Published online 7 February 2008 *Nucleic Acids Research, 2008, Vol. 36, No. 5* ***1681–1689***
doi:10.1093/nar/gkn006

Solution structure of stem-loop α of the hepatitis B virus post-transcriptional regulatory element

Martin Schwalbe[1], Oliver Ohlenschläger[1], Aliaksandr Marchanka[1], Ramadurai Ramachandran[1], Sabine Häfner[1], Tilman Heise[2] and Matthias Görlach[1,*]

[1]Leibniz-Institut für Altersforschung/Fritz-Lipmann-Institut, Beutenbergstr. 11, D-07745 Jena and
[2]Heinrich-Pette-Institut für Experimentelle Virologie und Immunologie, Martinistr. 52, D-20251 Hamburg, Germany

Received October 12, 2007; Revised and Accepted January 7, 2008

ABSTRACT

Chronic hepatitis B virus (HBV) infections may lead to severe diseases like liver cirrhosis or hepatocellular carcinoma (HCC). The HBV post-transcriptional regulatory element (HPRE) facilitates the nuclear export of unspliced viral mRNAs, contains a splicing regulatory element and resides in the 3′-region of all viral transcripts. The HPRE consists of three sub-elements α (nucleotides 1151–1346), β1 (nucleotides 1347–1457) and β2 (nucleotides 1458–1582), which confer together full export competence. Here, we present the NMR solution structure (pdb 2JYM) of the stem-loop α (SLα, nucleotides 1292–1321) located in the sub-element α. The SLα contains a CAGGC pentaloop highly conserved in hepatoviruses, which essentially adopts a CUNG-like tetraloop conformation. Furthermore, the SLα harbours a single bulged G residue flanked by A-helical regions. The structure is highly suggestive of serving two functions in the context of export of unspliced viral RNA: binding sterile alpha motif (SAM-) domain containing proteins and/or preventing the utilization of a 3′-splice site contained within SLα.

INTRODUCTION

The Hepatitis B virus (HBV) belongs to the hepatotropic DNA virus family, the *Hepadnaviridae*. It is a small, enveloped virus with a circular, partially double-stranded DNA genome of 3.2 kb, which replicates exclusively in the liver and causes hepatitis B (1). In spite of available vaccines, worldwide 350 million people are chronically infected with HBV, of which 1 million die per year from liver cirrhosis and HCC (2) underscoring the need for novel therapeutic approaches.

After infection of the cell, the viral genome is delivered into the nucleus. The partially double-stranded viral genome is converted by host enzymes to a covalently closed, circular, double-stranded DNA molecule (cccDNA). The minus strand of the cccDNA represents the template for all viral transcripts. Starting at four different promoters (Figure 1A) the six different viral mRNAs are transcribed by cellular RNA polymerase II. All viral mRNAs are capped and share an identical 3′-end due to the utilization of the same polyadenylation site. The unspliced viral transcripts are exported into the cytoplasm, where translation of the viral polymerase, the core protein, the X protein and the three different surface proteins takes place. The pre-genomic RNA encodes the polymerase and the core protein, is reverse transcribed into genomic minus strand DNA and is also encapsidated (1–3).

Although the viral RNA contains a variety of splice donor and splice acceptor sites the majority of the viral RNAs escape the splicing machinery and are exported in a splice-independent manner. The nuclear export of the unspliced RNAs encoding the surface proteins is facilitated by the *cis*-active hepatitis B virus post-transcriptional regulatory element (HPRE) (4,5), whereas the export mechanism for the unspliced pregenomic RNA is still elusive (6). The HPRE is contained in all viral transcripts at their 3′-end and partially overlaps with the ORFs encoding the protein X and the polymerase.

Splicing of viral RNAs (7–14), the expression of the HBV splice-generated protein (HBSP) (15,16) and a polymerase and surface fusion protein (17) are well documented. However, the function of splicing in the viral life cycle and/or during infection is still enigmatic. In addition to its export function, the HPRE contains a splicing enhancer element indicative for an additional splicing regulatory function (6).

*To whom correspondence should be addressed. Tel: +49 3641 656220; Fax: +49 3641 656225; Email: mago@fli-leibniz.de
Present address:
Martin Schwalbe, Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand.
Aliaksandr Marchanka, Max-Planck-Institut für Bioanorganische Chemie, Stiftstrasse 34-36, D-45470 Mülheim, Germany.
Tilman Heise, Medical University of South Carolina, 173 Ashley Avenue, Charleston SC 29425, USA.

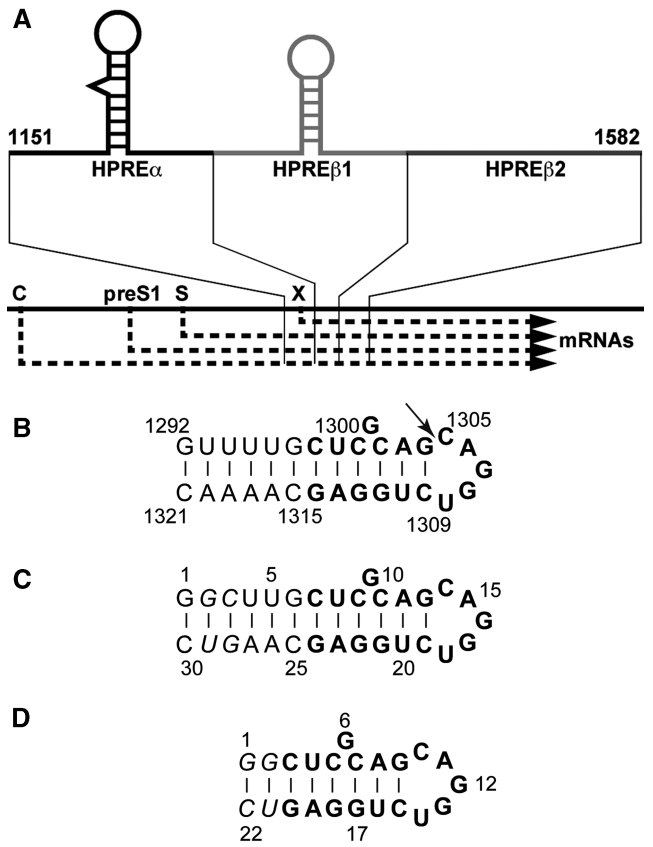

Figure 1. (**A**) Schematic representation of the HBV post-transcriptional regulatory element (HPRE) sub-elements and of the HBV genome. The promoter sites (C, preS1 and X) are indicated. (**B**) Sequences of HPRE SLα and the two synthetic RNAs HSLA (**C**) and HSLAap (**D**). The arrow in (B) indicates the position of the 3′ splice site (see Discussion section). Residues added for *in vitro* transcription purposes are shown in 'italics' in (C) and (D). Residues that are identical in all molecules are given in bold face. Numbering of HPRE SLα sequence is based on the GenBank sequence of HBV strain D00329.

