Structure of the K-turn U4 RNA: a combined NMR and SANS study

Melanie Falb[1], Irene Amata[1], Frank Gabel[2], Bernd Simon[1] and Teresa Carlomagno[1,*]

[1]European Molecular Biology Laboratory, Structural and Computational Biology Unit, D-69117 Heidelberg, Germany and [2]Institut de Biologie Structurale Jean-Pierre Ebel, CEA, CNRS, UJF UMR 5075, 38027 Grenoble, France

ABSTRACT

K-turn motifs are universal RNA structural elements providing a binding platform for proteins in several cellular contexts. Their characteristic is a sharp kink in the phosphate backbone that puts the two helical stems of the protein-bound RNA at an angle of 60°. However, to date no high-resolution structure of a naked K-turn motif is available. Here, we present the first structural investigation at atomic resolution of an unbound K-turn RNA (the spliceosomal U4-Kt RNA) by a combination of NMR and small-angle neutron scattering data. With this study, we wish to address the question whether the K-turn structural motif assumes the sharply kinked conformation in the absence of protein binders and divalent cations. Previous studies have addressed this question by fluorescence resonance energy transfer, biochemical assays and molecular dynamics simulations, suggesting that the K-turn RNAs exist in equilibrium between a kinked conformation, which is competent for protein binding, and a more extended conformation, with the population distribution depending on the concentration of divalent cations. Our data shows that the U4-Kt RNA predominantly assumes the more extended conformation in the absence of proteins and divalent cations. The internal loop region is well structured but adopts a different conformation from the one observed in complex with proteins. Our data suggests that the K-turn consensus sequence does not *per se* code for the kinked conformation; instead the sharp backbone kink requires to be stabilized by protein binders.

INTRODUCTION

The past two decades have seen an impressive increase in the structural information available for RNA molecules, mostly due to the atomic structures of the ribosome (1–3) and of large catalytic RNAs (4). A wealth of RNA structural motifs other than A-form helical regions have been identified, such as U-turns, S-turns, A-platforms, loop-E motifs, Z-anchors, ribose zippers and others (5). These non-helical RNA regions underline the multiplicity of conformations that the RNA can assume in spite of the limited diversity of the nucleotide building blocks and confer the RNAs a high versatility either in the interaction with protein partners or in the optimization of folds for specific catalytic activity. On the other hand, the vast conformational space available to RNA molecules represents a challenge for structural biologists in the identification of clear structure–sequence relationships and in the formulation of rules to predict RNA structure from the primary nucleotide sequence.

In a cellular environment, RNA molecules are found associated with proteins in ribonucleoprotein (RNP) complexes or with other RNAs, as for example in the case of RNA enzymes or regulatory RNAs. The assembly of these complexes is generally regulated by the propensity of the RNA to assume the bound conformation, an event that might or might not require the presence of cofactors, such as divalent cations. An understanding of the sequence–structure and structure–function relationships in RNA requires collecting structural information for a large number of RNA molecules, both in the presence and absence of their binding partners or cofactors.

K-turn motifs are universal RNA structural elements found in rRNAs (6), small nuclear (sn) RNAs (7), untranslated regions of mRNAs (8,9) and small nucleolar RNAs (snoRNAs) (10) in Archea, Prokarya and Eukarya. They represent ubiquitous protein-binding platforms and are central components of RNP complexes of

---

diverse functions, such as ribosome, spliceosomal U4 and RNA processing enzymes. The K-turn is a two stranded helix–loop–helix motif (Figure 1A). The internal asymmetric loop, usually consisting of 3 nt on one strand and none on the other strand, is flanked by two stems: the non-canonical (NC)-stem starts with two non-canonical base pairs, usually sheared G–A base pairs, while the canonical (C)-stem is a normal duplex commonly starting with a C–G base pair. The K-turn motif owes its name to the sharp bend in the phosphodiester backbone observed when the RNA is bound to its cognate proteins (Figure 1B; 6,7,11); this bend puts the two NC- and C-stems at an angle of 60° to each other. In the asymmetric loop, the nucleotide on the 5′ site of the longer strand, usually a purine, stacks on the C-stem, while the following one, commonly also a purine, stacks on the NC-stem and the third nucleotide protrudes into the protein-binding pocket (12). The K-turn fold is stabilized by long-range contacts between the minor groove edges of the adenines of the G–A base pairs and the minor groove of the C-stem (A-minor interaction) (13).

The characteristic sharp kink in the phosphate backbone has been observed for K-turn RNAs in complex with proteins or in the context of large RNP complexes, while structural information at atomic resolution is not available for the K-turn consensus sequence in isolation. Thus, the question arises whether the consensus sequence of the K-turn motif shown in Figure 1A does *per se* code for the K-turn structural element or whether the kinked backbone conformation is induced and stabilized by electrostatic interactions with the positively charged side chains of the cognate protein or in alternative requires the presence of divalent cations.

