Published online April 6, 2005

Nucleic Acids Research, 2005, Vol. 33, No. 6 ***2003–2011***
doi:10.1093/nar/gki501

A novel cGUUAg tetraloop structure with a conserved yYNMGg-type backbone conformation from cloverleaf 1 of bovine enterovirus 1 RNA

Yvonne Ihle, Oliver Ohlenschläger, Sabine Häfner, Elke Duchardt[1], Martin Zacharias[2], Simone Seitz[3], Roland Zell[3], Ramadurai Ramachandran and Matthias Görlach*

Molekulare Biophysik/NMR-Spektroskopie, Institut für Molekulare Biotechnologie e.V., Beutenbergstraße 11, D-07745 Jena, Germany, [1]Institut für Organische Chemie, Johann-Wolfgang-Goethe-Universität, Marie-Curie-Straße 11, D-60439 Frankfurt/M., Germany, [2]International University Bremen, School of Engineering and Science, Campus Ring 1, D-28759 Bremen, Germany and [3]Institut für Virologie und Antivirale Therapie, Friedrich-Schiller-Universität, Winzerlaer Straße 10, D-07745 Jena, Germany

Received December 23, 2004; Revised March 7, 2005; Accepted March 21, 2005

ABSTRACT

The 5′-terminal cloverleaf (CL)-like RNA structures are essential for the initiation of positive- and negative-strand RNA synthesis of entero- and rhinoviruses. SLD is the cognate RNA ligand of the viral proteinase 3C (3C^{pro}), which is an indispensable component of the viral replication initiation complex. The structure of an 18mer RNA representing the apical stem and the cGUUAg D-loop of SLD from the first 5′-CL of BEV1 was determined in solution to a root-mean-square deviation (r.m.s.d.) (all heavy atoms) of 0.59 Å (PDB 1Z30). The first (*anti*G) and last (*syn*A) nucleotide of the D-loop forms a novel 'pseudo base pair' without direct hydrogen bonds. The backbone conformation and the base-stacking pattern of the cGUUAg-loop, however, are highly similar to that of the coxsackieviral uCACGg D-loop (PDB 1RFR) and of the stable cUUCGg tetraloop (PDB 1F7Y) but surprisingly dissimilar to the structure of a cGUAAg stable tetraloop (PDB 1MSY), even though the cGU*U*Ag BEV D-loop and the cGU*A*Ag tetraloop differ by 1 nt only. Together with the presented binding data, these findings provide independent experimental evidence for our model [O. Ohlenschläger, J. Wöhnert, E. Bucci, S. Seitz, S. Häfner, R. Ramachandran, R. Zell and M. Görlach (2004) *Structure*, 12, 237–248] that the proteinase 3C^{pro} recognizes structure rather than sequence.

INTRODUCTION

The enteroviruses consist of at least 8 species comprising more than 90 serotypes and belong to the family of *Picornaviridae* (1). They are the causative agents of a number of diseases including neurological diseases and myocarditis. Enteroviruses are positive-strand RNA viruses. Their genome consists of a 7.5 kb RNA encoding a polyprotein of ~2200 amino acids, which is processed post-translationally by the virus-encoded proteinases 2A^{pro}, 3C^{pro} and 3CD^{pro}. The open reading frame is flanked on both sides by non-translated regions (NTRs). The 5′-NTR contains important signals necessary for the initiation of translation and replication (2,3). Viral and cellular proteins bind to the 5′-cloverleaf (CL), a structure at the 5′ end of the positive-strand RNA, to form a functional ribonucleoprotein complex essential for the initiation of viral RNA replication. This 5′-CL is highly conserved among all enteroviruses. It consists of ~90 nt and contains four joint subdomains: stem A and the stem–loops B, C and D (Figure 1). While stem–loop B is the ligand of the human poly(rC)-binding protein 2 (PCBP2), stem–loop D (SLD) binds the viral 3C^{pro} and 3CD^{pro}, respectively (4–8). A number of *in vivo* and *in vitro* studies, also including chimeric enteroviral constructs, indicated that the SLD:3C^{pro} interaction is not serotype-specific (9–13). Instead, a degenerate consensus sequence residing in the apical tetraloop of SLD, the D-loop, was shown to be important for the RNA:3C^{pro} interaction (7). Structural analysis of the SLD of coxsackievirus B3 (CVB3) indicated that its apical D-loop (uCACGg) forms a tetraloop (8,14) highly similar in structure to stable cUNCGg-type tetraloops, thereby extending the consensus sequence for this tetraloop family (cYNMGg) (15) to yYNMGg. Analysis

*To whom correspondence should be addressed. Tel: +49 3641 656220; Fax: +49 3641 656225; Email: mago@imb-jena.de

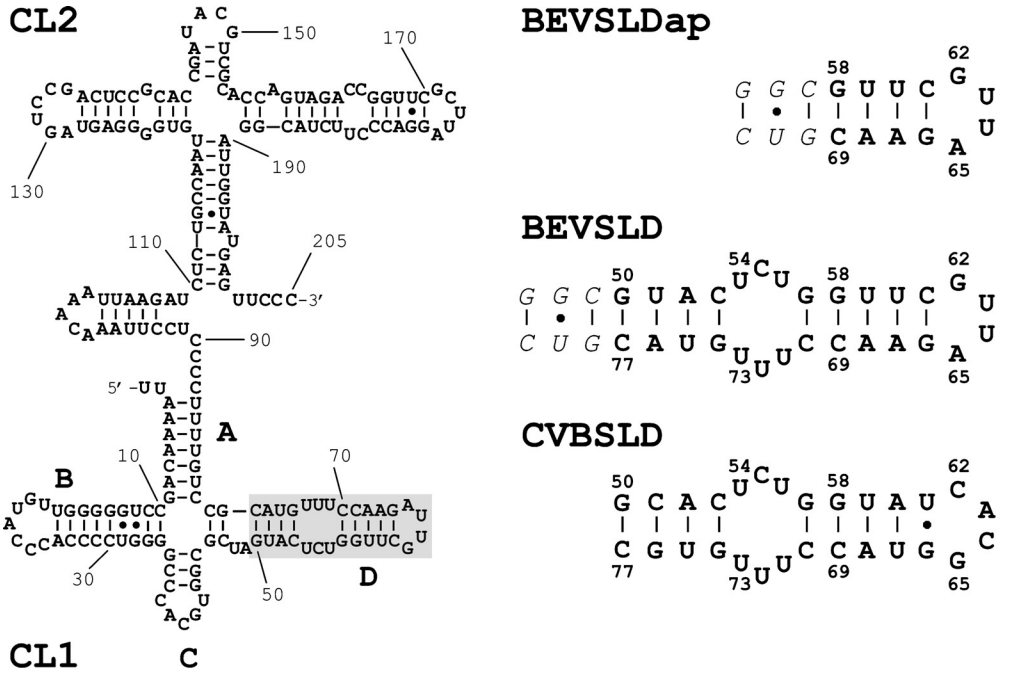

Figure 1. RNA sequences and secondary structure predictions. Left panel: secondary structure of the two cloverleaf motifs (CL1, CL2) in the 5′-NTR of BEV1. The stem (A) and stem–loop (B–D) subdomains of CL1 are indicated and the RNA sequence of SLD is boxed in grey. Right panel: 18mer RNA representing the apical D-loop (BEVSLDap); the 34mer RNA construct representing the SLD of BEV1 CL1 (BEVSLD); and the SLD of coxsackievirus B3 (CVBSLD). The numbering scheme corresponds to the respective native nucleotide sequences of the BEV1 and the CVB3 genomic RNAs. Residues indicated in italics were added for *in vitro* transcription purposes.

in vivo and by NMR spectroscopy revealed that, in addition to 2 nt adjacent to a triple pyrimidine mismatch of the SLD stem, this D-loop structure constitutes an essential element for the specific $3C^{pro}$:SLD interaction (8). The D-loop sequence of CVB3 (uCACGg) is prototypical for the human enteroviruses, and based upon our structural and binding studies, we proposed that the common structural features rather than the sequence of individual D-loops determine the specificity of RNA:protein interaction in this system (7,8).