The HPRE is an ~500 nt long tripartite element consisting of the sub-elements α (nucleotides1151–1346), β1 (nucleotides 1347–1457) and β2 (nucleotides 1458–1582) (Figure 1A) (18). All three sub-elements alone can facilitate nuclear export albeit with very low efficiency indicating that all three elements are required for full HPRE activity (18–20). Two stem-loops have been identified in the sub-elements, SLα (nucleotides 1292–1321) and SLβ1 (nucleotides 1410–1434). Disruption of the stem integrity results in a decrease of HPRE activity comparable to the loss of activity after deletion of the whole corresponding sub-element. Compensatory mutations that restore the stem integrity also restored export activity (18). These experiments led to the prediction that both hairpins serve as protein-binding sites for nuclear export factors. Note, that the sequence of the SLα is highly conserved between HBV and the closely related woodchuck hepatitis virus (18).

Secondary structure prediction based on the conserved sequences of SLα indicates that the stem is closed by a pentaloop and that the stem is interrupted by a bulged guanine (Figure 1B). The HPRE facilitates the nuclear export of viral transcripts most likely by recruiting *trans*-acting host proteins. So far, three proteins, including the poly-pyrimidine tract binding protein (PTB), have been described to bind to the HPRE (21–23). Binding sites for the PTB were identified in the 3′-region of the HPRE and functional PTB-binding sites were shown to be important for HPRE-dependent reporter gene expression (23,24) even though its direct role in HBV mRNA export remains to be shown [(25); see also (26)]. As RNA export is mediated by specific protein factors (26) and RNA–protein interactions frequently involve non-canonical RNA structures (27,28), it is conceivable that the non-canonical parts of the SLα, i.e. the pentaloop and the bulged guanine, might constitute protein recognition sites functionally important in HBV mRNA export. This idea is consistent with recent findings that RNA hairpins containing pentaloop sequences similar to but non-identical with the SLα pentaloop bind SAM domain containing proteins in the yeast *Saccharomyces cerevisiae* (29,30). Nuclear export of unspliced HBV RNAs is a bottleneck in the viral life cycle and represents a potential target for therapeutic intervention. The documented contribution of the SLα domain to the function of the HPRE and its potential to be used as a drug target served as incentive to determine the structure of SLα.

MATERIALS AND METHODS

Synthesis, purification and preparation of RNA

Labelled nucleotides and the stem-loop α of the HPRE (HSLA, Figure 1C, nucleotides 1295–1318) were prepared and purified in uniformly ^{15}N-labelled and in uniformly ^{13}C,^{15}N-labelled form by *in vitro* transcription with T7 RNA polymerase as described previously (31,32). In addition, a short version of the apical part of the stem-loop α of the HPRE (HSLAap, Figure 1D, nucleotides 1298–1315) was prepared in ^{13}C,^{15}N-labelled form. The NMR samples had final concentrations of 0.6 mM [^{15}N]-HSLA, 1.0 mM [^{13}C,^{15}N]-HSLA in H_2O and ^2H_2O, respectively, and 1.0 mM [^{13}C,^{15}N]-HSLAap in H_2O and ^2H_2O, respectively. The NMR buffer contained 10 mM KH_2PO_4/K_2HPO_4, pH 6.2, 40 mM KCl, 0.2 mM EDTA and 10% ($^v/_v$) ^2H_2O or 99.99% ($^v/_v$) ^2H_2O, respectively.

In order to ensure the folding of samples into a monomeric stem-loop species, all samples were denatured by heating to 95°C for 10 min and re-annealed by rapid quenching of the sample temperature by 10-fold dilution with ice cold H_2O. After incubation for 1 h on ice the folding state was analysed by a native polyacrylamide gel electrophoresis (data not shown). Folded samples were exchanged into NMR buffer using Centricon ultrafiltration devices.

NMR spectroscopy

NMR spectra were recorded at temperatures of 283 K (exchangeable protons) and 293 K (exchangeable and non-exchangeable protons) on Varian $^{\text{UNITY}}$*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, USA) and analysed using the program CARA (33). Chemical shifts were referenced as described previously (34). Resonance assignment of the exchangeable and non-exchangeable protons of HSLA and HSLAap was performed as previously described in (31,32,34–36) using (^{1}H,^{15}N)-HSQC (37), (^{1}H,^{15}N)-CPMG-NOESY (38), (^{1}H,^{1}H)-H5(C5C4N)H (C,U) (39), (^{1}H,^{1}H)-HCCNH-TOCSY (G) (40), (^{1}H,^{1}H)-HCCH-TOCSY (A) (41), (^{1}H,^{1}H)-HCCH-COSY (C,U) (42), (^{1}H,^{15}N)-2bond-HSQC (43), (^{1}H,^{13}C)-H(N)CO (44), (^{1}H,^{13}C)-HSQC (37,45), 3D (^{1}H,^{1}H,^{13}C)-HCCH-COSY (sugar), 3D (^{1}H,^{1}H,^{13}C)-HCCH-TOCSY (sugar) (42) and (^{1}H,^{15}N)-HCN (46). (^{1}H,^{1}H)-NOESY (47) and 3D (^{1}H,^{1}H,^{13}C)-NOESY-HSQC (48) spectra were acquired for sequential assignments. The virtually full assignment for HSLAap allowed for completion of assignments for corresponding nucleotides in HSLA.

Partial alignment of RNA for residual dipolar coupling (RDC) measurements was achieved by adding 20 mg/ml of Pf1 bacteriophages (Profos AG, Regensburg, Germany) to the uniformly labelled sample (49).

^{13}C relaxation parameters T_1, $T_{1\rho}$ and {^{1}H}-^{13}C steady-state heteronuclear NOE (HETNOE) for C1′ and aromatic carbons were measured by using gradient-enhanced pulse sequences to minimise water saturation (50). T_1 and $T_{1\rho}$ rates were determined by fitting the peak intensities at multiple relaxation delays to the equation $I = I_0 \exp(-(\tau/T_1))$ (51) using RELAXFIT (52) (see also Supplementary Data). The programs ROTDIF (53) and DYNAMICS (52) were employed for relaxation analysis. {^{1}H}-^{13}C steady-state heteronuclear NOEs were calculated as the ratio of ^{1}H-^{13}C correlation peak intensities in the spectra acquired with and without 3 s of proton saturation during the 5s recycle delay. Data were analysed for the individual spin pairs without taking into account other relaxation pathways in the uniformly labelled sample according to Shajani and Varani (54).

