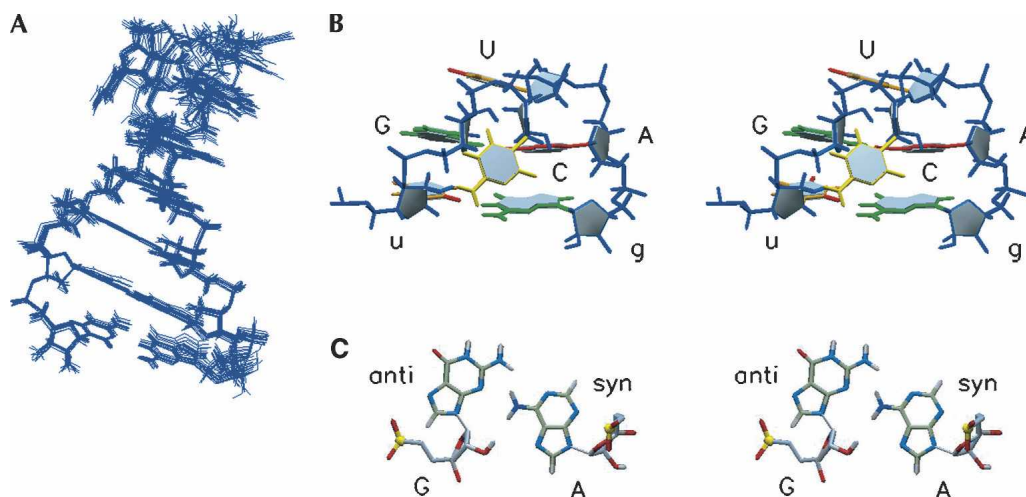


FIGURE 3. (A) Overlays of the 20 lowest-energy hairpin structures fitted to all residues. (B) Stereo view into the minor groove of the apical uGCUAg loop of the lowest energy structure. The A-residue bases are in red, the G-residue bases are in green, the C-residue bases are in yellow, and the U-residue bases are in orange. Uracil C–O bonds are in red. The sugar-phosphate backbone is in blue. (C) Stereo top view down the helical axis of the lowest energy structure, showing configuration of the G6–A9(*syn*) base pair.

TABLE 2. Thermodynamic parameters for the formation of RNA hairpins with tetraloops in 10 mM sodium phosphate, 1 M NaCl, 0.1 mM Na₂EDTA at pH 7.0

Sequence	T_m (°C)	ΔH (kcal/mol)	ΔS (e.u.)	ΔG^{37} (hairpin) (kcal/mol)	ΔG^{37} (loop) (kcal/mol)
GGUAU-GCUA-GUACC	54.3	-38.7	-118.2	-2.0	4.9
GGUAU-UUCG-GUACC	57.3	-56.1	-169.8	-3.4	3.7
GGUAU-UUCG-GUACC	59.7	-47.9	-143.9	-3.3	3.7
GGUAC-UUCG-GUACC	75.2	-63.7	-182.9	-7.0	1.2
GGUAU/GUACC ^a	—	-36.2	-94	-7	—
GGUAC/GUACC ^a	—	-40.7	-104.7	-8.2	—
GGAC-GCUA-GUCC	68.7	-40.1	-119.3	-3.8	3.5
GGAC-UUCG-GUCC ^b	76.2	-55.9	-159.9	-6.3	1
GGAC-UACG-GUCC ^b	73.8	-53.6	-154.5	-5.7	1.6
GGAC/GUCC ^a	—	-35.7	-91.4	-7.3	—

^aParameters for the RNA duplexes are calculated from Turner's rules (Serra and Turner 1995).

^bLiterature data (Antao and Tinoco 1991).

Table 2, the lower stability is primarily caused by changing the closing base pair from C-G to U-G, which has a profound influence on the hairpin stability, lowering the T_m by >10°C. Using the available data, the T_m of a 5-bp hairpin capped with a uUACGg sequence, present in five enteroviruses, is estimated to be 56.1°C, putting it in the class of moderately stable hairpins as well. The majority of the naturally occurring apical D-YNMG loops in the enterovirus cloverleaf structures are closed by a U-G base pair (75%), while the remainder is closed by a C-G base pair. Thus, it is reasonable to assume that the majority of these D-stem-loops are not unusually stable.

DISCUSSION

The GCUA loop is structurally a member of the YNMG rather than the GNRA family

The structure of the classical UUCG loop is the most stable and best-characterized member of the YNMG family and

has been determined at atomic resolution both by NMR and X-ray crystallography (Allain and Varani 1995; Ennifar et al. 2000). The structural details of the GCUA loop described here are the same as in the UUCG loop. The UUCG loop structure (PDB 1F7Y) (Fig. 4A) is characterized by a U-G(*syn*) reverse wobble (U-G *trans* sugar edge/Watson-Crick) base pair; folding of the second and third loop nucleotides into the minor groove and major groove sides, respectively; and stacking of the third nucleotide, C or U, on top of the U-G reverse wobble. On the other hand, the GAAA loop, i.e., the best structurally characterized member of the GNRA family (Jucker et al. 1996; Batey et al. 2001), clearly has a different fold with a sheared G-A (*trans* sugar edge/Hoogsteen) base pair, all nucleotides in the *anti* conformation, and stacking extending from the 3'-side up to the second (N) nucleotide, after which a sharp turn between the first and the second nucleotide reverses the chain direction (PDB 1HQ1) (Fig. 4B).