In contrast to other enteroviruses, a characteristic feature of the bovine enteroviral 5′-NTR is the presence of two 5′-CLs in tandem. The second CL enhances the translational activity of the internal ribosome entry site (13,16), carries a pentanucleotide (cGCUUAg) D-loop and does not bind CVB3 $3C^{pro}$ *in vitro* (7). In contrast, the first CL binds CVB3 $3C^{pro}$ *in vitro* and carries a cGUUAg D-loop (7), which does not fit to the abovementioned yYNMGg consensus for $3C^{pro}$ binding rendering the structural basis for this interaction less clear. However, the viability of CVB3 chimeras, where the cognate CL was replaced by the cloverleaves of BEV1 (13), suggested that this *in vitro* observation carries biological relevance. In addition, one of the three porcine enteroviruses carries a D-loop matching the BEV1 cGUUAg sequence (17). To address the structural basis for the SLD:$3C^{pro}$ interaction in these viruses, the solution structure of the apical part of the SLD from CL1 of BEV1 (Figure 1, right panel, BEVSLDap) was determined. This construct carries the cGUUAg D-loop representing the most significant deviation in sequence as compared with other highly conserved enteroviral SLDs binding to CVB3 $3C^{pro}$ (7,8).

The D-loop of BEV1 presented here constitutes a novel tetraloop structure. However, its backbone adopts a conformation similar to the cUNCGg-like conformations of the yYNMGg sequence family (8,14,15,18). In addition, an analysis of the interaction between an RNA representing the SLD from CL1 of BEV1 (Figure 1, right panel, BEVSLD) with $3C^{pro}$ is reported. Together, this analysis and the structural findings strongly support our model (8) of a structure-specific recognition of enteroviral SLDs by $3C^{pro}$.

MATERIALS AND METHODS

Synthesis, purification and preparation of RNA and $3C^{pro}$

Labelled nucleotides and the apical portion of SLD RNA (BEVSLDap, Figure 1, nucleotides 58–69) from bovine enterovirus 1 cloverleaf 1 (BEV1CL1) were prepared and purified in uniformly ^{15}N-labelled and uniformly ^{15}N,^{13}C-labelled form by *in vitro* transcription with T7 RNA polymerase as described previously (7,19). The NMR samples had final concentrations of 0.9 mM *U*-[^{13}C,^{15}N]-BEVSLDap in H_2O and D_2O, respectively, and 1.0 mM *U*-[^{15}N]-BEVSLDap. The NMR buffer contained 10 mM KH_2PO_4/K_2HPO_4, pH 6.2 with 40 mM KCl, 0.2 mM EDTA and 10% v/v ^2H_2O or 99.99% v/v ^2H_2O, respectively. For binding experiments with $3C^{pro}$, a 34mer (BEVSLD, Figure 1) representing nucleotides 50–77 of the BEV1CL1 was prepared in uniformly ^{15}N-labelled and uniformly ^{15}N,^{13}C-labelled form. The binding experiments were performed at 60 μM *U*-[^{13}C,^{15}N]-BEVSLD and 170 μM $3C^{pro}$ in 20 mM KOAc, pH 5.5, 300 mM KCl. Overexpression and purification of $3C^{pro}$ was carried out as described previously (7). After exchange into NMR buffer (20 mM KOAc, pH 5.5, 300 mM KCl) by

dialysis, the protein was concentrated to 180 μM. The BEVSLD U64A mutant RNA was purchased in unlabelled form from Dharmacon (Lafayette), deprotected as suggested by the manufacturer and purified via high-performance liquid chromatography as detailed for the labelled RNAs above.

NMR spectroscopy

NMR spectra were recorded at temperatures of 283 K (exchangeable protons) and 298 K (exchangeable and non-exchangeable protons) on Varian ^{UNITY}*INOVA* 600 MHz or ^{UNITY}*INOVA* 750 MHz four channel NMR spectrometers equipped with pulse field gradient accessories and triple resonance probes with actively shielded Z gradient coils. The NMR spectra were processed with VNMR (Varian Assoc., Palo Alto) and analysed using the program XEASY (20). Chemical shifts were referenced as described previously (21). Resonance assignment of the exchangeable and non-exchangeable protons of BEVSLDap was performed using (^{1}H,^{15}N)-HSQC (22), (^{1}H,^{15}N)-CPMG-NOESY (23), (^{1}H,^{1}H)-H5(C5C4N)H (C,U) (24), (^{1}H,^{1}H)-HCCNH-TOCSY (G) (25), (^{1}H,^{1}H)-HCCH-TOCSY (A) (26), (^{1}H,^{1}H)-HCCH-COSY (C,U) (27), (^{1}H,^{15}N)-2bond-HSQC (28), (^{1}H,^{13}C)-H(N)CO (29), (^{1}H,^{13}C)-HSQC, 3D (^{1}H,^{1}H,^{13}C)-HCCH-COSY (sugar), 3D (^{1}H,^{1}H,^{13}C)-HCCH-TOCSY (sugar) (27) and (^{1}H,^{15}N)-HCN (30). For sequential assignments 3D (^{1}H,^{1}H,^{13}C)-NOESY-HSQC (31) and (^{1}H,^{31}P)-HCP-CCH-TOCSY (32) spectra were acquired. (^{1}H,^{15}N)-HSQC, (^{1}H,^{1}H)-NOESY, (^{1}H,^{13}C)-HSQC, H5(C5C4N)H, (^{1}H,^{1}H)-HCCNH-TOCSY, (^{1}H,^{1}H)-HCCH-TOCSY and (^{1}H,^{1}H)-HCCH-COSY experiments on a 700 μM *U*-[^{13}C,^{15}N]-BEVSLD and a 320 μM *U*-[^{15}N]-BEVSLD sample in conjunction with the virtually full assignment for BEVSLDap (see Figure 1) and the chemical shift information for the highly homologous CVBSLD were used for assignments of the BEVSLD base moiety resonances.