Both experimental and theoretical approaches have been used in the past to characterize the structural preferences of K-turn RNAs. An extensive experimental study applies both gel mobility and fluorescence assays to the ribosomal Kt-7 (6), to assess the presence of the K-turn

structural motif in the absence of protein binders and in dependence of the concentration of divalent cations (14). In this work, it has been suggested that K-turn RNAs inter-convert between an extended and a kinked conformation with the populations' distribution being dependent on the concentration of divalent cations.

In another study, the stability of the K-turn fold in the absence of proteins has been assessed by molecular dynamics (MD) simulations (15,16) for the spliceosomal U4 K-turn RNA (7,11,17; Figure 1). The MD trajectories of the unbound RNA show a transition from the kinked to a more extended conformation, occurring in concurrence with opening of the G–A base pairs and ultimately with the disruption of the whole NC-stem. In another MD study (18), the kinked conformation was found to be stable also in the absence of protein binders for a variety of K-turn RNAs, while later, the same laboratory reported on the transition from the kinked conformation to a second structure with a significantly larger angle between the two stems (19).

Despite the deep insights that these studies provided into the stability of the K-turn motif, they lack experimental structural data at an atomic level. While it seems clear that the K-turn fold is stabilized by the interaction with proteins, the question whether the consensus sequence of Figure 1A codes for a kinked, tightly structured RNA motif in the absence of binding partners remains open.

In an early stage during spliceosome assembly, the U6 snRNA is associated with the U4 snRNA in the U4/U6–U5 complex (20,21). The formation of the U4/U6 complex is initiated by the recognition of the 5′-stem–loop (SL) region of the U4 snRNA (U4-Kt1; Figure 1A), located between the two U4/U6 base paired regions, by the 15.5K protein (Supplementary Figure S1; 17). The K-turn motif of the U4 snRNA serves as a binding platform for 15.5K (7,11).

Here, we present the structure of the U4 snRNA K-turn motif solved by NMR spectroscopy in solution and in the absence of protein binders and divalent cations (U4-Kt2; Figure 1A). In this work, we demonstrate that both stems of the U4-Kt RNA are stably folded; in contrast to the results of published MD studies (15,16), we observe that both G–A base pairs of the NC-stem are readily present in solution in the absence of proteins and divalent cations. The average inter-stem angle is also well defined and deviates by ~45° from the value of 60° observed for the sharply kinked crystallographic K-turn. The solution structure of the U4-Kt RNA is more extended, as confirmed independently by NMR and small-angle neutron scattering (SANS) data. The internal asymmetric loop assumes a well-defined conformation, which differs from that observed in complex with proteins.
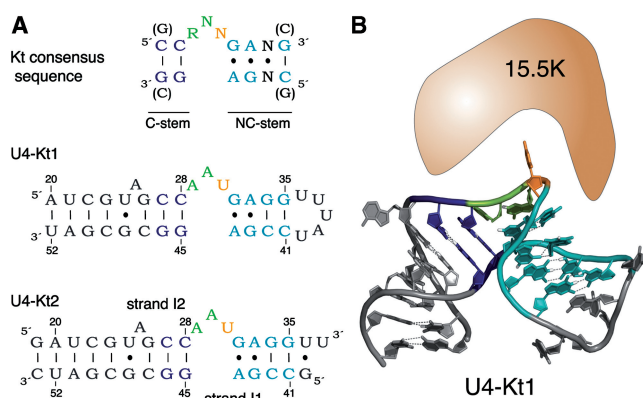

Figure 1. (**A**) Consensus sequence of the K-turn motif, U4 5′-SL RNA (U4-Kt1) containing the K-turn element, and construct derived from the U4 5′-SL RNA (U4-Kt2) used for the NMR analysis. Conserved nucleotides are colour-coded: C-stem, blue; NC-stem, light blue; internal loop, green and orange. (**B**) Structure of the U4-Kt1 bound to 15.5K in the context of the U4 5′-SL RNA–15.5 K–hPrp31 complex (11).

MATERIALS AND METHODS

Preparation of RNA samples

The RNA strands 5′-GCCGAGGCGCGAUC-3′ and 5′-GAUCGUAGCCAAUGAGGUU-3′ were designed to form the secondary structure shown in Figure 1A. This 'open-loop' version of the U4-Kt RNA (U4-Kt2) was

chosen instead of U4-Kt1, as *in vitro* transcription of U4-Kt1 did not yield enough RNA for NMR structural studies at reasonable costs. The duplex U4-Kt2 is competent for binding 15.5K (data not shown). The $^{13}C/^{15}N$-labelled RNA was synthesized by *in vitro* transcription using T7-polymerase produced *in-house*, $^{13}C/^{15}N$-labelled rNTPs (Spectra Stable Isotopes) and synthetic single-stranded DNA templates (IBA Göttingen). The RNA was purified by denaturing 15% polyacrylamide gel electrophoresis. Unlabelled *trans*-acting hammerhead RNA was used to overcome the 3′ heterogeneity (22). Unlabelled RNA strands were purchased (IBA, Göttingen).