A comparison of Figures 3 and 4 reveals that the GCUA loop has the same topology as the UUCG loop and is

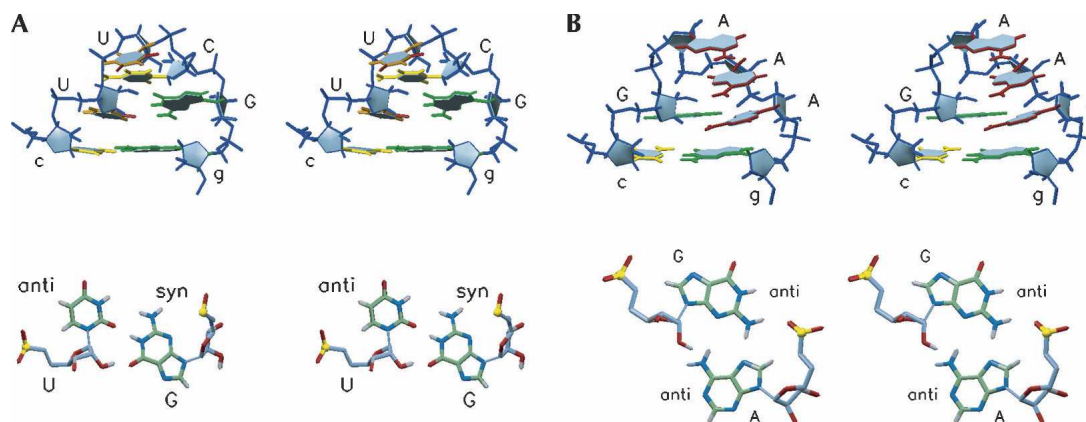


FIGURE 4. (A) Stereo view into the minor groove of the uUUCGg loop, and top view down the helix of the loop U-G(*syn*) base pair. (B) Stereo view into the minor groove of the cGAAA loop, and top view down the helix of the loop G-A base pair. Coloring is as in Figure 3.

markedly different from the GAAA loop. In particular, the stacking pattern with the second and third bases on the opposite sides of the backbone, and presence of a *trans* sugar edge/Watson-Crick base pair with a *syn* nucleotide provide for easily recognizable features indicating the structural equivalence. In a recent study, Gorlach and coworkers (Ihle et al. 2005) solved the structure of the bovine enterovirus 1 (BEV1) apical D-loop with a uGUUAg sequence, the sole naturally occurring enterovirus cloverleaf structure with a deviating GNYA sequence. However, no G-A(*syn*) base pair but rather a displaced pseudo-base pair without direct hydrogen bonds was found for the BEV1 GUUA loop (Ihle et al. 2005). Although we do not have a direct explanation for this difference, it might be a subtle sequence dependency of the second nucleotide (the second cytidine in the GCUA structure descends lower into the minor groove than the second uridine in the GUUA structure). Despite this subtle difference, the overall structures of the two loops are very similar with the characteristic folding of the second and third loop nucleotides and the fourth nucleotide *syn*.

Notwithstanding a great overall resemblance between the GCUA and UUCG loops, small differences are present. In the UUCG loop the amino group of the third loop nucleotide, C, forms a hydrogen bond with the second phosphate-pro-R-oxygen, which is absent in the GCUA loop. In the YNMG loop structures, the U-G or C-G base pairs can substitute for each other, preserving the structural motif because they are isosteric *trans* sugar edge/Watson-Crick base pairs. The GCUA loop also contains a *trans* sugar edge/Watson-Crick (G-A) base pair; however, it is not exactly isosteric with the Y-G base pairs due to a lateral shift in hydrogen-bond positions. Thus, in the UUCG loop, the 2'-hydroxyl of the first loop nucleotide, U, forms a hydrogen bond with the opposite guanosine O6-moiety, whereas in the GCUA loop, the 2'-hydroxyl of the first loop nucleotide, G, is hydrogen-bonded to the uridine pro-R-phosphoryl oxygen. G-A *trans* sugar edge/Watson-Crick base pairs more isosteric to Y-G, including a G2'O-ANH6 hydrogen bond, have been found in a distorted uGAAA(*syn*)Gg pentaloop of *Haloarcula marismortui* LSU RNA (Klein et al. 2004) and a (G-C)-A(*anti*) base triple in a frameshifting pseudoknot (Su et al. 1999). In the GCUA loop, the A(*syn*) is twisted more into the major groove than the G(*syn*) in the UUCG loop and does not lie directly over the next residue H1', which is in accordance with the normal chemical shift position of the G10 H1' resonance. In the case of the UUCG loop, a large upfield shift of ~ 1.5 ppm is observed, created by a large current effect of the loop G base on the next residue H1' (Varani et al. 1991). Also, the G6 nucleotide is twisted more into the major groove, shifting its 2'-hydroxyl to the proper position for donating a hydrogen bond. Overwinding of G6 and A9 narrows the cross-strand C1'-C1' distance to 8.5 Å in the GCUA loop and leaves the A9 N7-moiety less accessible than the

corresponding G N7 in the UUCG loop. Finally, the second loop nucleotide, C, descends more steeply into the minor groove, also rendering the minor groove side of the non-canonical G-A(*syn*) base pair less accessible to solvent. These subtle differences have consequences for the D-loop 3CD protein interaction and can be used to provide further understanding for the 3CD recognition. It is, for instance, important to note that the similarity between the UUCG and GCUA loops is larger on the major groove side than on the minor groove side.