Structure calculations

NOE constraints for non-exchangeable hydrogens were obtained from 3D (^{1}H,^{1}H,^{13}C)-NOESY-HSQC spectra using the *U*-[^{13}C,^{15}N]-BEVSLDap sample. Separate data sets were acquired for aromatic and aliphatic carbons with mixing times of 120 and 180 ms. Resulting NOEs were classified into four groups representing strong, medium, weak and very weak cross peaks and assigned to upper distances of 2.8, 3.8, 5.4 and 6.4 Å, respectively. Furthermore, a 2D (^{1}H,^{1}H)-NOESY (150 ms mixing time) and a 2D (^{1}H,^{15}N)-CPMG-NOESY (80 ms) experiment (23) were analysed. NOE cross peaks of exchangeable hydrogens were related to upper limit distance constraints of 2.8, 3.8 and 5.4 Å. Intra-residue H5–H6 and intra-base pair imino to amino NOEs were used for the calibration of cross-peak intensities. Hydrogen bond constraints were introduced for all base pairs exhibiting cross peaks in an HNN-COSY spectrum (33) and for the G56U71 wobble pair. Four upper and four lower distance constraints for every AU and GU and six upper and six lower distance constraints for every GC pair, respectively, were used. For all residues except for G66, a ^{31}P resonance could be unambiguously assigned in the (^{1}H,^{31}P)-HCP-CCH-TOCSY spectrum. Because none of them showed any downfield-shift, the backbone torsion angles α and ζ of base-paired residues were restricted to a range of ±60° of A-form helical values and those of the non-base-paired residues to 0° ± 120° according to (34,35). For restraining the sugar pucker $^3J_{H1'H2'}$ coupling constants were extracted from a 2D (^{1}H,^{13}C)-HCCH-E.COSY spectrum (36).

Further torsion angle constraints were derived from local conformational analysis with the FOUND module (37) using 518 NOE derived upper limit distance constraints as well as the angle and hydrogen bond constraints mentioned above. As a result, 184 constraints defining 156 torsion angles could be used for the distance geometry structure calculations employing CYANA (38) in addition to the distance constraints. Energy minimization of the 15 out of 100 conformers with the lowest target function was performed with OPAL (39). The implemented AMBER94 force field (40) was used and a scaling of the electrostatic interaction by 0.5 to avoid excessive charge on the RNA backbone for *in vacuo* calculations was applied. MOLMOL (41) was used for visualization of the resulting structures.

Molecular dynamics

Molecular dynamics (MD) simulations were performed on the tetraloop segment (5′-CGUUCGUUAGAACG) using the Sander module of the Amber8 (Assisted Model Building with Energy Restraints) package (42) starting from the average NMR structure. The simulations were performed in a periodic box including ∼3000 TIP3 water molecules (43), sodium counter ions and using the force field described previously (40). A 9 Å cut-off for the short-range non-bonded interactions was used in combination with the particle mesh Ewald option (44) using a grid spacing of ∼0.9 Å to account for long-range electrostatic interactions. Following initial energy minimization, the system was gradually heated from 50 to 300 K with positional restraints on the RNA atoms over a period of 0.1 ns. During another 0.1 ns simulation time at 300 K, the positional restraining force constant was gradually reduced from 50 kcal mol^{-1} Å^{-2} to zero. The simulations were continued for a total simulation time of 4 ns. Conformations of the final 2 ns were used for comparison with the experimentally determined NMR structures.

Binding assays

RNA:protein interactions *in vitro* were studied by employing electrophoretic mobility shift assays (EMSAs) as detailed previously (7). Interactions *in vivo* were analysed in the yeast three-hybrid system using the RNA–protein Hybrid Hunter^{TM} Kit (Invitrogen) according to the manufacturer's instructions. The gene regions of CVB3 3C^{pro} and 3CD^{pro} were cloned into the pYESTrp3 vector. The catalytic cysteine was substituted by a glycine to inactivate the proteinases. DNA encoding the desired RNAs was cloned into the vector pRH5′ (45). *Saccharomyces cerevisiae* strain L40-*ura3* was used for transformation with the protein prey plasmids, the RNA bait plasmids containing SLD encoding BEV1 sequences and pHybLex/Zeo-MS2. Transformed colonies of *S.cerevisiae* strain L40-*ura3* were grown on selective agar plates, transferred to filter paper, lysed and assayed for β-galactosidase activity 3–6 days after transformation. For documentation, colonies were grown overnight in supplemented minimal salt medium and spotted onto selective x-gal (40 mg/l)

phosphate-buffered (pH 7.0) minimal salt plates. Blue yeast colonies were scored after 5–6 days of incubation at 30°C.

RESULTS AND DISCUSSION

Structure determination

Recently, we demonstrated that the CVB3 proteinase 3C ($3C^{pro}$) interacts *in vitro* and *in vivo* with CLs from a number of entero- and rhinoviruses and that this binding occurs to the SLD subdomain of the CL (7,8). We solved the structure of SLD from CVB3 and mapped the interaction of $3C^{pro}$ to the apical D-loop and to residues adjacent to the internal triple pyrimidine mismatch region of SLD. The apical D-loop of SLD was shown to adopt a structure highly similar to the canonical cUNCGg-type tetraloop in spite of its deviating sequence uCACGg. Together with the binding data, we proposed that the specific SLD:$3C^{pro}$ interaction critically depends upon the cUNCGg-like backbone conformation common to D-loops with the rather degenerate consensus yYNYGg (7,8). The observed interaction of the CL1 of the BEV1 5′-NTR (Figure 1) with $3C^{pro}$ (7) and the viability of CVB3 chimeras, where the cognate CL was replaced by the CLs of BEV1 (13), posed the intriguing question whether the SLD of the BEV1 CL1 (BEVSLD, Figure 1) is similar in structure to CVBSLD (Figure 1) as the D-loop sequence (cGUUAg) of BEVSLD much more closely resembles that of the GNRA-tetraloop family (46–48) than that of the UNCG-like family consensus (yYNMGg) (14,15,18). Although the remainder of the two SLDs is highly homologous, the most significant difference between BEVSLD and the CVBSLD lies in the apical D-loop (Figure 1). Hence, we determined the structure of an RNA (BEVSLDap, Figure 1) representing the apical stem and the D-loop of BEVSLD in order to study the structural basis of the RNA:$3C^{pro}$ interaction further.

For structure determination, BEVSLDap was produced in ^{15}N- and $^{15}N/^{13}C$-labelled form by *in vitro* transcription. Virtually complete resonance assignments were achieved by standard triple resonance experiments. An extremely broadened resonance typical for an amino group was observed at 6.81 p.p.m./80.8 p.p.m. which could not be assigned. This amino group made an NOE contact observed in the CPMG-NOESY at 10°C to the imino group of G66 and, by exclusion, it belongs either to the exocyclic amino group of G62 or of A65, respectively. The imino resonance of G62 at 10.44 p.p.m./147.8 p.p.m. was broadened even at 10°C. No imino proton resonances were observed for U63 and U64, indicating a rapid exchange with the solvent. For the observable imino groups, hydrogen bonding in SLD was analysed by the HNN-COSY experiment. This analysis revealed six correlations (data not shown) between U and G imino groups and the N1 of A and N3 of C, respectively, consistent with 4 GC and 2 AU pairs in BEVSLDap. For the loop residues (G62, U63, U64 and A65), no NH...N hydrogen bonds were detected in the HNN-COSY experiment. The chemical shifts of the C2 and C4 carbonyl nuclei of U59 and U60, as analysed in a H(N)CO experiment, were found to be in the range typical for canonical AU base pairs (8,49). The U residue of the penultimate GU wobble pair, belonging to the extension at the basis of the BEVSLDap stem (Figure 1, residues in italics), exhibited a ^{13}C resonance of its carbonyl C2 shifted downfield