Structure calculations

NOE constraints for non-exchangeable hydrogens were obtained from 3D (^{1}H,^{1}H,^{13}C)-NOESY-HSQC spectra for both RNAs. Separate data sets were acquired for ribose and aromatic carbons with mixing times of 80 and 100 ms for HSLAap and 120 and 150 ms, respectively, for HSLA. Resulting NOEs were classified into three groups representing strong, medium and weak cross peaks and assigned to upper distances of 2.8, 4.3 and 5.9 Å, respectively. NOE constraints for exchangeable hydrogens were derived from (^{1}H,^{1}H)-NOESY and (^{1}H,^{15}N)-CPMG-NOESY spectra. Mixing times were 100 and 250 ms for HSLAap and 150 and 250 ms for HSLA. NOEs of exchangeable hydrogens were related to upper limit distance constraints of 4.3 and 5.9 Å. For calibration of cross peak intensities, intra-residue H5-H6 and intra-base pair imino to amino NOEs were used. Hydrogen bond constraints were introduced for all base pairs exhibiting cross peaks in an HNN-COSY spectrum (55) as well as for the penultimate GU wobble pair. For every AU and GU pair, four upper and lower distance constraints were used and six for every GC pair, respectively. 1D ^{31}P spectra showed no downfield shifted resonances, therefore the backbone torsion α and ζ of base-paired residues were restricted to a range of $\pm 60^\circ$ of A-form helical values according to (56,57) resulting in a total of 31 phosphorus shift-derived constraints for α and ζ. No such restrictions were imposed for loop and bulge residues. 14 $^3J_{H1'H2'}$, 7 $^3J_{H2'H3'}$ and 4 $^3J_{H3'H4'}$ coupling constants, extracted from a 2D (^{1}H,^{13}C)-HCCH-E.COSY (58), were used to restrain the sugar pucker for HSLAap.

Further torsion angle constraints were derived from local conformational analysis with the FOUND module (59) using 529 NOE derived upper limit distance constraints as well as the scalar couplings, torsion angle and hydrogen bond constraints mentioned earlier. As a result, 207 constraints defining 185 torsion angles in HSLAap were used for the distance geometry structure calculations employing CYANA (60) in addition to the distance constraints. No planarity constraints for restraining the base pairs were used. Energy minimization of the 15 out of 100 conformers with the lowest target function was performed with OPAL (61). The implemented AMBER94 force field (62) was used and a scaling of the electronic interaction by 0.5 to avoid excessive charge on the RNA backbone for *in vacuo* calculations was applied. MOLMOL (63) was used for visualization of the resulting structures. Back-calculation of the RDCs was performed using the program PALES (64).

RESULTS

Resonance assignment and qualitative structural features

The HSLA molecule representing the native SLα of the HPRE (Figure 1A and B) was modified by replacing two base pairs in the lower stem for *in vitro* transcription purposes (Figure 1C). Spectral overlap in this 30 nt long RNA initially prevented a full sequential assignment and extraction of structural parameters e.g. $^3J_{H1'H2'}$ coupling constraints. However, for a 22 nt long HSLAap molecule (Figure 1D) where the A-helical stem was shortened as compared to HSLA a virtually full resonance and sequential assignment was achieved. The diminished spectral overlap allowed for the extraction of a comprehensive set of structure restraining parameters. A comparison of (^{1}H,^{15}N)-HSQC (data not shown) and (^{1}H,^{13}C)-HSQC spectra (Figure 2A) of both molecules showed very similar or identical chemical shifts for the apical part of each molecule (residues 9–22 for HSLA and 5–18 for HSLAap, respectively; Figure 2A), indicating the same conformation for the apical part of SLα. To simplify matters, the structure of the HSLAap will be discussed here in lieu of HPRE SLα, as it contains all irregular elements of the full-length molecule.

Analysis of hydrogen bonding in HSLAap by HNN-COSY revealed seven correlations between U and G imino groups and N1 of A and N3 of C, respectively (Figure 2B). This result is consistent with the formation