The YNMG and GCUA loops closed by a U-G base pair are not unusually stable

The presence of YNMG and GNRA tetraloops closed with a C-G base pair found in the ribosomal RNAs has been correlated with unusual stability (Uhlenbeck 1990; Antao and Tinoco 1991; Antao et al. 1991; Hilbers et al. 1994; Proctor et al. 2002). To understand whether functional equivalence of the YNMG and the novel GCUA loops in the viral 5'-cloverleaf structures are due to structural or stability reasons or both, we determined the thermodynamic properties of several minihairpins containing wild-type and mutant D-loop sequences (Table 2). The results show that U-G base-pair-closed YNMG loops, representing the majority of D-loops in the enteroviral and rhinoviral 5'-cloverleaf structures, as well as the novel uGCUAg loop give rise to hairpins, which can be classified as moderately stable with T_m 's ranging from 54.3 to 57.3°C, while 25% of the D-loops are YNMG loops closed by a C-G base pair that yield unusually stable hairpins with T_m 's around 70°C. Thus, the unusual stability believed to be the key feature of the large number of YNMG and GNRA loops in the ribosomal RNAs is not the reason for the presence of YNMG and GCUA sequences in the enterovirus D-loops.

Phylogenetic comparison

The results described here clearly demonstrate a great structural resemblance between the YNMG and GCUA hairpin loops. This raises the question why nature appears to have selected against GNYA loops, which are virtually absent in the secondary structure models of the ribosomal and group I/II intron RNAs (<http://www.rna.icmb.utexas.edu>; Cannone et al. 2002). A survey of phylogenetic secondary structure models of 415 small subunit (SSU) rRNAs, 81 large subunit (LSU) rRNAs, and 20 group I and six group II introns reveals that GNYA loops are only present in hairpins 15, 39, and 41 of SSU rRNA, and hairpins 78 and 96 of LSU rRNA (Table 3). The vast majority of the hp15 sequences contain a GNRA tetraloop with only a few examples of YNMG or GNYA, and hp15 has a GNRA fold in the crystal structure of the ribosome (Schluzen et al. 2000; Wimberly et al. 2000). This suggests accidental occurrence of the YNMG and GNYA sequences in hp15

TABLE 3. Occurrence and phylogeny of GNYA hairpin loops in ribosomal RNA

Motif	Occurrence (no.)				
	SSU RNA ^a			LSU RNA ^a	
	Hairpin 15 ^b	Hairpin 39 ^c	Hairpin 41a ^d	Hairpin 78 ^d	Hairpin 96 ^d
GYYA	3	6	2	42	40
GRYA	1	0	29	0	2
YNMG	3	35	0	6	0
GNRA	249	8	281	0	6
Tetraloop	257	107	317	55	56
Others	12	182	14	2	1
Total no. of inspected sequences	269	289	331	57	57
Fold in ribosome X-ray structures	GNRA	YNMG	GNRA	N.D.	3'-stack

^aHairpins are numbered using standard classification.

^bAbsent (triloop) in eukaryotes (cytoplasm).

^cAbsent/fuzzy in organelles.

^dAbsent in mitochondria.

and that the function of hp15 can be fulfilled by structures not necessarily having a GNRA-like fold. Hp39 with six GYYA examples shows only a low abundance of YNMG and GNRA loops, while the majority have a different sequence showing no clear preference for or significance of either a YNMG or GNRA loop. Hp41a has a GNRA fold involved in a tertiary A-minor interaction in the crystal structures of the ribosome (Nissen et al. 2001). Again, the vast majority of these loops are represented by GNRA, with no YNMG and only two GYYA examples. There are, however, 29 examples of a GRYA loop, which, given the structural importance of the GNRA fold to form an A-minor interaction, suggests that the GRYA loops might adopt a GNRA fold as well. The LSU hp78 is of particular interest for this discussion (Table 4). Here, the majority is represented by a GYYA (GCCA, GCUA, or GUCA) sequence, which covaries with YNMG. Most interestingly, the loop sequences also covary with a proposed tertiary interaction involving the 5'-nucleotide of loop 77: hp78's with GYYA loops are proposed to interact with a U at the 5'-side of loop 77, while hp78's with YNMG loops are proposed to interact with a C. This regularity highly suggests a structural role for these loops at this site, in line with our proposal that they have a common fold. Unfortunately, neither hairpin 78 nor its interacting nucleotide is resolved in the available crystal structures of the large ribosomal subunit (Harms et al. 2001; Klein et al. 2004). However, on the basis of the UUCG and GCUA loop structures, a (U-G)-C or (G-A)-U base triple with the third base making a major groove interaction is most likely. Hp96 also contains a large number of GYYA loops, in this case covarying with GNRA. This is unlike the situation with hp78. However, inspection of the crystal structures reveals a distorted loop structure, uncommon to either GNRA or YNMG, with the four loop nucleotides forming a continuous stack from the 3'-side of the helix (Harms et al. 2001; Klein et al. 2004).

As mentioned above, a likely reason for the low abundance of GNYA loops in the ribosomal and group I/II introns is that YNMG loops evolved there primarily because of their unusual stability to serve, for instance, as nucleation sites for the folding of larger RNAs. On the other hand, GCUA loops, when present, might have evolved predominantly for structural reasons. The available structural and phylogenetic data—that is, (1) GCUA and GUUA loops are structurally similar to YNMG loops, and (2) a YNMG loop with a possible structural role covaries with GCUA, GUCA, and GCCA—provide evidence for extending the YNMG family with GYYA.

Implications for 3CD binding

A wide body of evidence is available that shows that a direct interaction between the cloverleaf apical D-loop and 3CD protein is essential for poliovirus replication. Earlier phylogenetic, mutational, and structural (inferred from the high-resolution NMR) data have mapped the RNA-binding

TABLE 4. Occurrence and sequence alignment of hairpin 78 in 23S rRNA

Motif	Sequence	Occurrence (no.)
GYYA	cGCCAg	19
	cGCUAg	16
	cGUCAg	4
	cGUCA-	1
	uGCUAg	1
YNMG	gGCCAg	1
	cUUCGg	2
	uUCCGg	2
	cUCCGg	1
	cCUCGg	1

site to the apical D-loop with a common YNMG-like structure (Andino et al. 1990; Zell et al. 2002; Du et al. 2003; Ohlenschlager et al. 2004). The GCUA loop structure can be used to further evaluate the specificity of binding.