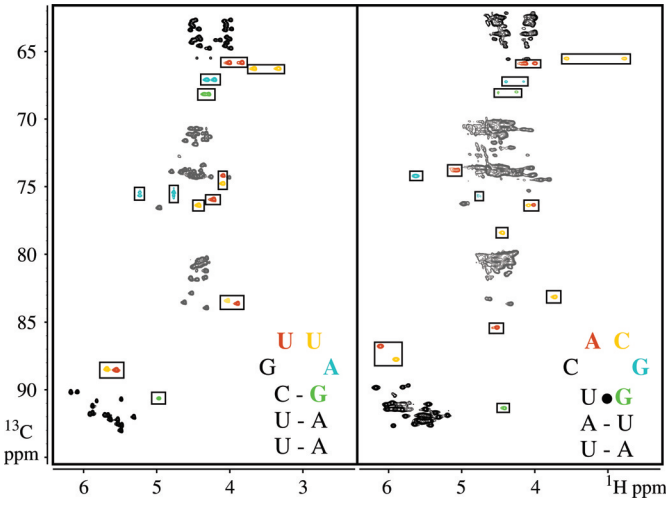

Figure 2. ($^1H,^{13}C$)-HSQC spectra of the BEV (left) and the CVB3 stem–loop D (right) RNAs. Characteristic cross peaks originating from the tetraloops and their closing base pairs are coloured (loop nucleotide 2, red; nucleotide 3, yellow; nucleotide 4, blue; and G of closing base pair, green).

with respect to the C2 resonances of the U residues involved in canonical AU base pairs. Together with the observed strong cross peaks in a ($^1H,^1H$)-NOESY spectrum between the imino protons of both residues this indicated a standard GU wobble pair. During the process of assignment, we noticed characteristic chemical shifts for ribose carbon and hydrogen nuclei of the BEV D-loop, e.g. H5′/H5″/C5′ of U63, U64 and A65, H4′/C4′ of U64, H3′/C3′ of A65, which were reminiscent of the corresponding shifts found for the CVB3 D-loop (8), thus exhibiting a similar 'fingerprint' for the respective nuclei in the ($^1H,^{13}C$)-HSQC spectra (Figure 2). The structure of BEVSLDap was calculated from 518 NOE constraints derived from ($^1H,^1H$)-NOESY, ($^1H,^{15}N$)-CPMG-NOESY and ^{13}C-edited 3D-NOESY-HSQC spectra, 36 hydrogen bond constraints deduced from the HNN-COSY and the H(N)CO experiments, 17 torsion angle constraints for angle α and ζ as derived from ^{31}P chemical shift analysis and 17 $^3J_{H1'H2'}$ coupling constants extracted from a HCCH-E.COSY defining the torsion angle ν_1.

BEVSLDap is capped by a novel tetraloop structure

The structure of BEVSLDap represented by the ensemble of 15 energy minimized structures with the lowest target function is shown in Figure 3A. The structural statistics is given in Table 1. The structure consists of an A-helix capped by a tetraloop. The helical portion of the molecule involves residues 58–61 and 66–69 of the BEVSLD (see also Figure 1) and includes the residues at the base of the stem added for *in vitro* transcription purposes (Figure 1, residues in italics). The entire stem of BEVSLDap adopts a standard A-form conformation and will not be discussed any further.

The apical cGUUAg D-loop represents a well-defined tetraloop structure (Figure 3C) resting upon the canonical C61:G66 slightly buckled closing base pair. Three of its bases are engaged in stacking interactions (Figure 3B). This tetraloop exhibits several novel properties (Figure 3B and C). G62 and A65 are found in an unusual staggered juxtaposition, which apparently does not allow for the formation of direct

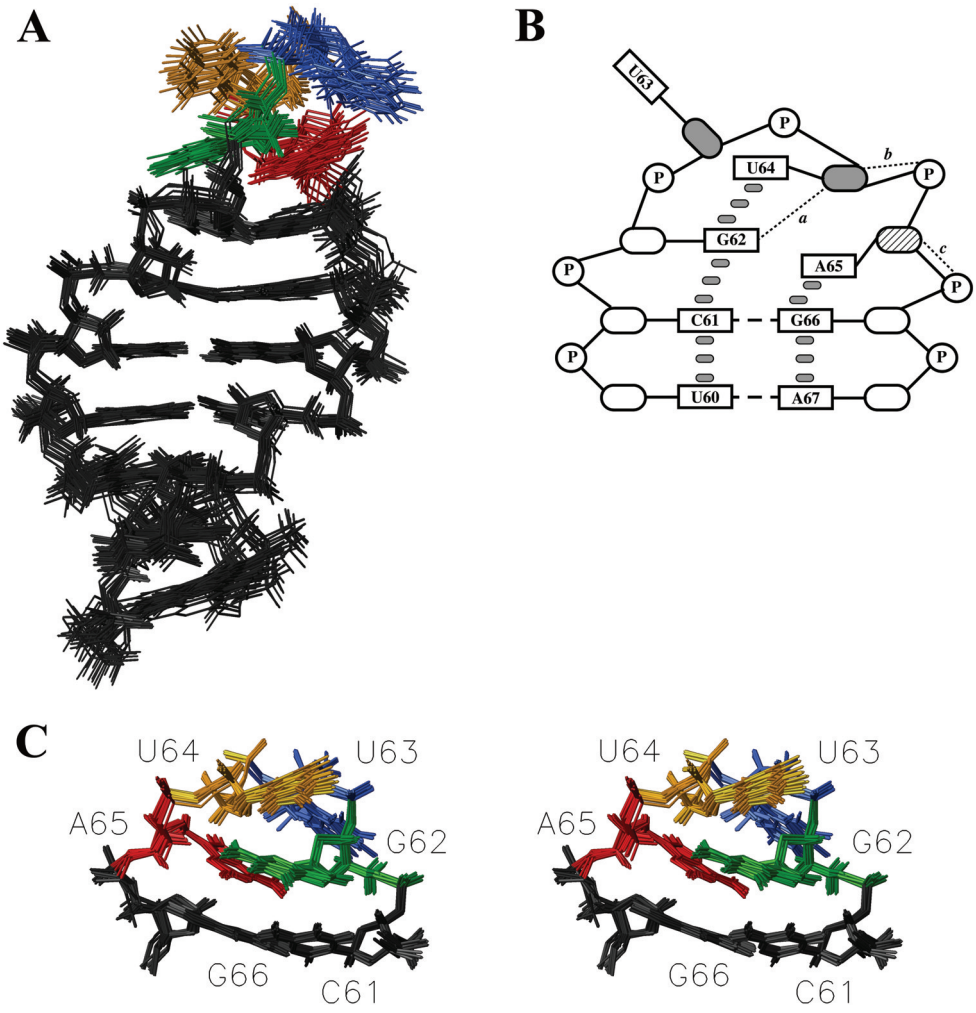

Figure 3. The NMR structure of BEVSLDap. (**A**) Superimposition (all heavy atoms) of 15 energy minimized conformers of 100 calculated structures. Residues of the A-helical stem are shown in black. The tetraloop residues are coloured (G62, green; U63, blue; U64, yellow; A65, red). (**B**) Schematic representation of stacking interactions and possible hydrogen bonding consistent with the calculated BEVSLDap structure: (*a*) G62H22–U64O4′, (*b*) U64HO2′–A65OP2 and (*c*) A65HO2′–G12OP1. Filled ellipses represent residues with a C2′-*endo*, empty ellipses a C3′-*endo* sugar pucker, respectively. The ribose moiety of A65 is shown as hatched ellipsis to indicate a possible C2′-*endo*, C3′-*endo* equilibrium (see text). (**C**) Stereoview (side-by-side) of the local superimposition (heavy atoms) of the apical tetraloop (G62–A65) and its closing base pair (C61–G66) of BEVSLDap. View into the major groove.