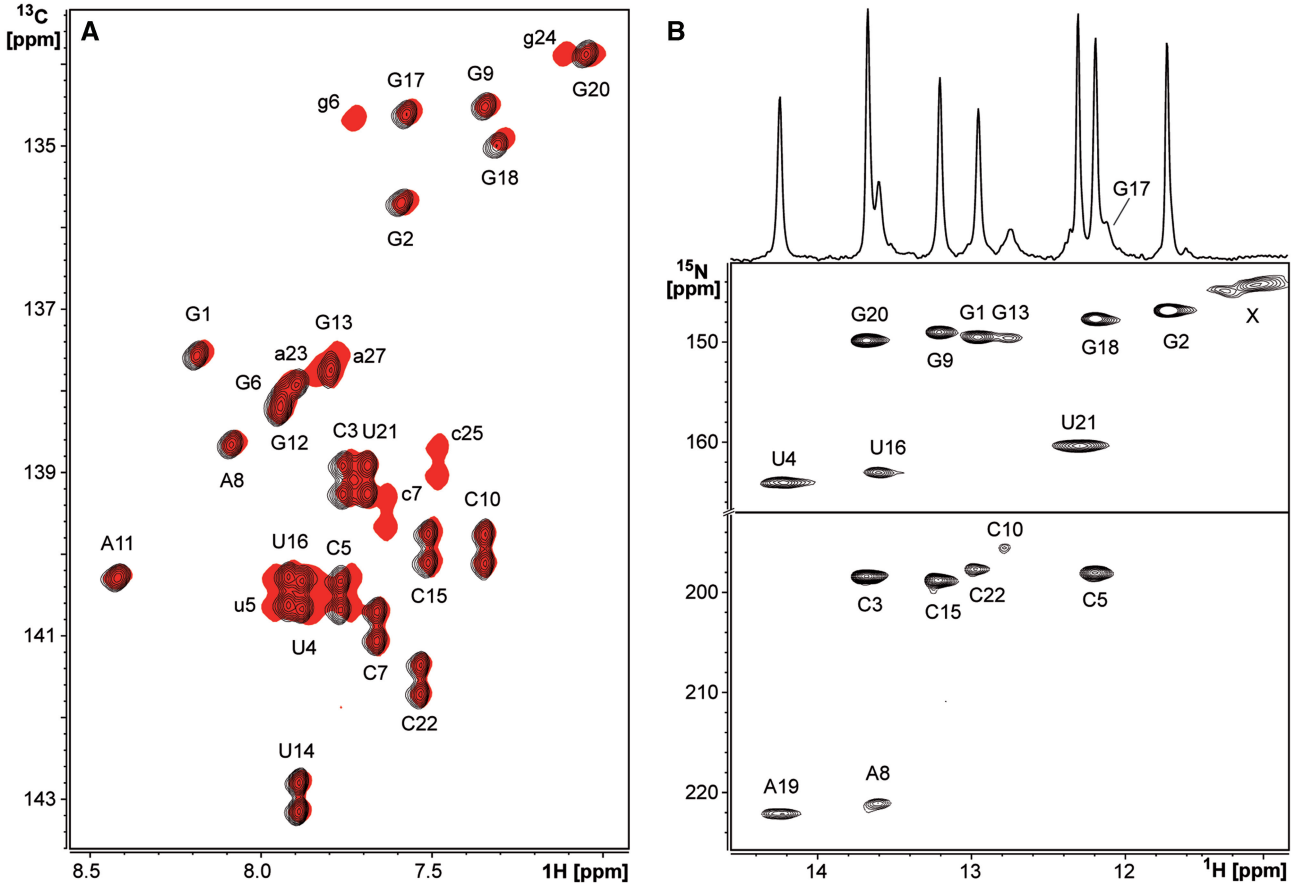

Figure 2. (**A**) Superimposition of the aromatic regions of the (^{1}H,^{13}C)-HSQC spectra of HSLA (red solid contours) and HSLAap (black contours) recorded at 20°C. The assignments for HSLAap are indicated. Resonances arising from residues only present in HSLA are indicated by lower case letters. (**B**) HNN-COSY spectrum of HSLAap recorded at 10°C; 1D spectrum of the imino region shown on top of (B). The assignments for the imino groups of HSLAap are indicated, X denotes broad overlapping signals as discussed in the text.

of 2 AU and 5 GC base pairs in HSLAap. Importantly, the formation of a Watson–Crick GC base pair by the first and fourth loop residues C10 and G13 is indicated by an—albeit weak—correlation in the HNN-COSY spectrum (Figure 2B) and NOEs between the imino proton of G13 and the amino groups of C10 (see Figure S1A in the Supplementary Data). An unusual base pair geometry for C10G13, as listed in a comprehensive survey (65), would not be compatible with these observations. In the (^{1}H,^{15}N)-HSQC spectrum (data not shown, cf. Figure 2B upper panel), the 1D ^{1}H spectrum (Figure 2B top) and the 2D (^{1}H,^{1}H)-NOESY spectrum (see Figure S2 in the Supplementary Data) further imino groups for G residues and one U residue were found showing no correlation in the HNN-COSY spectrum. The U imino group and one G imino group exhibited a strong cross peak in the (^{1}H,^{1}H)-NOESY spectrum (Figure S2) and were assigned to the penultimate G2U21 wobble pair. One broad G imino resonance (Figure 1B, top) giving rise to weak NOE cross peaks to the amino group of C7 (see Figure S1B in Supplementary Data) indicated a Watson–Crick GC base pair and was assigned to G17. No signal for the imino group of U14 and only degenerate candidate imino signals, most likely owing to increased exchange with the solvent, for the bulged G6 residue and the loop G12 could be detected (Figure 2B, X).

The HPRE SLα pentaloop forms a tetraloop-like structure

The structure of the HPRE SLα is represented by the ensemble of 15 energy-minimized structures of HSLAap with the lowest target function (Figure 3A). The structural statistics is given in Table 1. The structure consists of an A-helical stem, a G-bulge and a CAGGU-loop residing upon a GC closing base pair.

A-helix. The calculated NMR solution structure of HSLAap contains an A-helical portion formed by the residues 1–5, 7–9 and 15–22. Sequential NOE connectivities for H1′ to H8 of purines or H6 of pyrimidines, respectively, indicative for an A-form helix are shown in Figure S3 of the Supplementary Data. In addition, residual dipolar couplings determined for the G and U imino groups via (^{1}H,^{15}N)-HSQC measurements closely match those back-calculated for the A-helical portion of the HSLAap structure (Figure S4). The elongated stem region in HSLA (residues 1–7 and 24–30), which is missing in HSLAap, also adopts a standard A-form conformation

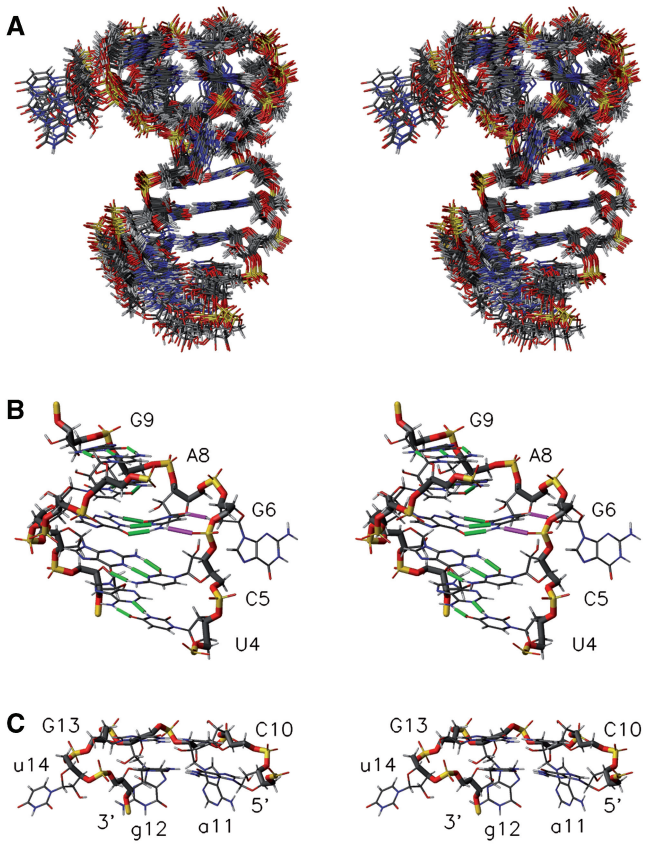

Figure 3. (**A**) Superimposition of the 15 energy-minimized conformations representing the solution structure of HSLAap. (**B**) Close-up of the G6-bulge region. Canonical hydrogen bonds are shown in green and hydrogen bonds stabilising the bulged G in magenta. View rotated by 90° relative to (A). (**C**) Close-up of the loop region. The trace of the backbone is emphasised. Non-base-paired residues are marked by lower case characters. View rotated by −90° relative to (A). Side-by-side stereo depictions are shown.