Two NMR samples at 0.3 mM duplex concentration were prepared by dissolving an unlabelled RNA strand and a $^{13}C/^{15}N$-labelled strand in 0.3 ml buffer (20 mM HEPES, 120 mM NaCl, pH 7.6). For a uniformly $^{13}C/^{15}N$-labelled NMR sample and for the SANS sample, the duplex was further purified by gel filtration chromatography to eliminate the excess of one of the two strands. NMR experiments involving exchangeable protons were performed in a $H_2O:D_2O$ mixture (9:1). All other experiments were performed in 99.96% D_2O (Sigma-Aldrich). For residual dipolar coupling (RDC) experiments 10–15 mg/ml Pf1 phages (ASLA Biotech, Riga, Latvia) were added, resulting in a splitting of the deuterium solvent line of ~10 Hz.

NMR spectroscopy

NMR spectra were recorded on Bruker Avance 600, 700, 800 and 900 MHz spectrometers. A variety of homo- and heteronuclear experiments were used to assign the RNA resonances and to generate structural data: 3D (^1H, ^{13}C, ^{15}N) H_bCN_b (23), 3D (^1H, ^{13}C, ^{15}N) H_sCN_b (24,25), 3D HCCH-COSY (correlated spectroscopy)-TOCSY (total correlated spectroscopy) (mixing time 5.4 ms) (26), 3D HCP (27), 3D HCCH-E.COSY (exclusive COSY) (28) and 3D (^1H, ^{13}C, ^1H) edited/filtered NOESY (nuclear Overhauser enhancement spectroscopy) (mixing times 150 and 300 ms) (29). The ribose spin systems were assigned from the 3D HCCH-COSY-TOCSY spectrum, while the correlation between the H1′ and the H6/H8 protons, allowing the intra-nucleotide connection of ribose and base spin systems, was obtained from the 3D HCN spectra and confirmed from the 3D ^{13}C-edited NOESY spectrum. Sequential assignment was obtained from the 3D ^{13}C-edited NOESY and confirmed from the 3D HCP spectrum. Imino protons were assigned from 2D NOESY spectra with 50, 150 and 300 ms mixing times. To detect $^2J_{NN}$ couplings across hydrogen bonds in Watson–Crick (WC) and sheared G–A base pairs, 2D HNN-COSY spectra were recorded using correlations to both exchangeable and non-exchangeable protons (30,31). The completeness of the ^1H chemical shift assignments was 98% for non-exchangeable and 50% for exchangeable protons. ^1H chemical shifts were referenced to an external standard of 2,2-Dimethyl-2-silapentane-5-sulfonic acid (DSS), and ^{13}C and ^{15}N chemical shifts were referenced to the ^1H shift, as recommended (32).

Spectra were analysed with Felix (FELIX NMR, San Diego, CA, USA). The integration of the nuclear Overhauser enhancement (NOE) volumes and the calibration of the distances were performed by an internal routine of the program.

RDCs were calculated as the difference between ^1H–^{13}C couplings measured for isotropic and partially aligned samples. RDCs were obtained for the following inter-nuclear vectors: H8–C8 (Pu), H6–C6 (Py), H5–C5 (Py), H2–C2 (Ade) and H1′–C1′, H2′–C2′, H3′–C3′, H4′–C4′ (ribose). Seventy-nine RDCs could be measured (35 in the bases and 44 in the riboses).

Structure calculations

Structures were calculated using the Aria 1.2/CNS 1.1 set-up (33,34). Six hundred and twenty-three NOE distances were categorized as weak (1.8–5.0 Å), medium (1.8–3.4 Å) or strong (1.8–2.6 Å) (Table I). Eighty dihedral angles were experimentally restrained. The ribose conformation of 19 nucleotides was restrained to the *C3′-endo* range by estimating the magnitude of the $^3J_{H1'-H2'}$ scalar couplings in the HCCH-E.COSY 3D experiment. The same experiment was used to estimate the magnitude of $^3J_{H4'-H5'/H5''}$ scalar couplings, providing 16 restraints for γ dihedral angles to the *gauche$^+$* range. The phosphate dihedral angles α and ζ were restrained to the *trans* conformation for 15 nucleotides on the basis of their ^{31}P chemical shifts (35). The χ angles of 15 nucleotides were restricted on the basis of the intensities of the intranucleotide H8–H1′ (Pu) and H6/H5–H1′ (Py) NOEs. Loose non-experimental restraints were added for the β (180° ± 60°) and ε (−135° ± 60°) angles of nucleotides in helical region excluding A25, C28-ε, A29, A30, U31, G32-β, U31, G32-β, A44-ε and G45-β. The rationale behind these restraints is to loosely restrain the β and ε angles of helical regions to the value assumed in A-form helices, while