Table 1. Structural statistics for the BEVSLDap RNA

Number of experimental constraints	
Distance constraints	518
Torsion angle constraints	156
Number of hydrogen bond constraints	36
CYANA target function before energy minimization (Å^2)	0.0463 ± 0.0022
AMBER energies (kcal/mol) after energy minimization:	
Physical energy	−159.28 ± 8.4
Van der Waals energy	−206.19 ± 3.2
Electrostatic energy	−370.24 ± 6.2
Number of violations	
NOE violation >0.2 Å	0
Torsion angle restraints violation >2.5°	0
Mean deviation from idealized covalent geometry	
Bond length (Å)	0.00452 ± 0.000086
Bond angles (°)	1.34654 ± 0.001190
Heavy atom r.m.s.d. from mean structure (Å)	
All	0.509
Best to mean	0.419
Tetraloop (C61–G66)	0.295

hydrogen bonds between their base moieties. This conformation contrasts the situation in UNCG-type (50,51) and GNRA-type (46–48,52) tetraloops where the first and the fourth loop residue are paired via hydrogen bonds. However, the calculated structure could accommodate a hydrogen bond of the amino group of G62 and the O4′ ribose oxygen of U64 (Figure 3B). Likewise, hydrogen bonds between 2′OH of U64 and a non-bridging oxygen of A65 as well as the 2′OH of A65 and the non-bridging oxygen of G66 (Figure 3B) would be consistent with the calculated structure. G62 adopts A-form conformation, including a glycosidic angle in *anti* conformation and a canonical C3′-*endo* sugar pucker, similar to the first residue in the two tetraloop families mentioned above. The calculated structure, however, shows an unusual χ angle in the *syn* range for the fourth loop residue, A65. A *syn*G in this position is typical for the UNCG-type tetraloop, whereas the fourth loop nucleotide of GNRA tetraloops adopts *anti* conformation. The *syn* torsion angle of A65 is reflected by a strong intra-residue H8–H1′ NOE cross peak and $^3J_{C8/C4–H1'}$ coupling

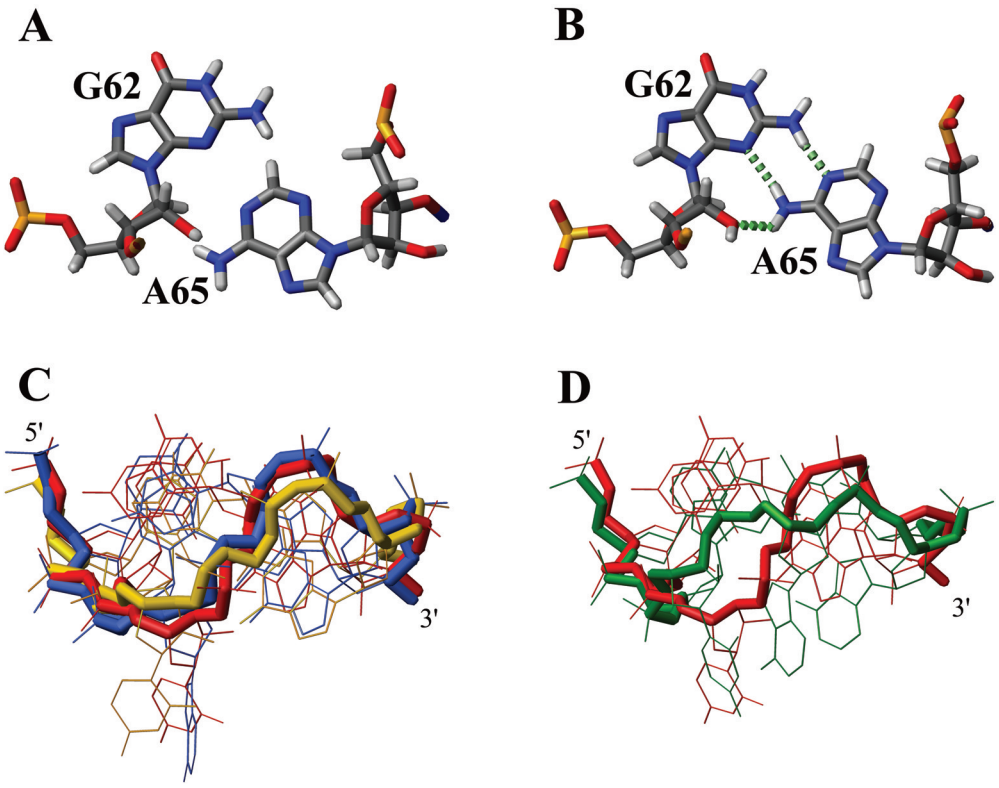

Figure 4. Structural features of the D-loop of BEVSLD. (**A**) The G62:*syn*A65 'pseudo base pair' as found in the experimentally determined mean structure of BEVSLDap. (**B**) The G62:*syn*A65 base pair forming three hydrogen bonds (indicated as broken green lines) as derived from a 4 ns MD simulation using the calculated mean structure of BEVSLDap as starting structure. (**C**) Superimposition (backbone atoms of loop and closing base pair, top view) of the apical loop regions of the BEVSLDap structure (nt 61–66, red) with SLD of coxsackievirus B3 (uCACGg, blue, PDB 1RFR) and the TL1 tetraloop from 16S rRNA (cUUCGg, yellow, PDB 1F7Y). (**D**) Superimposition (backbone atoms of loop and closing base pair, top view) of the apical loop region of the BEVSLDap structure (nt 61–66, red) with the crystal structure of the cGUAAg tetraloop (blue, PDB 1MSY). The 5′ and 3′ ends of the tetraloop backbones are indicated.

constants of 3.40 and 3.49 Hz, respectively (53). The magnitude of the $^{3}J_{H1'H2'}$ coupling constant (4 Hz) for the ribose of A65 is different from the respective values observed in BEVSLDap for residues in C3′- (0.0–2.5 Hz) or C2′-*endo* pucker (7 Hz), suggesting an equilibrium between C3′-*endo* and C2′-*endo* conformation (54). The base of A65 stacks onto G66 of the closing base pair at the 3′ side of the loop (Figure 3B and C). Prompted by the unusual conformation of the experimentally determined G:*syn*A 'pseudo base pair' (Figure 4A), an MD simulation for a total of 4 ns and in the presence of explicit solvent and counter ions was performed. The MD trajectory shows a stable hydrogen bonding between the amino group of G62 and the N1 of A65, between the amino group of A65 and the N3 of G62 as well as the 2′OH of G62. However, the base–base alignment needed for this G:*syn*A is inconsistent with a number of NMR observations. This includes the lack of defined amino resonances for G62 and A65, missing NOEs between the amino group of G62 and the aromatic H2 as well as the amino group of A65. Moreover, the observation of several NOEs, i.e. H1′/H2′ of G62 to H2 of A65, H3′ of U64 to H8 of A65, H1′ of U63 to H5′ of U64 and H2′ of G62 to H2 of A65, would not be expected for the MD-derived loop structure (Figure 4B).