Table 1. NMR and refinement statistics for the HBV HPRE stem-loop α structure (HSLAap)

NMR distance and dihedral constraints	
Distance constraints	529
Hydrogen bond constraints	48
Total dihedral angle restraints	207
Structure statistics	
Violations	
Target function (Å^2)	0.260 (0.00378)
Distance constraints > 0.1 Å	3.87 (0.96)
Max. distance constraint violation (Å)	0.18 (0.02)
Dihedral angle constraints	0 (0)
Max. dihedral angle violation (°)	1.71 (0.07)
AMBER physical energies (kcal/mol)	−161.68 (8.76)
Deviations from idealized geometry	
Bond lengths (Å)	0.0047 (0.00006)
Bond angles (°)	1.4115 (0.01533)
Mean global r.m.s.d. (Å)	
Heavy atoms	1.22 (0.40)
Pentaloop (residues 9–15)	0.74 (0.25)

(data not shown). This A-helical portion of SLα will not be discussed any further.

G-bulge. In the calculated structure the G6 residue is bulged out of the A-helix (Figure 3B) consistent with interresidual NOEs connecting the G6H8 to H2′, H3′ and H4′ of C5 as well as an NOE connecting the G6H8 to H5 of residue C7 (see Figure S1C). The H8G6 resonates at 7.94 p.p.m., which is very close to the H8 shift of GMP under identical buffer conditions (8.06 p.p.m.; data not shown) strongly suggesting the absence of significant ring current effects (66). In addition, neither imino (Figure 2B) nor amino protons were observed, indicating fast exchange with the solvent. Taken together, these observations are consistent with a bulged-out conformation of G6. The bulged G6 does not significantly disturb the A-helical conformation of the stem region. (^{1}H,^{1}H)-NOE spectra of a synthetic deletion mutant of HSLAap lacking the bulged G6 residue resulted in an imino–imino NOE pattern virtually identical to the respective wild-type HSLAap pattern (data not shown). The $^3J_{H1'H2'}$ coupling constant of the bulged G6 suggests an equilibrium between C3′- and C2′-*endo* conformations (67) with a fraction of approximately 30% C3′-*endo*. Hydrogen bonds between the OH2′ of G6 and O4′ of C7 as well as a phosphate oxygen of G6 and an amino proton of the C7 base stabilizing the bulged G conformation can be inferred from the calculated structure (Figure 3B). The glycosidic torsion angle χ resides in the high-*anti* region.

gCAGGUc-loop. The pentaloop (Figure 3C) adopts a well-defined, essentially tetraloop-like conformation. The observed ^{1}H/^{13}C resonances of the ribose CH groups A11-U14 were well separated from the respective helical ones (e.g. Supplementary Figure S5 for C1′H1′ groups), which allowed for their reliable assignment. Such unusual chemical shifts for the loop nucleotides are reminiscent of observations for other tetraloop conformations [cf. (32, 36, 66, 68)]. The first and fourth loop residues C10 and G13 form a Watson–Crick base pair, which stacks upon the closing base pair G9C15 as also indicated by a weak imino–imino NOE signal between G9 and G13 (Figure S2). The rise between both base pairs is 3.22 Å as is typical for A-form RNA. While all other loop nucleotides show χ angle in standard *anti* conformations the formation of the C10G13 base pair leads to a mean G13 χ angle of −23°, an intermediate between the high-*anti* and *syn* ranges. This is reflected by an NOE signal between the H8 and the H1′ of G13, which is significantly stronger than observed for the G residues in the helical part [Figure S6; see also (69)]. $^3J_{H1'H2'}$ coupling constants between 5 and 6 Hz observed for the three loop residues C10, G13 and U14 indicate a C3′-/C2′-*endo* conformational equilibrium of 30%/70%. The second loop residue A11 and the third loop residue G12 protrude into the minor groove. Both residues display a C2′-*endo* pucker as evident from the $^3J_{H1'H2'}$ coupling constants of 8.0 and 7.5 Hz, respectively. We note that for G12 the chemical shifts of its ribose protons are shifted to lower parts per

million values (H4′: 4.00 p.p.m., H5′: 4.21 p.p.m., H5″: 4.00 p.p.m. and H2′: 4.34 p.p.m.). Similarly, shifted resonances were observed for the equivalent loop residue U7 in the CUU_7G tetraloop (70). Such chemical shifts were discussed as being indicative for a backbone turn at such positions (66,71). The base planes of A11 and G12 are oriented nearly perpendicular to the base planes of the loop pair C10G13 and the G9C15 closing base pair. The structure suggests that the conformation of A11 is stabilized by two sets of mutually exclusive hydrogen bond interactions: either via N1A11–HO2′U16 together with N7A11–H22G9 or, alternatively, via H61A11–O2′G17 and N7A11–HO2′G9. In addition, a hydrogen bond H62A11–O2′G9 and a NH-π-hydrogen bond (72) between the amino group of G9 and the aromatic system of A11 may form. Other loop stabilizing hydrogen bond interactions consistent with the calculated structure can be deduced for: OH2′A11–O2C10, OH2′G13–OP2C15, OH2′G9–O5′C10, O2′G9–HN62A11, OH2′G12–O2′C15, O2′G12–OH2′C15, H22G13–O2′C12, respectively. Finally, the tetraloop-like structure is accommodated by extruding the base moiety of the fifth loop residue U14 into the solvent. This orientation is reflected by a lack of *inter*residual NOEs and by the chemical shift of the H5U14 at 6.07 p.p.m. (66).

The HSLAap shows an overall rotational correlation time of 4.04–4.06 ns when using the axial-symmetric approximation in the model-free approach for data analysis (52,53). The anisotropy value (D∥/D⊥ ratio) of 1.6 derived from the latter approach matches with the value obtained from HYDRONMR (73) using the HSLAap structure. Heteronuclear NOE measurements and relaxation analysis revealed that the sugar and base moieties of all nucleotides including the G6 bulge and the loop residues C10, A11, G12 and G13 display limited flexibility on the ps to ns time scale (see Figure S7 in the Supplementary Data). We however note a slight line broadening for the H2 of A11 but not for its H1′ (not shown) and the H1′ of G12 (cf. Supplementary Figure S5), but not for its H8 proton. While this suggests a limited local 'breathing' on a slower time scale of an otherwise stable tetraloop conformation, the base unit of the fifth loop residue U14 exhibits considerable motion already on the ps to ns time scale as evident from the low order parameter of 0.64.