Table 1. Structure Statistics

NOE distance restraints	623
Intra-nucleotide	440
Inter-nucleotide	183
Inter-strands	27
Hydrogen bonds	86
Planarity restraints	12
Dihedral angles restraints	80
RDC restraints	79
Base	35
Sugar	44
RMSD (for heavy atoms)[a]	
Overall (Å)	0.51 (0.61)
C-stem (Å)	0.44 (0.50)
NC-stem (Å)	0.45 (0.44)
Number of NOE violations (>0.6 Å)	0
Number of dihedral violations (>3°)	0

[a]The RMSD values refer to the 10 lowest energy structures after RDC refinement and prior to water refinement. The values in parentheses refer to the 10 lowest energy structures after water refinement. C-stem and NC-stem RMSDs are calculated considering the residues ranges (A20–U24, G26–C28 and G45–U52) and (G32–G35 and C41–A44) for the C-stem and the NC-stem, respectively. The overall RMSD includes the NC- and C-stem residues with the addition of A29 and A30.

allowing the β and ε angles of nucleotides in non-helical regions (or at the end of helices) to assume any value in agreement with the experimental data.

Hydrogen bonds of WC base pairs were detected in HNN correlations and in NOESY experiments. An intense NOESY peak between the imino resonances of G48 and U24 confirmed the presence of the G48-U24 wobble base pair. The two non-canonical G–A base pairs in the NC-stem were detected through a G–A-selective HNN experiment, where magnetization is transferred from the N7 of the As to the N2 of the Gs through the hydrogen bond-mediated $^2J_{NN}$ scalar coupling. During the calculations, hydrogen bonds were maintained by distances restraints, while planarity was enforced through weak planarity restraints (5 kcal mol^{-1} Å^{-2}).

Hundred structures were calculated in one iteration without using the automated assignment or the distance calibration options of Aria 1.2. The simulated annealing (SA) protocol starts with a high-temperature torsion angle SA phase with 50 000 steps at 10 000 K (time step of 28 fs). This is followed by a torsion angle dynamic cooling phase from 10 000 K to 2000 K in 10 000 steps and by two cartesian dynamic cooling phases with a time step of 3 fs (from 2000 K to 1000 K in 50 000 steps and from 1000 K to 50 K in 20 000 steps, respectively).

In a second step, the structures were refined adding RDCs data to the structural restraints. The initial values for the rhombic (r) and axial (D_a) components of the alignment tensor were initially obtained by evaluating the RDCs pattern distribution (36). This method provided the values $D_a = 12.0$ Hz and $r = 0.2$. An intensive grid search was performed around these values for both D_a and r, where the dipolar coupling energy term was evaluated as a function of the alignment tensor. Examination of the dipolar coupling energy profiles revealed a minimum for $D_a = 14.0$ and $r = 0.2$. Thus, this tensor was employed in the refinement.

The final ensemble of the 10 structures was refined in a shell of water molecules (37–39).

The final RDC refined structures showed no NOE (>0.6 Å) or dihedral angle (>3°) violations. The final structures were analysed using MolMol (40). Figures were prepared with Pymol (http://www.pymol.org).

SANS data

The U4-Kt2 RNA sample used for the SANS experiments consisted of a purified duplex, 0.75 mM in concentration dissolved in the same buffer used for the NMR experiments (in the absence of Mg^{2+}) with 100% H$_2$O. The data were acquired on the small-angle diffractometer D22 (41) at the Institute Laue-Langevin (ILL; Grenoble, France). A volume of 200 µl was measured in a Hellma$^{®}$ Qs quartz cuvette with an optical path length of 1 mm. The incident neutron wavelength was $\lambda = 6$ Å and a detector/collimator set-up of 2 m × 2 m was chosen. The temperature of the sample was maintained at 298 K throughout the experiment. Exposure times were 120 and 60 min for the RNA and its buffer, respectively. In addition, the empty quartz cuvette, a H$_2$O sample and

boron were measured for the background subtraction and detector calibration. The beam centre was determined using a Teflon sample. The scattering intensities of both the buffer and the RNA-containing solution were corrected for electronic noise, detector efficiency and sample holder scattering and integrated azimuthally using an ILL in-house data reduction suite (42). Sample and buffer intensities were subtracted using the software PRIMUS (43), to correct for the scattering effect of the buffer and the empty cuvette, as well as for the electronic background at the detector.