Despite the compact structure of the loop with numerous hydrogen-bond interactions, many hydrogen-bond acceptors and donors remain available for protein binding (Figs. 3,5A). On the major groove side, the amino, imino, O6, and N7 of G6; the N3 of A9; and the three Watson-Crick sites of U8 (O2, O4, and N3) form a cluster of accessible sites. On the minor groove side, the three Watson-Crick sites of C7 (amino, O2, and N3) are available, and possibly the A9 N7, although this site seems to be sequestered by the overwinding, and the C7 base that descends steeply into the minor groove in most of the structures. Also, the various backbone phosphorous, sugar oxygens and 2'-hydroxyls are exposed to solvent.

Two motifs of the 3C protein have been implicated in RNA binding, the so-called $_{154}\text{TGK}_{156}$ and $_{82}\text{KFRDI}_{86}$ motifs (Andino et al. 1990, 1993; Walker et al. 1995). In our in vivo SELEX study, some small plaque viruses containing a GNYA loop acquired a compensatory mutation after some passages, T_{154} to I, again showing a direct relation between the apical D-loop and TGK motif (D.V. Bakhmutov, A.P. Gmyl, E.V. Khitrina, M.M. Erokhin, J. Zoll, W.J.G. Melchers, H.A. Heus, and V.I. Agol, in prep.). Studies of engineered 3C mutants revealed several residues outside the known TGK and KFRDIR motifs, which might be involved in RNA binding (Andino et al. 1993). The protein residues implicated in RNA binding are indicated in Figure 5B. On the available high-resolution structure of 3C obtained by X-ray crystallography (Mosimann et al. 1997), the residues

map to the presumed RNA-binding site that involves a cluster of the TGK and KFRDIR motifs, and amino acids from the domain linker, the amphipatic helices C and D, and several β -hairpin turns on the protein surface.

The crystal structure of 3C, together with the NMR structures of the apical D-loops, provide for a framework to further understand the specificity of binding. For instance, it has been suggested that, although the second nucleotide of the YNMG loops that folds into the minor groove is not relevant for the isolated hairpin loop structures, its sequence conservation (with 52% A and 23% G) might indicate possible involvement in binding (Du et al. 2004). However, the in vivo SELEX study yielded a GCUA loop, and the BEV1 D-loop contains a uridine at the second loop position. Thus, the N-position may be occupied by all four possible residues, which adopt an *anti* conformation as deduced from the NMR structures. This gives a variety of hydrogen-bond donors or acceptors at comparable sites, making a specific role of the second (N) base less likely. Using a larger CVB3 D-stem construct including the three noncanonical pyrimidine base pairs, specific residues have been implicated in binding on the basis of NMR line broadening and chemical shift changes upon protein 3C addition (Ohlenschlager et al. 2004). Specifically, it was suggested that the minor groove side of the tetraloop together with the major groove side at the C54–C71 region (Fig. 1A) form a bipartite interdependent recognition site for 3C. However, in the GCUA structure, overwinding of the G6 and G9 residues and steep descent of the second variable nucleotide cover possible backbone hydrogen-bond donors and acceptors on the minor groove side of the tetraloop, which is more in favor of the minor groove

side being silent in recognition. Also, the similarity between the UUCG and GCUA loops is larger on the major groove side than on the minor groove side. Therefore, the importance for binding of the major groove side of the tetraloop, with a cluster of more accessible hydrogen-bond donors and acceptors, should not be overlooked. The 872uGAAAGg878 pentaloop present in hp35a of *H. marismortui* LSU RNA closely resembles the GCUA loop. Here also a G-A *trans* sugar edge/Watson-Crick base pair is formed, and the fifth base, G879, is bulged out, leaving a tetraloop with similar stacking properties as the GCUA loop, be it that the second loop base, A874, is pointing away from the minor groove to form a tertiary stacking contact with A857. Ribosomal protein L2 makes several specific contacts with hp35a in the LSU crystal structure (Klein et al. 2004). Arg2 forms

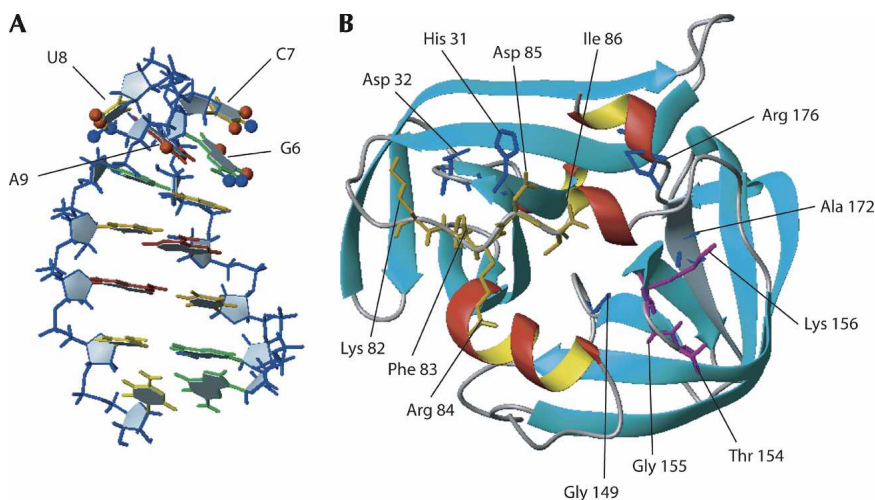


FIGURE 5. (A) Side view of the RNA hairpin with GCUA tetraloop, with possible hydrogen-bond acceptors and donors of the tetraloop indicated by red and by blue spheres, respectively. (B) Ribbon diagram of the crystal structure of 3C with the presumed RNA-binding site facing the viewer. RNA bases and protein amino acids implicated in RNA-protein binding are annotated. The TGK motif is in magenta, the KFRDIR motif is in gold, and other residues implicated in binding are in blue.