DISCUSSION

The structural analysis of SLα revealed that its apical pentaloop accommodates a tetraloop fold by bulging out the fifth loop nucleotide and that it belongs to the $CNGG(N)_n$ loop family (30,74), which has been discussed as an extension of the CUNG tetraloop family (70). The results presented here clearly show that significant similarities are apparent to the thermostable CUUG tetraloop (PDB code: 1RNG) (70). Both SLα and the CUUG loop have a GC closing base pair and a CG base pair is formed by the first and fourth loop residue that is stacked upon the closing base pair. A difference is observed for the glycosidic angle of G13, which tends towards a *syn* conformation only in the SLα gCAGGUc pentaloop. The conformation of the second base (U) of the CUUG loop is similar to the conformation of the equivalent residue A11 in the gCAGGUc pentaloop (Figure 4) and the backbone r.m.s.d. for the loop residues (excluding the fifth residue for gCAGGUc) amounts to 1.39 Å (2.06 Å incl. closing base pair). The main difference between the two structures is found for the respective third loop base (U versus G12), which locates to opposite sites of the RNA backbone. Nevertheless, the apical pentaloop of SLα constitutes another example for the sequence plasticity of CUNG-like tetraloop structures. The sequence plasticity observed for CUNG-like tetraloop structures here and elsewhere (30,74,75) resembles the situation which has been described for the other 'canonical' tetraloops UNCG (32,36,69) and GNRA (76,77), respectively. For the latter tetraloop-type a pentaloop sequence may also adopt a tetraloop fold (78). Taken together, these findings attest to a limited structure space available for RNA to achieve a 180° reversal of direction—as necessary in hairpin loop elements—for shorter loop segments and, hence, frequently comes in (subtle) variations of common themes represented by the three 'canonical' tetraloops CUNG, GNRA and UNCG (27,79).

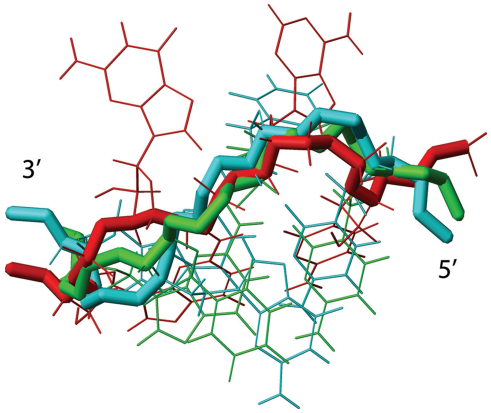

Figure 4. Superimposition of the HSLAap gCAGGUc loop (red) with the SRE uCUGGCa pentaloop (cyan; PDB 2ES5) and the CUUG tetraloop (green; PDB 1RNG). The residues CAGG for HSALap, CUGG for the SRE loop and CUUG for the canonical tetraloop were superimposed and are shown here. The backbone trace is emphasized.

The other members of the $CNGG(N)_n$ family are represented by the uCUGGCa pentaloop of the yeast Smaug recognition element (SRE) and a synthetic uUUGACa pentaloop, both of which form stable complexes with the SAM domain of Vts1p (29,30,75). A superimposition of the gCAGGUc pentaloop with the uCUGGCa pentaloop of the unbound SRE (PDB code: 2ES5) shows a similar trace of the sugar-phosphate-backbone (Figure 4). The backbone r.m.s.d. between gCAGGUc and uCUGGCa including all loop residues is 1.75 Å. Most prominent deviations occur for the first loop position, probably due to compensation effects for the difference in the closing base pair (gc versus ua), and for the third loop nucleotide. In the SRE pentaloop,

the conserved third residue (a G) engages in a stacking interaction with the purines of the loop CG pair and the closing UA Watson–Crick base pairs (30). Thus, it was surprising to observe that the G12 in HSLAap (gCAGGUc) did not adopt a similar stacking conformation. Therefore, we asked whether our data could also be consistent with a uCUGGCa-like stacking conformation for G12. This, however, was clearly not the case. For instance, NOEs observed for HSLAap between H1′G12 and H4′C15 as well as for H4′A11–H5′G12, H1′A11–H2′G12, H5′G13–H1′C15, H4′G13–H4′U14, H2′G13–H4′C15, H1′G13–H5′U14, H5″A11–H4′A11 would be violated for a loop conformation accommodating a stacked G12. In spite of these differences, yet being a member of the CNGG(N)$_n$family it is conceivable for SLα that a protein-binding partner constitutes a SAM-motif containing human homologue of the yeast Vts1p or the Drosophila Smaug with a function in RNA transport. Such a homologue might bind the SLα pentaloop in a manner as observed for Vts1p (29,30). Additional contacts to the G-bulge, only present in SLα, might be necessary for the *in vivo* function of such a homologue.

Importantly, it has been shown recently (80) that the SLα harbours a 3′-splice site at G1304/C1305 of the HBV genome corresponding to positions G9/C10 in HSLap (Figure 1B and D), which is used in conjunction with a 5′-splice site at position 458 to splice the subgenomic RNA. Structured regions of RNA are known to regulate the utilization of splice sites (81). Furthermore, splicing of the HBV related duck Hepatitis virus is suppressed by the formation of secondary RNA structures harbouring 5′- and 3′-splice sites (82). Since the nuclear export of unspliced subgenomic RNAs is essential for the synthesis of viral proteins and, thus, for viral propagation it is important to prevent the utilization of the 3′-splice site present in SLα in order to allow for efficient export of unspliced subgenomic RNAs. Hence, SLα may well serve that function—be it solely by virtue of its tertiary structure described here or in concert with as yet unidentified cellular factors.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank J. Wöhnert for critically reading the article. This work was supported in part by a grant of the Deutsche Forschungsgemeinschaft to T.H. (HE 2814/3-2). The Fritz Lipmann Institute is financially supported by the State of Thuringia and the Federal Government of Germany. Funding to pay the Open Access publication charges for this article was provided by FLI.

Conflict of interest statement. None declared.