The radius of gyration (R_g) was extracted from the scattering intensity data using the Guiner approximation (44):

$$I(Q) = I(0) \exp\left(-Q^2 R_g^2/3\right) \tag{1}$$

which is valid for $(QR_g) < 1.0\ldots1.3$, where $Q = 4\pi \sin\theta/\lambda$, and 2θ is the scattering angle. The experimental scattering curve and R_g were compared to those simulated for the lowest energy NMR structure and the protein-bound crystallographic structure (2OZB.pdb, in complex with 15.5K and hPrp31) (11) using the software CRYSON (45). The same program was used to verify the influence of the hydration shell properties on the SANS curve and the R_g by varying the density of the hydration shell from 100% to 130% of that of the bulk solvent. The construct used for this structural study (U4-Kt2) differs from the one used in the crystallographic study (U4-Kt1) at the apical pentaloop, which is substituted by a 3-nt open-loop construct in U4-Kt2 (Figure 1), and in the length of the C-stem. In order to compare the scattering curve predicted from the RNA protein-bound crystallographic structure to the experimental data, two nucleotides of the apical pentaloop were removed from the RNA of 2OZB.pdb and a base pair was added at the end of the C-stem with INSIGHT (Accelerys Software Inc., San Diego, CA, USA).

SANS data were used to perform *ab initio* modelling using DAMMIN (46). Several distance distribution functions $p(r)$ were calculated using GNOM (47), with different values of D_{max} (longest distance) ranging from 50 Å to 65 Å. The output of GNOM was used as input to perform a first set of shape modelling runs (5 for each D_{max}) with the software DAMMIN (46), which uses a SA procedure to calculate a 'best-fit' structural model consisting of dummy atoms. The best value of D_{max} (50 Å) was selected by comparing the scattering curve, the R_g and the solvent-excluded volume predicted for each model with the experimental data, as well as by visual inspection of the $p(r)$ function (smoothness and gentle drop to 0 as $r \rightarrow D_{max}$). The chosen data set was used for further modelling with DAMMIN in 20 different runs. The resulting 20 envelopes were aligned, averaged and filtered using the DAMAVER package (48). The R_g of the final model, its back-calculated scattering curve, as well as its solvent-excluded volume were in excellent agreement with the values extracted from the experimental data. The final model was aligned to the NMR lowest energy structure as well as to the X-ray structure using the program SUPCOMB20 (49).

The U4-Kt molecular weight in solution, $M_r(exp)$ was determined using absolute calibration against water (50):

$$M_r(exp) = \frac{1-T}{f4\pi T_s} \frac{I(0)}{I_{inc}(0)} \frac{10^3}{CN_A t} \left[\frac{(\sum b_i - \rho_s V)}{M_r} \right]^2 \quad (2)$$

$I(0)$ (= 0.055) and $I_{inc}(0)$ (= 1.0) are the scattered intensities in the forward direction of RNA and pure water (H_2O), respectively, T_s (= 0.49) and T (= 0.49) are the transmissions of the sample and H_2O, respectively, C (= 8.1 mg/ml) is the RNA concentration, t (= 0.1 cm) is the thickness of the quartz cuvette, f is a correction factor for the anisotropicity of the solvent scattering as a function of neutron wavelength (equal to 0.8 for $\lambda = 6 \text{ Å}$) (50), b_i are the scattering lengths of the RNA atoms ($\Sigma b_i = 350 \times 10^{-12}$ cm), ρ_s (-0.562×10^{10} cm^{-2}) is the solvent scattering density, V (equal to 10100 Å, from the U4-Kt sequence using Vosh and Gerstein, 2006) is the solvent-excluded volume of the RNA, M_r is the theoretical molecular weight of U4-Kt2 (10.8 kDa) and N_A is Avogadro's constant.

RESULTS AND DISCUSSION

NMR analysis

The K-turn RNA U4-Kt2 (Figure 1) was subjected to structural analysis by NMR. We chose this construct instead of the U4-Kt1, as the latter gives poor yields in *in vitro* transcription. The U4-Kt2 construct is competent to bind 15.5K (data not shown). Most of the NMR analysis has been conducted on two samples of the U4-Kt2 duplex, in each of which only one strand is $^{13}C/^{15}N$-labelled. This greatly simplifies the analysis of the NMR data by substantially reducing spectral overlap. The unlabelled strand was given in slight excess (1.5:1) to ensure complete saturation of the labelled strand. To verify that the excess of the unlabelled strand does not induce undesired conformational properties, a third sample, consisting of both $^{13}C/^{15}N$-labelled strands in complex with each other, was purified by size-exclusion chromatography prior to NMR analysis. All the spectral properties of this sample, containing the purified duplex and no excess of single strands, were equivalent to those of the two samples containing the unlabelled strand in excess. This third sample was also used to perform cross-strand NMR experiments, necessary to verify the presence of hydrogen bonding in the two stems.

The C–H correlation spectra of both strands are shown in Figure 2A for the C6/C8–H6/H8 region and in Supplementary Figure S2 for the C5–H5 region. Both the presence of one single set of resonances and the sharp line width of the resonances of both the stems and the internal loop exclude that the U4-Kt RNA exists in solution in two substantially populated conformations that are slowly inter-converting.