The second and third nucleotide of the loop, U63 and U64, projects their bases into the minor or the major groove, respectively. Both sugar moieties adopt a C2′-*endo* conformation. The base of U63 is not involved in interactions with any part of the tetraloop. The base of U64 stacks onto G62, which in turn is stacking on C61 of the closing base pair (Figure 3B and C). The calculated backbone conformation indicates a reversal of direction between the second and third loop nucleotide, i.e. U63 and U64, resulting in an 'S'-shaped overall appearance (Figure 4C). The observed chemical shift signature (see above), the stacking pattern, the projection of the second and third loop base into opposite helical grooves and the overall backbone conformation are very similar (Figure 4C) to the D-loop of CVBSLD (r.m.s.d. 0.87 Å) (8) and a cUUCGg stable tetraloop (r.m.s.d. 0.66 Å) (51). In contrast, in the cGUAAg tetraloop (46), which deviates from BEV D-loop by one residue only and superimposes with an r.m.s.d. of 1.66 Å (Figure 4D), the reversal of backbone direction occurs between the first and second loop nucleotide which in turn results in a continuous stacking of the bases of loop nucleotides 2–4 on the major groove side of the loop. Thus, a difference in 1 nt only causes a transformation between an UNCG-type and a GNRA-type tetraloop structure.

BEVSLD:3C^{pro} interaction

To address the functional significance of our structural observations, the binding of BEVSLD to CVB3 3C^{pro} was investigated. In addition to binding to CL1 of BEV (7), 3C^{pro} binds to BEVSLD *in vivo* in the yeast three-hybrid system shown in Figure 5A and *in vitro* in an EMSA (Figure 5D and data not

shown). Deleting 2 nt from the D-loop of BEVSLD renders the RNA non-binding (Figure 5A). Also, the SLD derived from the second CL of BEV1 did not bind in yeast to CVB3 $3C^{pro}$ (Figure 5A), matching the *in vitro* results obtained with the entire CL2 earlier (7). To analyse the interaction of BEVSLD and $3C^{pro}$ further (^1H,^{13}C)-HSQC spectra for aromatic resonances were recorded for the free BEVSLD and for the BEVSLD:$3C^{pro}$ complex under identical conditions. Owing to the limited solubility of $3C^{pro}$, those spectra had to be taken at elevated ionic strength and on dilute samples of 60 μM RNA and 170 μM $3C^{pro}$ (RNA:protein molar ratio 1:3). In the (^1H,^{13}C)-HSQC spectra of the complex (Figure 5B) resonances of the aromatic groups of C53, G62, U63, U64 and G66 are missing from their original positions. However, for the newly appearing resonances, the solubility-limited concentration of the complex precluded assignments. In addition, the resonances of C61, A65 and C69 are significantly broadened. This observation is consistent with binding at the intermediate to slow NMR timescale and with a K_D of 3.7 μM for the interaction of $3C^{pro}$ with the first 5′-CL of BEV1 (7). Assuming a bimolecular binding reaction, this K_D corresponds to a 97% saturation of the RNA with $3C^{pro}$ in the NMR experiment described above. Importantly, the affected residues are located in the D-loop of BEVSLD and adjacent to the internal triple pyrimidine mismatch and are equivalent to the residues of the CVBSLD showing a similar line broadening upon complex formation with $3C^{pro}$ (8). For BEVSLD, we also observed the expected line broadening for the imino resonances of G62 and G66 in (^1H,^{15}N)-HSQC spectra (data not shown). Additional residues (U59 and U60) showed such an imino line broadening, suggesting that they also might contribute to $3C^{pro}$ binding. This effect, however, was not reflected in the signals of their H6 base protons in the corresponding (^1H,^{13}C)-HSQC spectrum (Figure 5B).

The residues of BEVSLD implied in $3C^{pro}$ binding here are essentially the same as in CVBSLD (8) and indicated in Figure 5C. The most significant deviation in sequence between BEVSLD (cGUUAg) and CVBSLD (uCACGg), however, resides in their D-loops which, nevertheless, share the UNCG-type backbone conformation of the yYNMGg sequence family. Hence, the novel tetraloop structure presented here sheds new light on the sequence plasticity for such tetraloops. The conserved structural motif rather than a limited sequence family explains their specific interaction with $3C^{pro}$, which is essential for the initiation of enteroviral replication. On the other hand, a mutation of the D-loop sequence (cGUUAg) into cGUAAg would cause a transformation from an UNCG-type to a GNRA-type tetraloop structure. This transformation, however, would result in significantly