REFERENCES

1. Seeger,C. and Mason,W.S. (2000) Hepatitis B virus biology. *Microbiol. Mol. Biol. Rev.*, **64**, 51–68.
2. Hollinger,F.B. and Liang,T.J. (2001) *Hepatitis B Virus*, 4th edn. Lippincott, Williams & Wilkins, Philadelphia.
3. Strauss,J.H. and Strauss,E.G. (2002) *Viruses and Human Disease*. Academic Press, San Diego.
4. Huang,J. and Liang,T.J. (1993) A novel hepatitis B virus (HBV) genetic element with Rev response element-like properties that is essential for expression of HBV gene products. *Mol. Cell. Biol.*, **13**, 7476–7486.
5. Huang,Z.M. and Yen,T.S. (1994) Hepatitis B virus RNA element that facilitates accumulation of surface gene transcripts in the cytoplasm. *J. Virol.*, **68**, 3193–3199.
6. Heise,T., Sommer,G., Reumann,K., Meyer,I., Will,H. and Schaal,H. (2006) The hepatitis B virus PRE contains a splicing regulatory element. *Nucleic Acids Res.*, **34**, 353–363.
7. Su,T.S., Lai,C.J., Huang,J.L., Lin,L.H., Yauk,Y.K., Chang,C.M., Lo,S.J. and Han,S.H. (1989) Hepatitis B virus transcript produced by RNA splicing. *J. Virol.*, **63**, 4011–4018.
8. Suzuki,T., Masui,N., Kajino,K., Saito,I. and Miyamura,T. (1989) Detection and mapping of spliced RNA from a human hepatoma cell line transfected with the hepatitis B virus genome. *Proc. Natl Acad. Sci. USA*, **86**, 8422–8426.
9. Wu,H.L., Chen,P.J., Tu,S.J., Lin,M.H., Lai,M.Y. and Chen,D.S. (1991) Characterization and genetic analysis of alternatively spliced transcripts of hepatitis B virus in infected human liver tissues and transfected HepG2 cells. *J. Virol.*, **65**, 1680–1686.
10. Chen,P.J., Chen,C.R., Sung,J.L. and Chen,D.S. (1989) Identification of a doubly spliced viral transcript joining the separated domains for putative protease and reverse transcriptase of hepatitis B virus. *J. Virol.*, **63**, 4165–4171.
11. Terre,S., Petit,M.A. and Brechot,C. (1991) Defective hepatitis B virus particles are generated by packaging and reverse transcription of spliced viral RNAs in vivo. *J. Virol.*, **65**, 5539–5543.
12. Gunther,S., Sommer,G., Iwanska,A. and Will,H. (1997) Heterogeneity and common features of defective hepatitis B virus genomes derived from spliced pregenomic RNA. *Virology*, **238**, 363–371.
13. Sommer,G., van Bommel,F. and Will,H. (2000) Genotype-specific synthesis and secretion of spliced hepatitis B virus genomes in hepatoma cells. *Virology*, **271**, 371–381.
14. Marschenz,S., Endres,A.S., Brinckmann,A., Heise,T., Kristiansen,G., Nurnberg,P., Kruger,D.H., Gunther,S. and Meisel,H. (2006) Functional analysis of complex hepatitis B virus variants associated with development of liver cirrhosis. *Gastroenterology*, **131**, 765–780.
15. Rosmorduc,O., Petit,M.A., Pol,S., Capel,F., Bortolotti,F., Berthelot,P., Brechot,C. and Kremsdorf,D. (1995) In vivo and in vitro expression of defective hepatitis B virus particles generated by spliced hepatitis B virus RNA. *Hepatology*, **22**, 10–19.
16. Soussan,P., Garreau,F., Zylberberg,H., Ferray,C., Brechot,C. and Kremsdorf,D. (2000) In vivo expression of a new hepatitis B virus protein encoded by a spliced RNA. *J. Clin. Invest.*, **105**, 55–60.
17. Huang,H.L., Jeng,K.S., Hu,C.P., Tsai,C.H., Lo,S.J. and Chang,C. (2000) Identification and characterization of a structural protein of hepatitis B virus: a polymerase and surface fusion protein encoded by a spliced RNA. *Virology*, **275**, 398–410.
18. Smith,G.J. III, Donello,J.E., Luck,R., Steger,G. and Hope,T.J. (1998) The hepatitis B virus post-transcriptional regulatory element contains two conserved RNA stem-loops which are required for function. *Nucleic Acids Res.*, **26**, 4818–4827.
19. Donello,J.E., Beeche,A.A., Smith,G.J. III, Lucero,G.R. and Hope,T.J. (1996) The hepatitis B virus posttranscriptional regulatory element is composed of two subelements. *J. Virol.*, **70**, 4345–4351.
20. Huang,Z.M., Zang,W.Q. and Yen,T.S. (1996) Cellular proteins that bind to the hepatitis B virus posttranscriptional regulatory element. *Virology*, **217**, 573–581.
21. Heise,T., Guidotti,L.G. and Chisari,F.V. (1999) La autoantigen specifically recognizes a predicted stem-loop in hepatitis B virus RNA. *J. Virol.*, **73**, 5767–5776.
22. Zang,W.Q., Fieno,A.M., Grant,R.A. and Yen,T.S. (1998) Identification of glyceraldehyde-3-phosphate dehydrogenase as a cellular protein that binds to the hepatitis B virus posttranscriptional regulatory element. *Virology*, **248**, 46–52.