hydrogen bonds with G871-O6, U872-O4, and G878-O6; Arg8, with G873-O6; and Asp198, with the G873 imino and amino protons. All these contacts are on the major groove side of the hairpin loop and two flanking base pairs. Possibly similar interactions can be fulfilled by comparable residues of the stem-loop D apex, in agreement with the importance of its major groove side for 3C recognition.

Figure 5 gives an impression of the spatial relationship between the compact tetraloop that might snugly fit into a tight binding pocket of the 3C protein, but also reveals that the size of the D-loop region implicated in binding (ranging from the tetraloop down to C54) probably exceeds the size of the presumed RNA-binding site of 3C. Clearly, more studies are needed to address these issues.

In summary, the structural data presented here further explain the sequence requirements of the apical D-loop and support the notion that the tetraloop structure is the primary determinant for 3C recognition. This can be used to direct future studies on the role of the 5'-cloverleaf-3CD interaction in enterovirus RNA replication. The novel structure prompted us to re-evaluate the secondary structure database, and the phylogenetic comparison revealed another example of a GYYA loop with a structural role, possibly involved in a triple interaction. Also, the structurally comparable RNA stem-loop 35a from *H. marismortui* 50S ribosomal subunit is a major recognition site in L2 protein binding. These observations show precedence for a more widespread occurrence and importance of the GCUA fold we describe here. The structures of the GCUA and GCAA loops are very different. This exemplifies again the extreme structural flexibility of RNA, where a 1-nt change can make a huge structural difference, making it virtually impossible to propose a reliable structure from sequence alone, when noncanonical interactions are involved.

MATERIALS AND METHODS

Sample preparation

The 14-mer RNA oligonucleotide was purchased from ABI GmbH. The final concentration of the RNA samples was 0.5 mM. The pH was adjusted to 6.8 (meter reading). Weakly aligned NMR samples were prepared by adding 27 mg/mL filamentous phage Pfl (Hansen et al. 1998) purchased from ASLA Ltd.

NMR spectroscopy

NMR spectra were acquired on Varian Inova 600 MHz and 800 MHz spectrometers. For the assignment of exchangeable imino and amino protons, 2D NOESY spectra and natural abundance ^1H - ^{15}N HMQC were recorded in 90% H_2O , 10% D_2O at 15°C and 800 MHz using a jump return pulse for water suppression (Plateau and Guéron 1982). Other spectra were recorded at 25°C in D_2O at 600 and 800 MHz. Nonexchangeable protons were assigned using 2D NOESY (50, 100, and 300 msec mixing times) (Jeener et al. 1970), DQF-COSY (Shaka and

Freeman 1983), and TOCSY (Griesinger et al. 1988). Residual dipolar couplings (RDCs) were collected by measuring heteronuclear ^1H - ^{13}C one-bond couplings in the absence or presence of filamentous phage Pfl using a ^1H -proton coupled 2D-HSQC or the IPAP approach (Ottiger et al. 1998) at 600 MHz with cryoprobe technology. Typical acquisition times for the IPAP spectra at natural abundance were ~24 h with 48 scans per FID, 1024 t_2 data points, and 320 complex t_1 data points. Different data sets were collected involving the aromatic H8, H6, and H2, and the H1' and H5 with the INEPT-delay set to 2.5 and 2.9 msec, respectively. All data were analyzed with Sparky (T.D. Goddard and D.G. Kneller, University of California, San Francisco).

Structural restraints

Distances involving the nonexchangeable protons were calculated from NOE cross-peak intensities of the 2D NOESYs with 50, 100, and 300 msec mixing times using the modified isolated spin pair approximation approach (Barsukov and Lian 1993). The calculated distances were used with $\pm 20\%$ error bounds. Distance information involving exchangeable protons was obtained from 2D NOESY spectra recorded in H_2O . Hydrogen-bonding restraints were treated as NOE distance restraints and imposed on residues involved in Watson-Crick base-pairing, evidenced by the imino proton spectra and observed NOEs and chemical shifts.

Torsion angle restraints for the sugar moieties were obtained from DQF-COSY and TOCSY experiments, and interpreted as N or S-type sugar puckers when they were small (<3 Hz) or large (>8 Hz), respectively. Dihedral torsion restraints were applied as previously described (Lescrier et al. 2003). For the stem residues, the ϵ dihedral angle was restrained to $225^\circ (\pm 60^\circ)$. The α and ζ dihedral angles were constrained from chemical shift considerations of the phosphorous resonance to $0^\circ (\pm 120^\circ)$. RDCs were used with 20% error bounds.