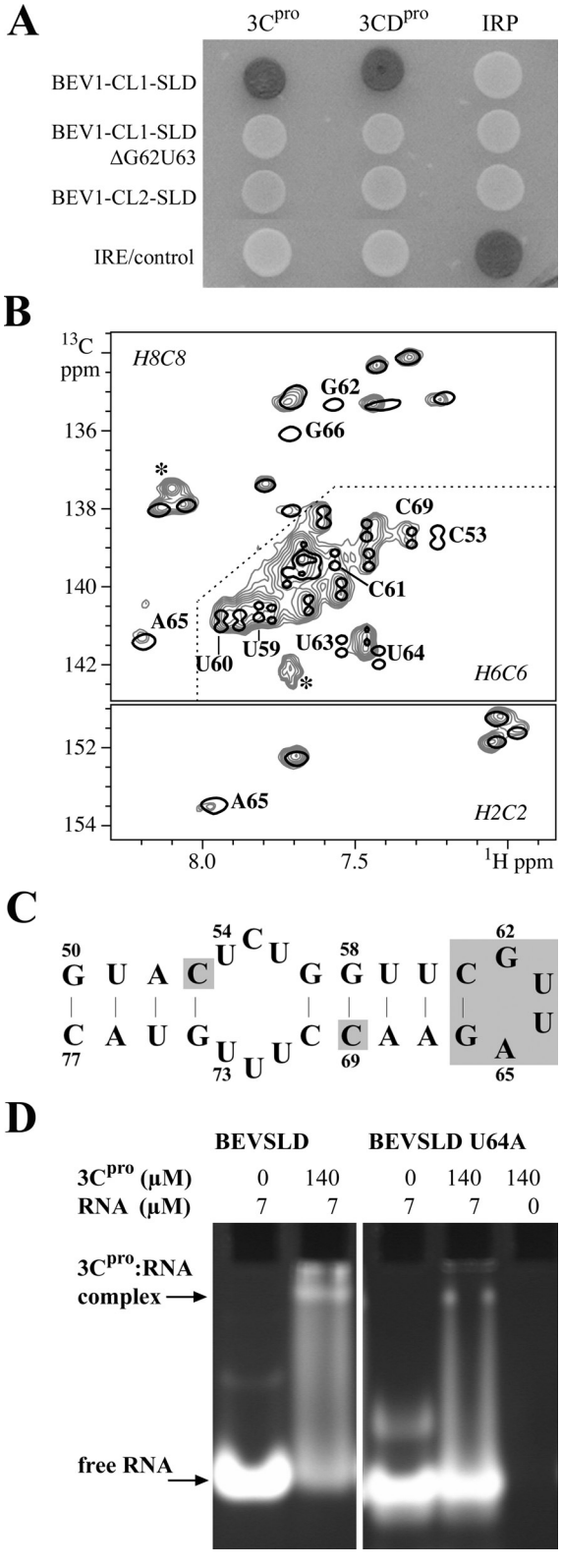

Figure 5. CVB3 $3C^{pro}$ binding studies. (**A**) Binding of CVB3 $3C^{pro}$ and $3CD^{pro}$ to the SLD of cloverleaf 1 and cloverleaf 2, respectively, of BEV1 as assayed in yeast three-hybrid experiments. Wild-type SLD of cloverleaf 1 (BEV1-CL1-SLD), mutant of BEVSLD, from which G62 and U63 were deleted (BEV1-CL1-SLD-ΔG62U63) and wild-type SLD of cloverleaf 2 (BEV1-CL2-SLD). The rabbit iron-regulatory protein 1 (IRP) and the rat ferritin light chain iron-responsive element (IRE) were used as control. (**B**) Superimposition of (^1H,^{13}C)-HSQC spectra for the aromatic regions (H2C2, H6C6 and H8C8) of BEVSLD acquired in the presence (grey contours) and absence of $3C^{pro}$ (thick black contours). Residues indicated by nucleotide type and residue number are discussed in the text. Asterisks denote new resonances appearing in the complex. (**C**) The residues most affected in the complex are indicated on the secondary structure of BEVSLD using grey boxes (see text for discussion). (**D**) Binding of CVB3 $3C^{pro}$ to wild-type and mutant BEVSLD RNAs as analysed by using EMSA. RNA was detected by ethidium bromide staining. The concentration of $3C^{pro}$ and of the respective RNA present in the binding reaction before gel electrophoresis is indicated. Wild-type SLD of cloverleaf 1 (BEVSLD), mutant of BEVSLD (BEVSLD U64A), where U64 was replaced by an A.

different D-loop surface properties thereby interfering with the $3C^{pro}$:RNA interaction. Consistent with this, a dramatically reduced binding of $3C^{pro}$ to the cGUAAg-mutant of BEVSLD under NMR buffer conditions is observed in an EMSA (Figure 5D). No binding was detected for the unrelated stem–loop B from CVB3 (data not shown). Based on our findings, one could speculate that such a mutation *in vivo* would lead to non-viable virus mutants. This notion is supported by the fact that D-loop sequences accommodating a GNRA-type structure have not been reported in the NCBI GenBank (www.ncbi.nih.gov/Genbank) or the Picorna Virus Sequence Database (www.iah.bbsrc.ac.uk/virus/Picornaviridae/SequenceDatabase/Index.html) to date.

The coordinates, distance and angle constraints as well as the chemical shifts of the BEVSLDap RNA have been deposited in the PDB (accession code 1Z30) and the BioMagResBank (accession code 6562), respectively.

ACKNOWLEDGEMENTS

The authors thank A. Heller and T. Kubitzki for technical assistance and J. Wöhnert for critically reading the manuscript. This work was supported in part by the Deutsche Forschungsgemeinschaft (DFG Ze 446/3). The IMB is a member of the Science Association 'Gottfried Wilhelm Leibniz' (WGL) and is financially supported by the Federal Government of Germany and the State of Thuringia. Funding to pay the Open Access publication charges for this article was provided by IMB.

Conflict of interest statement. None declared.

REFERENCES

1. King,A.M.Q., Brown,F., Christian,P., Hovi,T., Hyypiä,T., Knowles,N.J., Lemon,S.M., Minor,P.D., Palmenberg,A.C., Skern,T. *et al.* (2000) In Van Regenmortel,M.H.V., Fauquet,C.M., Bishop,D.H.L., Calisher,C.H., Carsten,E.B., Estes,M.K., Lemon,S.M., Maniloff,J., Mayo,M.A., McGeoch,D.J., Pringle,C.R. and Wickner,R.B. (eds), *Virus Taxonomy. Seventh Report of the International Committee for the Taxonomy of Viruses*. Academic Press, New York, San Diego, pp. 657–673.
2. Ehrenfeld,E. and Teterina,N.L. (2002) In Semler,B.L. and Wimmer,E. (eds), *Molecular Biology of Picornaviruses* ASM Press, Washington DC, pp. 159–170.
3. Paul,A.V. (2002) In Semler,B.L. and Wimmer,E. (eds), *Molecular Biology of Picornaviruses* ASM Press, Washington DC, pp. 227–246.
4. Andino,R., Rieckhof,G.E. and Baltimore,D. (1990) A functional ribonucleoprotein complex forms around the 5′ end of poliovirus RNA. *Cell*, **63**, 369–380.
5. 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.
6. 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.
7. Zell,R., Sidigi,K., Bucci,E., Stelzner,A. and Görlach,M. (2002) Determinants of the recognition of enteroviral cloverleaf RNA by coxsackievirus B3 proteinase 3C. *RNA*, **8**, 188–201.
8. Ohlenschläger,O., Wöhnert,J., Bucci,E., Seitz,S., Häfner,S., Ramachandran,R., Zell,R. and Görlach,M. (2004) The structure of the stem–loop D subdomain of coxsackievirus B3 cloverleaf RNA and its interaction with the proteinase 3C. *Structure (Camb.)*, **12**, 237–248.
9. Johnson,V.H. and Semler,B.L. (1988) Defined recombinants of poliovirus and coxsackievirus: sequence-specific deletions and functional substitutions in the 5′-noncoding regions of viral RNAs. *Virology*, **162**, 47–57.
10. Todd,S., Towner,J.S., Brown,D.M. and Semler,B.L. (1997) Replication-competent picornaviruses with complete genomic RNA 3′ noncoding region deletions. *J. Virol.*, **71**, 8868–8874.
11. Xiang,W., Harris,K.S., Alexander,L. and Wimmer,E. (1995) Interaction between the 5′-terminal cloverleaf and 3AB/3CDpro of poliovirus is essential for RNA replication. *J. Virol.*, **69**, 3658–3667.
12. Zell,R., Klingel,K., Sauter,M., Fortmüller,U. and Kandolf,R. (1995) Coxsackieviral proteins functionally recognize the polioviral cloverleaf structure of the 5′-NTR of a chimeric enterovirus RNA: influence of species-specific host cell factors on virus growth. *Virus Res.*, **39**, 87–103.
13. Zell,R., Sidigi,K., Henke,A., Schmidt-Brauns,J., Hoey,E., Martin,S. and Stelzner,A. (1999) Functional features of the bovine enterovirus 5′-non-translated region. *J. Gen. Virol.*, **80**, 2299–2309.
14. Du,Z., Yu,J., Andino,R. and James,T.L. (2003) Extending the family of UNCG-like tetraloop motifs: NMR structure of a CACG tetraloop from coxsackievirus B3. *Biochemistry*, **42**, 4373–4383.
15. 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.
16. Zell,R. and Stelzner,A. (1997) Application of genome sequence information to the classification of bovine enteroviruses: the importance of 5′- and 3′-nontranslated regions. *Virus Res.*, **51**, 213–229.
17. Krumbholz,A., Dauber,M., Henke,A., Birch-Hirschfeld,E., Knowles,N.J., Stelzner,A. and Zell,R. (2002) Sequencing of porcine enterovirus groups II and III reveals unique features of both virus groups. *J. Virol.*, **76**, 5813–5821.
18. 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.
19. Stoldt,M., Wöhnert,J., Ohlenschläger,O., Görlach,M. and Brown,L.R. (1999) The NMR structure of the 5S rRNA E-domain-protein L25 complex shows preformed and induced recognition. *EMBO J.*, **18**, 6508–6521.
20. Bartels,C., Xia,T., Billeter,M., Güntert,P. and Wüthrich,K. (1995) The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR*, **6**, 1–10.
21. Sich,C., Ohlenschläger,O., Ramachandran,R., Görlach,M. and Brown,L.R. (1997) Structure of an RNA hairpin loop with a 5′-CGUUUCG-3′ loop motif by heteronuclear NMR spectroscopy and distance geometry. *Biochemistry*, **36**, 13989–14002.
22. Bodenhausen,G. and Ruben,D.J. (1980) Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.*, **69**, 185.
23. Mueller,L., Legault,P. and Pardi,A. (1995) Improved RNA structure determination by detection of NOE contacts to exchange-broadened amino-protons. *J. Am. Chem. Soc.*, **117**, 11043–11048.
24. Wöhnert,J., Ramachandran,R., Görlach,M. and Brown,L.R. (1999) Triple-resonance experiments for correlation of H5 and exchangeable pyrimidine base hydrogens in 13C,15N-labeled RNA. *J. Magn. Reson.*, **139**, 430–433.
25. Sklenar,V., Dieckmann,T., Butcher,S.E. and Feigon,J. (1996) Through-bond correlation of imino and aromatic resonances in 13C-, 15N-labeled RNA via heteronuclear TOCSY. *J. Biomol. NMR*, **7**, 83–87.
26. Marino,J.P., Prestegard,J.H. and Crothers,D.M. (1994) Correlation of adenine H2/H8 resonances in uniformly C-13 labeled RNAs by 2D HCCH-TOCSY: a new tool for H-1 assignment. *J. Am. Chem. Soc.*, **116**, 2205–2206.
27. Nikonowicz,E.P. and Pardi,A. (1993) An efficient procedure for assignment of the proton, carbon and nitrogen resonances in 13C/15N labeled nucleic acids. *J. Mol. Biol.*, **232**, 1141–1156.
28. Sklenar,V., Peterson,R.D., Rejante,M.R. and Feigon,J. (1994) Correlation of nucleotide base and sugar protons in a 15N-labeled HIV-1 RNA oligonucleotide by 1H-15N HSQC experiments. *J. Biomol. NMR*, **4**, 117–122.
29. Ikura,M., Kay,L.E. and Bax,A. (1990) A novel approach for sequential assignment of 1H, 13C, and 15N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. *Biochemistry*, **29**, 4659–4667.