23. Zang,W.Q., Li,B., Huang,P.Y., Lai,M.M. and Yen,T.S. (2001) Role of polypyrimidine tract binding protein in the function of the hepatitis B virus posttranscriptional regulatory element. *J. Virol.*, **75**, 10779–10786.
24. Li,B. and Yen,T.S. (2002) Characterization of the nuclear export signal of polypyrimidine tract-binding protein. *J. Biol. Chem.*, **277**, 10306–10314.
25. Kamath,R.V., Leary,D.J. and Huang,S. (2001) Nucleocytoplasmic shuttling of polypyrimidine tract-binding protein is uncoupled from RNA export. *Mol. Biol. Cell*, **12**, 3808–3820.
26. Kohler,A. and Hurt,E. (2007) Exporting RNA from the nucleus to the cytoplasm. *Nat. Rev.*, **8**, 761–773.
27. Hermann,T. and Patel,D.J. (1999) Stitching together RNA tertiary architectures. *J. Mol. Biol.*, **294**, 829–849.
28. Hermann,T. and Westhof,E. (1999) Non-Watson-Crick base pairs in RNA-protein recognition. *Chem. Biol.*, **6**, R335—R343.
29. Aviv,T., Lin,Z., Ben-Ari,G., Smibert,C.A. and Sicheri,F. (2006) Sequence-specific recognition of RNA hairpins by the SAM domain of Vts1p. *Nat. Struct. Mol. Biol.*, **13**, 168–176.
30. Oberstrass,F.C., Lee,A., Stefl,R., Janis,M., Chanfreau,G. and Allain,F.H. (2006) Shape-specific recognition in the structure of the Vts1p SAM domain with RNA. *Nat. Struct. Mol. Biol.*, **13**, 160–167.
31. Stoldt,M., Wohnert,J., Ohlenschlager,O., Gorlach,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.
32. 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.
33. Keller,R. (2004) *The Computer Aided Resonance Assignment Tutorial* CANTINA, Goldau.
34. Sich,C., Ohlenschlager,O., Ramachandran,R., Gorlach,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.
35. Ohlenschlager,O., Wohnert,J., Ramachandran,R., Sich,C. and Gorlach,M. (2003) Nuclear magnetic resonance studies of ribonucleic acids. *Spectroscopy*, **17**, 537–547.
36. 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 yYNMGg-type backbone conformation from cloverleaf 1 of bovine enterovirus 1 RNA. *Nucleic Acids Res.*, **33**, 2003–2011.
37. Bodenhausen,G. and Ruben,D.J. (1980) Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.*, **69**, 185–188.
38. 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.
39. Wohnert,J., Ramachandran,R., Gorlach,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.
40. 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.
41. Marino,J.P., Prestegard,J.H. and Crothers,P.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.
42. 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.
43. 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.
44. 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.
45. Santoro,J. and King,G.C. (1992) A constant-time 2D Overbodenhausen experiment for inverse correlation of isotopically enriched species. *J. Magn. Reson.*, **97**, 202–207.
46. 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.
47. Kumar,A., Ernst,R.R. and Wuthrich,K. (1980) A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochem. Biophys. Res. Commun.*, **95**, 1–6.
48. 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.
49. 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.
50. Farrow,N.A., Muhandiram,R., Singer,A.U., Pascal,S.M., Kay,C.M., Gish,G., Shoelson,S.E., Pawson,T., Forman-Kay,J.D. and Kay,L.E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry*, **33**, 5984–6003.
51. Viles,J.H., Duggan,B.M., Zaborowski,E., Schwarzinger,S., Huntley,J.J., Kroon,G.J., Dyson,H.J. and Wright,P.E. (2001) Potential bias in NMR relaxation data introduced by peak intensity analysis and curve fitting methods. *J. Biomol. NMR*, **21**, 1–9.
52. Fushman,D., Cahill,S. and Cowburn,D. (1997) The main-chain dynamics of the dynamin pleckstrin homology (PH) domain in solution: analysis of 15N relaxation with monomer/dimer equilibration. *J. Mol. Biol.*, **266**, 173–194.
53. Walker,O., Varadan,R. and Fushman,D. (2004) Efficient and accurate determination of the overall rotational diffusion tensor of a molecule from (15)N relaxation data using computer program ROTDIF. *J. Magn. Reson.*, **168**, 336–345.
54. Shajani,Z. and Varani,G. (2005) 13C NMR relaxation studies of RNA base and ribose nuclei reveal a complex pattern of motions in the RNA binding site for human U1A protein. *J. Mol. Biol.*, **349**, 699–715.
55. 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**, 1601–1602.
56. Gorenstein,D.G. (1984) *Phosphorus-31 NMR: Principles and Applications*. Academic Press, Orlando, FL.
57. Varani,G., Cheong,C. and Tinoco,I. Jr. (1991) Structure of an unusually stable RNA hairpin. *Biochemistry*, **30**, 3280–3289.
58. Griesinger,C. and Eggenberger,U. (1992) Determination of proton-proton coupling constants in 13C-labeled molecules. *J. Magn. Reson.*, **97**, 426–434.
59. Guntert,P., Billeter,M., Ohlenschlager,O., Brown,L.R. and Wuthrich,K. (1998) Conformational analysis of protein and nucleic acid fragments with the new grid search algorithm FOUND. *J. Biomol. NMR*, **12**, 543–548.
60. Guntert,P., Mumenthaler,C. and Wuthrich,K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.*, **273**, 283–298.
61. Luginbuhl,P., Guntert,P., Billeter,M. and Wuthrich,K. (1996) The new program OPAL for molecular dynamics simulations and energy refinements of biological macromolecules. *J. Biomol. NMR*, **8**, 136–146.
62. 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.
63. Koradi,R., Billeter,M. and Wuthrich,K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.*, **14**, 51–55.
64. Zweckstetter,M. and Bax,A. (2000) *J. Am. Chem. Soc.*, **122**, 3791–3792.

65. 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.
66. Cromsigt,J.A., Hilbers,C.W. and Wijmenga,S.S. (2001) Prediction of proton chemical shifts in RNA. Their use in structure refinement and validation. *J. Biomol. NMR*, **21**, 11–29.
67. van Wijk,J., Huckriede,B.D., Ippel,J.H. and Altona,C. (1992) Furanose sugar conformations in DNA from NMR coupling constants. *Methods Enzymol.*, **211**, 286–306.
68. Furtig,B., Richter,C., Wohnert,J. and Schwalbe,H. (2003) NMR spectroscopy of RNA. *Chembiochem.*, **4**, 936–962.
69. Allain,F.H. and Varani,G. (1995) Structure of the P1 helix from group I self-splicing introns. *J. Mol. Biol.*, **250**, 333–353.
70. Jucker,F.M. and Pardi,A. (1995) Solution structure of the CUUG hairpin loop: a novel RNA tetraloop motif. *Biochemistry*, **34**, 14416–14427.
71. Cromsigt,J., van Buuren,B., Schleucher,J. and Wijmenga,S. (2001) Resonance assignment and structure determination for RNA. *Methods Enzymol.*, **338**, 371–399.
72. Weiss,M.S., Brandl,M., Suhnel,J., Pal,D. and Hilgenfeld,R. (2001) More hydrogen bonds for the (structural) biologist. *Trends Biochem. Sci.*, **26**, 521–523.
73. Garcia de la Torre,J., Huertas,M.L. and Carrasco,B. (2000) HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J. Magn. Reson.*, **147**, 138–146.
74. Aviv,T., Amborski,A.N., Zhao,X.S., Kwan,J.J., Johnson,P.E., Sicheri,F. and Donaldson,L.W. (2006) The NMR and X-ray structures of the Saccharomyces cerevisiae Vts1 SAM domain define a surface for the recognition of RNA hairpins. *J. Mol. Biol.*, **356**, 274–279.
75. Johnson,P.E. and Donaldson,L.W. (2006) RNA recognition by the Vts1p SAM domain. *Nat. Struct. Mol. Biol.*, **13**, 177–178.
76. 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.
77. 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.
78. Legault,P., Li,J., Mogridge,J., Kay,L.E. and Greenblatt,J. (1998) NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. *Cell*, **93**, 289–299.
79. Varani,G. (1995) Exceptionally stable nucleic acid hairpins. *Annu. Rev. Biophys. Biomol. Struct.*, **24**, 379–404.
80. Hass,M., Hannoun,C., Kalinina,T., Sommer,G., Manegold,C. and Gunther,S. (2005) Functional analysis of hepatitis B virus reactivating in hepatitis B surface antigen-negative individuals. *Hepatology*, **42**, 93–103.
81. Buratti,E. and Baralle,F.E. (2004) Influence of RNA secondary structure on the pre-mRNA splicing process. *Mol. Cell. Biol.*, **24**, 10505–10514.
82. Loeb,D.D., Mack,A.A. and Tian,R. (2002) A secondary structure that contains the 5′ and 3′ splice sites suppresses splicing of duck hepatitis B virus pregenomic RNA. *J. Virol.*, **76**, 10195–10202.