Intrigued by the literature reports proposing a massive unfolding of the NC-stem in the absence of protein binders (15,16), we set out to verify whether the base pairs of the NC-stems are present in the U4-Kt2 RNA in solution. The imino resonances of all Gs can be seen and assigned in 1D (Figure 2B) and 2D NOESY experiments including those of G34 and G35 that are involved in the two WC base pairs of the NC-stem. The non-selective HNN correlation (Supplementary Figure S3) reveals that the two imino groups of G34 and G35 are indeed involved in H-bonds with the N3 of Cs, confirming that the two canonical base pairs of the NC-stem are present in solution. These resonances are weaker than the imino resonances of WC C–G base pairs in the C-stem, as expected for terminal base pairs or for base pairs belonging to very short helical stems, which are more easily accessible to the solvent. However, a rough estimation of the H-bond-mediated $^2J_{NN}$ coupling (51), obtained by quantification of the cross-peaks in the HNN experiment, indicates scalar couplings of similar size for both the C- and NC-stems, which confirms that the NC-stem is stably formed in solution. The presence of the two sheared G–A base pairs was verified in a second HNN experiment, where magnetization is transferred in a selective manner from the H8 to the N7 of As and subsequently from the N7 of As to the N2 of Gs through the H-bond-mediated $^2J_{NN}$ coupling constant. The presence of two peaks correlating the H8s of A33 and A44 with the N2 of Gs confirms that even the G–A base pairs of the NC-stem are formed in the U4-Kt RNA in the absence of protein binders, in contrast to what was proposed on the basis of MD simulations (15,16). Interestingly, the imino protons of G33 and G43 are visible as well (Figure 2 and Supplementary Figure S6), although they are not involved in hydrogen bonding. This fact points to a reduced solvent accessibility of these sites, as it has been observed previously for G–A base pairs (52–55). The chemical shifts of these imino protons are in the expected range for imino protons of G–A base pairs (up to 12.2 ppm) (52–55).

MD simulations performed on the same U4 K-turn RNA studied in this work suggested that the NC-stem unfolds in the absence of protein binders as a consequence of the opening of both the non-canonical G–A and WC G–C base pairs (16). The unfolding of the NC-stem had been proposed to explain the higher accessibility to Kethoxal modification of G32, G34 and G35 with respect to the Gs of the C-stem and in the absence of 15.5K. In contrast, our NMR data clearly show that both the G–C and G–A base pairs are present in the U4-Kt RNA in solution. Indeed, the higher accessibility of G32, G34 and G35 to Kethoxal modifications could be explained without invoking a complete unfolding of the NC-stem. G32 and G35 both belong to terminal base pairs of the NC-stem and thus their N1 and N2 sites are easily accessible in the free RNA. G34 is very poorly modified with respect to G32 and G35. Its moderate accessibility to Kethoxal is probably a consequence of the faster opening rate of the G34–C42 base pair in the short NC-stem. However, it should be noticed that, while the G–C base pairs of the NC-stem seem to be kinetically less stable than those in the C-stem, as indicated by the lower intensity of their resonances in the HNN spectrum (Supplementary Figure S3), thermodynamically they are as stable as all other G–C base pairs, as indicated by similar values of the hydrogen bond-mediated $^2J_{NN}$ coupling constants.