Structure calculations

Structure calculations were performed with the torsion angle dynamics protocol implemented within XPLOR-NIH (Stein et al. 1997; Schwieters et al. 2003) and included simulated annealing starting from extended strands. Out of starting sets of 100 randomly extended structures, 20 acceptable structures were obtained with no dihedral angle violations $>5^\circ$ and no NOE violations $>0.5 \text{ \AA}$ (Table 1) and subjected to refinement using RDCs. Initial values for the axial and rhombic components of the alignment tensor were estimated from the distribution of the RDC values (Clare et al. 1998; Bax et al. 2001). The subsequent grid search yielded optimal values of $Da = -50 \text{ Hz}$ and $R = 0.4 \text{ Hz}$. Da was varied from -35 Hz to -55 Hz , and R was varied from 0 Hz to 0.6 Hz in steps of 1 Hz and 0.1 Hz , respectively. The correlation coefficient for the back-calculated RDCs increased from 0.46 to 0.997 after inclusion of the RDCs in the structure calculations. Color figures were generated using MOLMOL (Koradi et al. 1996). Helical parameters were calculated using the programs 3DNA (Lu and Olson 2003) or freehelix98 (Dickerson and Chiu 1997; Dickerson 1998).

UV melting curves and thermodynamic parameters

UV absorbance melting curves were recorded at both 260 and 280 nm simultaneously on a Cary-300 spectrometer. Samples

($A_{260} = 0.7\text{--}1.5 \text{ cm}^{-1}$; corresponding to $\sim 7\text{--}15 \text{ }\mu\text{M}$) were prepared in 10 mM sodium phosphate, 1 M NaCl, and 0.1 mM Na_2EDTA . The samples were heated a few minutes to 90°C and quickly cooled to 1°C . The melting curves were measured at a heating rate of $0.5^\circ\text{C}/\text{min}$ from 15°C to 90°C with data collected every 0.25°C increment. The melting curves were run in triplicate at two concentrations and showed identical profiles indicating reversible folding and single unimolecular transitions corresponding to unfolding of hairpins. Melting temperatures (T_m) and thermodynamic parameters were derived assuming a two-state model using the Cary-WIN software. Enthalpy values and other thermodynamic parameters were calculated from the first derivative of the absorbance with respect to temperature, using the equation (Gralla and Crothers 1973):

$$\Delta H = \frac{-3.2}{\frac{1}{T_{1/2}} - \frac{1}{T_{3/4}}}$$

In this equation, $T_{1/2}$ is the temperature at the maximum, which is equal to the melting temperature T_m , and $T_{3/4}$ is the temperature at the upper half-height of the derivative of the melting curve.

Atomic coordinates

Coordinates of the ensemble of lowest energy structures have been deposited in the RCSB Protein Data Bank (accession code 2EVY).

ACKNOWLEDGMENTS

This research was supported by grants of the Council of Chemical Sciences of the Netherlands Organization for Scientific Research (NWO-CW 700.50.021 and NWO-RFBR 047.017.023) and the European Communities (INTAS). The Moscow group was also supported by the Russian Foundation for Basic Research and the Program for Support of Leading Scientific Schools.

Received April 11, 2006; accepted June 10, 2006.