30. Sklenar,V., Peterson,R.D., Rejante,M.R. and Feigon,J. (1993) Two- and three-dimensional HCN experiments for correlating base and sugar resonances in 15N, 13C-labeled RNA oligonucleotides. *J. Biomol. NMR*, **3**, 721–727.
31. Zwahlen,C., Legault,P., Vincent,S.J.F., Greenblatt,J., Konrat,R. and Kay,L.E. (1997) Methods for measurement of intermolecular NOEs by multinuclear NMR spectroscopy: application to a bacteriophage 1 N-peptide/boxB RNA complex. *J. Am. Chem. Soc.*, **119**, 6711–6721.
32. Marino,J.P., Schwalbe,H., Anklin,C., Bermel,W. and Crothers,D.M. (1995) Sequential correlation of anomeric ribose protons and intervening phosphorus in RNA oligonucleotides by a H-1,C-13,P-31 triple resonance experiment: HCP-CCH-TOCSY. *J. Biomol. NMR*, **5**, 87–92.
33. Dingley,A.J. and Grzesiek,S. (1998) Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide 2JNN couplings. *J. Am. Chem. Soc.*, **120**, 8293–8297.
34. Gorenstein,D.G. (1984) *Phosphorus-31 NMR: Principles and Applications*. Academic Press, New York, NY.
35. Varani,G., Cheong,C. and Tinoco,I.,Jr (1991) Structure of an unusually stable RNA hairpin. *Biochemistry*, **30**, 3280–3289.
36. Griesinger,C. and Eggenberger,U. (1992) Determination of proton-proton coupling constants in 13C-labeled molecules. *J. Magn. Reson.*, **97**, 426–434.
37. Güntert,P., Billeter,M., Ohlenschläger,O., Brown,L.R. and Wüthrich,K. (1998) Conformational analysis of protein and nucleic acid fragments with the new grid search algorithm FOUND. *J. Biomol. NMR*, **12**, 543–548.
38. Güntert,P., Mumenthaler,C. and Wüthrich,K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.*, **273**, 283–298.
39. Luginbühl,P., Güntert,P., Billeter,M. and Wüthrich,K. (1996) The new program OPAL for molecular dynamics simulations and energy refinements of biological macromolecules. *J. Biomol. NMR*, **8**, 136–146.
40. Cornell,W.D., Cieplak,P., Bayly,C.I., Gould,I.R., Merz,K.M., Ferguson,D.M., Spellmeyer,D.C., Fox,T., Caldwell,J.W. and Kollman,P.A. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, **117**, 5179–5197.
41. Koradi,R., Billeter,M. and Wüthrich,K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.*, **14**, 51–55, 29–32.
42. Pearlman,D.A., Case,D.A., Caldwell,J.W., Ross,W.S., Cheatham,T.E., Debolt,S., Ferguson,D., Seibel,G. and Kollman,P.A. (1995) Amber, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput. Phys. Commun.*, **91**, 1–41.
43. Jorgensen,W.L., Chandrasekhar,J., Madura,J., Impey,R.W. and Klein,M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys*, **79**, 926–935.
44. Darden,T., York,D. and Pedersen,L. (1993) Particle mesh Ewald: an *N*log*N* method for Ewald sums in large systems. *J. Chem. Phys.*, **98**, 11089–10092.
45. SenGupta,D.J., Zhang,B., Kraemer,B., Pochart,P., Fields,S. and Wickens,M. (1996) A three-hybrid system to detect RNA–protein interactions *in vivo*. *Proc. Natl Acad. Sci. USA*, **93**, 8496–8501.
46. Correll,C.C. and Swinger,K. (2003) Common and distinctive features of GNRA tetraloops based on a GUAA tetraloop structure at 1.4 A resolution. *RNA*, **9**, 355–363.
47. 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.
48. 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.
49. Fürtig,B., Richter,C., Wöhnert,J. and Schwalbe,H. (2003) NMR spectroscopy of RNA. *ChemBioChem.*, **4**, 936–962.
50. Allain,F.H. and Varani,G. (1995) Structure of the P1 helix from group I self-splicing introns. *J. Mol. Biol.*, **250**, 333–353.
51. 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.
52. Juneau,K., Podell,E., Harrington,D.J. and Cech,T.R. (2001) Structural basis of the enhanced stability of a mutant ribozyme domain and a detailed view of RNA–solvent interactions. *Structure (Camb.)*, **9**, 221–231.
53. Trantirek,L., Stefl,R., Masse,J.E., Feigon,J. and Sklenar,V. (2002) Determination of the glycosidic torsion angles in uniformly 13C-labeled nucleic acids from vicinal coupling constants 3JC2/4-H1′ and 3JC6/8-H1′. *J. Biomol. NMR*, **23**, 1–12.
54. de Leeuw,F.A.A.M. and Altona,C. (1982) Conformational analysis of beta-D-ribo-, beta-D-deoxyribo-, beta-D-arabino-, beta-D-xylo-, and beta-D-lyxo-nucleosides from proton–proton coupling constants. *J. Chem. Soc. Perkin Trans. II*, 375–384.