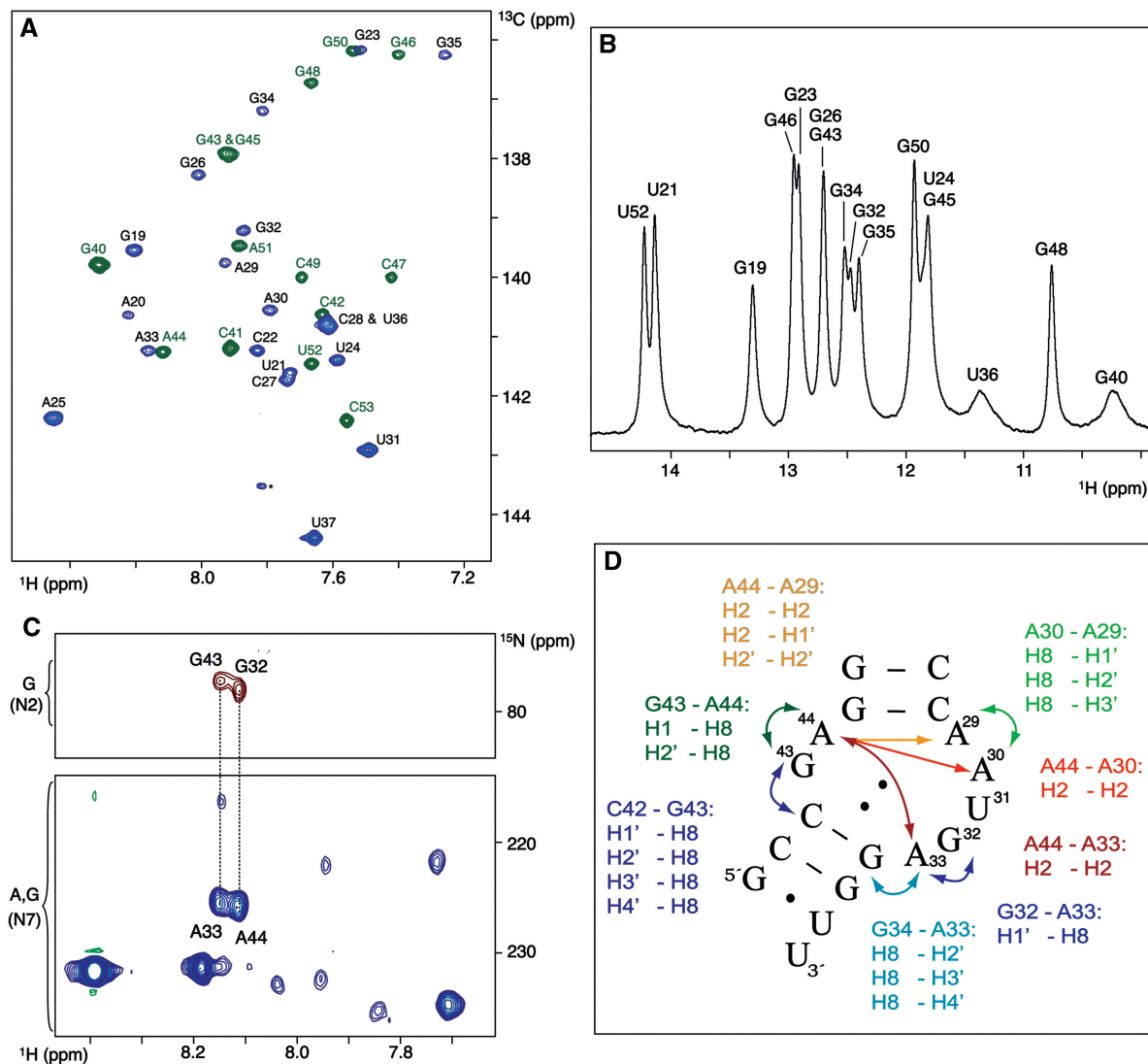**Figure 2.** (**A**) C–H correlation showing the base region of both the short (I1, green) and the long (I2, blue) strands of the U4-Kt2 RNA see Figure 1 for the definition of I1 and I2. (**B**) Imino spectrum of the U4-Kt2 RNA with assignment. G40 and U36 resonances are visible only at 5°C and are assigned by exclusion to the terminal G–U base pair of the NC-stem. (**C**) Selective HNN correlation designed to transfer the magnetization from the H8 of the As to the N7 of the As through the $^2J_{NH}$ scalar couplings and subsequently from the N7 of the As to the N2 of the Gs through the hydrogen-bond mediated $^2J_{NN}$ coupling. The presence of two peaks correlating the H8 protons of A33 and A44 with the N2 of two Gs confirms the presence of the G–A non-canonical base pairs. (**D**) Network of inter-nucleotide NOEs observed at the internal loop site. The presence of several NOEs confirms that the internal loop is well structured in solution even in the absence of proteins.

Binding of 15.5K results in the protection of all three Gs from Kethoxal modification. For G32, this can be attributed to the direct contacts of the side chain of E41 with the N1 and N2 sites of G32; for G34 and G35, it can be attributed to the extensive contacts of the side chains of K44 and R48 with the phosphate backbone of C42 and C41. These contacts slow down the opening process of the G–C base pairs of the NC-stem by fixing the helical backbone on the site of the Cs, thus making Kethoxal modification inefficient.

Analysis of the NOE spectra revealed a tight network of NOEs in the internal loop region (Figure 2D) including both cross-strand and (*i*, *i*+1) sequential NOEs. This is a strong indication that the internal loop assumes a preferred, well defined, major conformation even in the absence of proteins. However, the riboses of nucleotides around the bulged-out A25 of the C-stem and around the U31 of the internal loop (U24, A25, G26, A30, U31 and G32) show averaged values of the $^3J_{H1'H2'}$, which is indicative of riboses inter-converting between the *C3'-endo* and the *C2'-endo* conformations. In addition, the $R_{1\rho}/R_1$ ratios of relaxation rates for the C1' nuclei of U24, A25, G26, A30, U31 and G32 (but not for the C1' nuclei of A29 and A33) deviate from those of helical regions, indicating local plasticity of these riboses (data not shown).

From all our NMR data, it can be concluded that both the internal loop and the NC-stem of the U4-Kt RNA are well structured in solution. Next, we set out to verify whether the tight structural organization of the U4-Kt RNA in solution reflects the typical, sharply bent conformation observed in complex with proteins.