REFERENCES

- Ackerman, E.J., Saxena, S.K., and Ulbrich, N. 1988. α -Sarcin causes a specific cut in 28S rRNA when microinjected into *Xenopus* oocytes. *J. Biol. Chem.* **263**: 17076–17083.
- Allain, F.H. and Varani, G. 1995. Structure of the P1 helix from group I self-splicing introns. *J. Mol. Biol.* **250**: 333–353.
- Andino, R., Rieckhof, G.E., Trono, D., and Baltimore, D. 1990. Substitutions in the protease (3Cpro) gene of poliovirus can suppress a mutation in the 5' noncoding region. *J. Virol.* **64**: 607–612.
- Andino, R., Rieckhof, G.E., Achacoso, P.L., and Baltimore, D. 1993. Poliovirus RNA synthesis utilizes an RNP complex formed around the 5'-end of viral RNA. *EMBO J.* **12**: 3587–3598.
- Antao, V.P. and Tinoco Jr., I. 1991. Thermodynamic parameters for loop formation in RNA and DNA hairpin tetraloops. *Nucleic Acids Res.* **21**: 5901–5905.
- Antao, V.P., Lai, S.Y., and Tinoco Jr., I. 1991. Thermodynamic parameters for loop formation in RNA and DNA hairpin tetraloops. *Nucleic Acids Res.* **20**: 819–824.
- Barsukov, I.L. and Lian, L.Y. 1993. Structure determinations from NMR data I. Analysis of NMR data. In *NMR of macromolecules. A practical approach* (ed. G.C.K. Roberts), pp. 315–357. Oxford University Press, New York.
- Batey, R.T., Sagar, M.B., and Doudna, J.A. 2001. Structural and energetic analysis of RNA recognition by a universally conserved protein from the signal recognition particle. *J. Mol. Biol.* **307**: 229–246.
- Bax, A., Kontaxis, G., and Tjandra, N. 2001. Dipolar couplings in macromolecular structure determination. *Methods Enzymol.* **339**: 127–174.
- Cai, Z., Gorin, A., Frederick, R., Ye, X., Hu, W., Majumdar, A., Kettani, A., and Patel, D.J. 1998. Solution structure of P22 transcriptional antitermination N peptide-boxB RNA complex. *Nat. Struct. Biol.* **5**: 203–212.
- Cannone, J.J., Subramanian, S., Schnare, M.N., Collett, J.R., D'Souza, L.M., Du, Y., Feng, B., Lin, N., Madabusi, L.V., Muller, K.M., et al. 2002. The comparative RNA Web (CRW) site: An online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics* **3**: 2.
- Cate, J.H., Gooding, A.R., Podell, E., Zhou, K., Golden, B.L., Kundrot, C.E., Cech, T.E., and Doudna, J.A. 1996. Crystal structure of a group I ribozyme domain: Principles of RNA packing. *Science* **273**: 1678–1685.
- Cheong, C., Varani, G., and Tinoco Jr., I. 1990. Solution structure of an unusually stable RNA hairpin, 5'GGAC(UUCG)GUCC. *Nature* **346**: 680–682.
- Clore, G.M., Gronenborn, A.M., and Tjandra, N. 1998. Direct structure refinement against residual dipolar couplings in the presence of rhombicity of unknown magnitude. *J. Magn. Reson.* **131**: 159–162.
- Dickerson, R.E. 1998. DNA bending: The prevalence of kinkiness and the virtues of normality. *Nucleic Acids Res.* **26**: 1906–1926.
- Dickerson, R.E. and Chiu, T.K. 1997. Helix bending as a factor in protein/DNA recognition. *Biopolymers* **44**: 361–403.
- Du, Z., Yu, J., Andino, R., and James, T.L. 2003. Extending the family of UNGC-like tetraloop motifs: NMR structure of a CACG tetraloop from coxsackievirus B3. *Biochemistry* **42**: 4373–4383.
- Du, Z., Yu, J., Ulyanov, N.B., Andino, R., and James, T.L. 2004. Solution structure of a consensus stem loop D RNA domain that plays important roles in regulating translation and replication in enteroviruses and rhinoviruses. *Biochemistry* **43**: 11959–11972.
- Ennifar, E., Nikulin, A., Tishchenko, S., Serganov, A., Nevskaya, N., Garber, M., Ehresmann, B., Ehresmann, C., Nikonov, S., and Dumas, P. 2000. The crystal structure of UUCG tetraloop. *J. Mol. Biol.* **304**: 35–42.
- Gamarnik, A.V. and Andino, R. 1997. Two functional complexes formed by KH domain containing proteins with the 5' noncoding region of poliovirus RNA. *RNA* **3**: 882–892.
- . 1998. Switch from translation to RNA replication in a positive-stranded RNA virus. *Genes & Dev.* **12**: 2293–2304.
- Gralla, J., and Crothers, D.M. 1973. Free energy of imperfect nucleic acid helices. III. Small internal loops resulting from mismatches. *J. Mol. Biol.* **78**: 301–319.
- Griesinger, C., Otting, G., Wüthrich, K., and Ernst, R.R. 1988. Clean TOCSY for ^1H spin system identification in macromolecules. *J. Am. Chem. Soc.* **110**: 7870–7872.
- Gutell, R.R., Weiser, B., Woese, C.R., and Noller, H.F. 1985. Comparative anatomy of 16-S-like ribosomal RNA. *Prog. Nucleic Acid Res. Mol. Biol.* **32**: 155–216.
- Hansen, M.R., Mueller, L., and Pardi, A. 1998. Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nat. Struct. Biol.* **5**: 1065–1074.
- Harms, J., Schluenzen, F., Zarivach, R., Bashan, A., Gat, S., Agmon, I., Bartels, H., Franceschi, F., and Yonath, A. 2001. High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. *Cell* **107**: 679–688.
- Heus, H.A. and Hilbers, C.W. 2003. Structures of non-canonical tandem base pairs in RNA helices: Review. *Nucleosides Nucleotides Nucleic Acids* **22**: 559–571.

- Heus, H.A. and Pardi, A. 1991. Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops. *Science* **253**: 191–194.
- Hilbers, C.W., Heus, H.A., van Dongen, M.J.P., and Wijmenga, S.S. 1994. The hairpin elements of nucleic acid structure: DNA and RNA folding. In *Nucleic acids and molecular biology* (eds. F. Eckstein and D.M.J. Lilley), Vol. 8. pp. 56–104. Springer Verlag, Berlin.
- Ihle, Y., Ohlenschlager, O., Hafner, S., Duchardt, E., Zacharias, M., Seitz, S., Zell, R., Ramachandran, R., and Gorlach, M. 2005. A novel cGUUAg tetraloop structure with a conserved yNYMGg-type backbone conformation from cloverleaf 1 of bovine enterovirus 1 RNA. *Nucleic Acids Res.* **33**: 2003–2011.
- Jeener, J., Meier, B.H., Bachmann, P., and Ernst, R.R. 1970. Investigation of exchange processes by two dimensional NMR spectroscopy. *J. Chem. Phys.* **71**: 4546–4553.
- Jucker, F.M., Heus, H.A., Yip, P.F., Moors, E.H., and Pardi, A. 1996. A network of heterogeneous hydrogen bonds in GNRA tetraloops. *J. Mol. Biol.* **264**: 968–980.
- Klein, D.J., Moore, P.B., and Steitz, T.A. 2004. The roles of ribosomal proteins in the structure assembly, and evolution of the large ribosomal subunit. *J. Mol. Biol.* **340**: 141–177.
- Koradi, R., Billeter, M., and Wüthrich, K. 1996. MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**: 51–55.
- Legault, P., Li, J., Mogridge, J., Kay, L.E., and Greenblatt, J. 1998. NMR structure of the bacteriophage λ N peptide/boxB RNA complex: Recognition of a GNRA fold by an arginine-rich motif. *Cell* **93**: 289–299.
- Leontis, N.B., Stombaugh, J., and Westhof, E. 2002. The non-Watson-Crick base pairs and their associated isostericity matrices. *Nucleic Acids Res.* **30**: 3497–3531.
- Lescrinier, E.M.H.P., Tessari, M., van Kuppeveld, F.J.M., Melchers, W.J.G., Hilbers, C.W., and Heus, H.A. 2003. Structure of the pyrimidine-rich internal loop in the poliovirus 3'-UTR: The importance of maintaining pseudo-2-fold symmetry in RNA helices containing two adjacent non-canonical base-pairs. *J. Mol. Biol.* **331**: 759–769.
- Lu, X. and Olson, W.K. 2003. 3DNA: A software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* **31**: 5108–5121.
- Lyons, T., Murray, K.E., Roberts, A.W., and Barton, D.J. 2001. Poliovirus 5'-terminal cloverleaf RNA is required in *cis* for VPg uridylylation and the initiation of negative-strand RNA synthesis. *J. Virol.* **75**: 10696–10708.
- Michel, F. and Westhof, E. 1990. Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J. Mol. Biol.* **216**: 585–610.
- Mosimann, S.C., Cherney, M.M., Sia, S., Plotch, S., and James, M.N. 1997. Refined X-ray crystallographic structure of the poliovirus 3C gene product. *J. Mol. Biol.* **273**: 1032–1047.
- Murphy, F.L. and Cech, T.R. 1994. GAAA tetraloop and conserved bulge stabilize tertiary structure of a group I intron domain. *J. Mol. Biol.* **236**: 49–63.
- Nissen, P., Ippolito, J.A., Ban, N., Moore, P.B., and Steitz, T.A. 2001. RNA tertiary interactions in the large ribosomal subunit: The A-minor motif. *Proc. Natl. Acad. Sci.* **98**: 4899–4903.
- Ohlenschlager, O., Wohnert, J., Bucci, E., Seitz, S., Hafner, S., Ramachandran, R., Zell, R., and Gorlach, M. 2004. The structure of the stemloop D subdomain of coxsackievirus B3 cloverleaf RNA and its interaction with the proteinase 3C. *Structure* **12**: 237–248.
- Ottiger, M., Delaglio, F., and Bax, A. 1998. Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson.* **131**: 373–378.
- Pace, N.R., Smith, D.K., Olsen, G.J., and James, B.D. 1989. Phylogenetic comparative analysis and the secondary structure of ribonuclease P RNA—A review. *Gene* **82**: 65–75.
- Parsley, T.B., Towner, J.S., Blyn, L.B., Ehrenfeld, E., and Semler, B.L. 1997. Poly (rC) binding protein 2 forms a ternary complex with the 5'-terminal sequences of poliovirus RNA and the viral 3CD proteinase. *RNA* **3**: 1124–1134.
- Plateau, P. and Guéron, M. 1982. Exchangeable proton NMR without baseline distortion, using new strong-pulse sequences. *J. Am. Chem. Soc.* **104**: 7310–7311.
- Poritz, M.A., Strub, K., and Walter, P. 1988. Human SRP RNA and *E. coli* 4.5S RNA contain a highly homologous structural domain. *Cell* **55**: 4–6.
- Proctor, D.J., Schaak, J.E., Bevilacqua, J.M., Falzone, C.J., and Bevilacqua, P.C. 2002. Isolation and characterization of a family of stable RNA tetraloops with the motif YNMG that participate in tertiary interactions. *Biochemistry* **41**: 12062–12075.
- Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F., et al. 2000. Structure of functionally activated small ribosomal subunit at 3.3 Å resolution. *Cell* **102**: 615–623.
- Schwieters, C.D., Kuszewski, J.J., Tjandra, N., and Clore, G.M. 2003. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**: 66–74.
- Semler, B.L. and Wimmer, E., eds. 2002. *Molecular biology of picornaviruses*. ASM Press, Washington, DC.
- Serra, M.J. and Turner, D.H. 1995. Predicting thermodynamic properties of RNA. *Methods Enzymol.* **259**: 242–261.
- Shaka, A.J. and Freeman, R. 1983. Simplification of NMR spectra by filtration through multiple quantum coherence. *J. Magn. Reson.* **51**: 169–173.
- Stein, E.G., Rice, L.M., and Brünger, A.T. 1997. Torsion-angle molecular dynamics as a new efficient tool for NMR structure calculation. *J. Magn. Reson.* **124**: 154–164.
- Su, L., Chen, L., Egli, M., Berger, J.M., and Rich, A. 1999. Minor groove RNA triplex in the crystal structure of a ribosomal frameshifting viral pseudoknot. *Nat. Struct. Biol.* **6**: 285–292.
- Uhlenbeck, O.C. 1990. Tetraloops and RNA folding. *Nature* **346**: 613–614.
- van Knippenberg, P.H. 1986. Structural and functional aspects of the N6,N6 dimethyladenosines in 16S ribosomal RNA. In *Structure, function and genetics of ribosomes* (eds. B. Hardesty and B. Kramer), pp. 412–424. Springer-Verlag, New York.
- Varani, G., Cheong, C., and Tinoco Jr., I. 1991. Structure of an unusually stable RNA hairpin. *Biochemistry* **30**: 3280–3289.
- Walker, P.A., Leong, L.E., and Porter, A.G. 1995. Sequence and structural determinants of the interaction between the 5'-non-coding region of picornavirus RNA and rhinovirus protease 3C. *J. Biol. Chem.* **270**: 14510–14516.
- Wimberly, B.T., Brodersen, D.E., Clemons Jr., W.M., Morgan-Warren, R.J., Carter, A.P., Vornrhein, C., Hartsch, T., and Ramakrishnan, V. 2000. Structure of the 30S ribosomal subunit. *Nature* **407**: 327–339.
- Woese, C.R., Winker, S., and Gutell, R.R. 1990. Architecture of ribosomal RNA: Constraints on the sequence of “tetra-loops.” *Proc. Natl. Acad. Sci.* **87**: 8467–8471.
- Zell, R., Sidigi, K., Bucci, E., Stelzner, A., and Gorlach, M. 2002. Determinants of the recognition of enteroviral cloverleaf RNA by coxsackievirus B3 proteinase 3C. *RNA* **8**: 188–201.