Structural analysis

Structural restraints included 623 NOEs (440 intra-residue and 183 inter-residue NOEs), 80 experimentally determined dihedral angle restraints (19 restraints for δ, 16 for γ, 15 for α and ζ and 15 for χ) and 79 RDCs. In addition, H-bond restraints were imposed for all base pairs in the C- and NC-stems, for which the base pairing was verified in HNN experiments (for the G–U base pair of the C-stem, an intense NOE peak was observed between the imino protons of G48 and U24 in a NOESY spectrum with 50 ms mixing time). Details about the structure calculation protocols are given in the 'Materials and Methods' section.

The final ensemble of the 10 lowest energy structures prior to RDC refinement shows an excellent convergence for the NC- and C-stems separately [heavy atom root mean square deviation (RMSD): 0.31 and 0.61 Å, respectively]. On the other hand, the relative position of the two stems is less well defined (overall RMSD 0.98 Å) as a result of the lack of long-range structural information in NOEs and dihedral angle restraints. This long-range structural information is provided by the RDCs. After refinement including RDC restraints, followed by minimization in water, the 10 lowest energy structures converged to a total pairwise RMSD of 0.61 Å, while the RMSD for the NC- and C-stems are 0.44 and 0.50 Å, respectively (Table 1 and Figure 3). The structure of the U4-Kt RNA is very well defined both in the conformation of the two stems and in their relative orientation, as indicated by the excellent fit of the experimental RDCs to the back-calculated ones ($Q = 0.17$) (56). The 10 lowest energy structures have been deposited in the Protein Data Bank (PDB) under accession code 2KR8 together with the NMR restraints used for structure calculations.

The conformation of the U4-Kt RNA in solution is much more extended than that observed in protein-bound

K-turn motifs. In Cojocaru *et al.* (15), an angle φ is defined between the phosphates of C47, U31 and G35 to describe the kink in the RNA backbone. This angle assumes a value of 25° in the protein-bound crystallographic conformation of K-turn RNAs, while it becomes larger for more extended conformations. In our case, $\varphi = 69°$, which indicates that the RNA is not sharply kinked in solution in the absence of cognate proteins and divalent cations (Figure 4A). As a consequence of this, no A-minor interaction is observed between A33 of the NC-stem and G45 of the C-stem, as confirmed by the absence of NOEs between the ribose of G45 and either the ribose or the aromatic protons of A33 (Supplementary Figure S4).

A detailed analysis of the region between nucleotides 28–33 and 43–45 reveals that the K-turn consensus sequence is indeed well folded in the unbound RNA, but that the pattern of inter-nucleotide interactions substantially differs from that in complex with proteins. In our structure, A30 stacks on A29, which is situated on the top of C28 of the C-stem at an angle of ∼60° (Figure 4B). The stacking of A30 onto A29 is indicated by a series of sequential NOEs connecting the H8 of A30 with the ribose protons of A29, which would not be expected from the protein-bound K-turn structure (Figure 2D). In fact, in the sharply kinked structure, A30 does not stack on A29 but rather on the NC-stem (Figure 4B). The large relocation of A30 from the top of the NC-stem in the protein-bound structure to the top of A29, and thus of the C-stem, in the free RNA solution structure is accompanied by a change in the conformation of the ribose of A30 from *C2′-endo* to *C3′-endo* and of the χ dihedral angle form *syn* to *anti*. The calculated structural ensemble shows a very well-converged *C3′-endo* conformation for the ribose of A30, which was left unrestrained during the calculations; however, the value of 5.2 Hz measured for the $^3J_{H1'H2'}$ scalar coupling of A30 is indicative of a ribose interchanging between the *C3′-endo* and *C2′-endo* conformations, as expected for non-A-form helical regions and terminal base pairs of helices. The pyrimidine base of the internal loop, U31, also roughly stacks on A30 but is less well ordered. In agreement with this, relaxation rates suggest high fast dynamics for the base of U31, but not for those of A29 and A30 (data not shown).

SANS analysis

The NMR structural analysis indicates the presence of an extended conformation for the unbound U4-Kt2 RNA in solution and in the absence of divalent cations. To obtain a second independent measure of the overall shape of this RNA, we performed a SANS analysis. The scattering curve strictly depends on the spatial coordinates of the scattering nuclei and therefore on the shape of the molecule. The experimental scattering curve (Figure 5A) was compared to the scattering curves calculated from the atomic coordinates of the lowest energy NMR structure of the free U4-Kt RNA and for the crystallographic structure of the protein-bound U4-Kt RNA of 2OZB.pdb, as explained in the 'Material and Methods' section. The quality of the fitting is excellent for the NMR structure of the free RNA ($\chi^2 = 0.81$), while it is not acceptable for the sharply
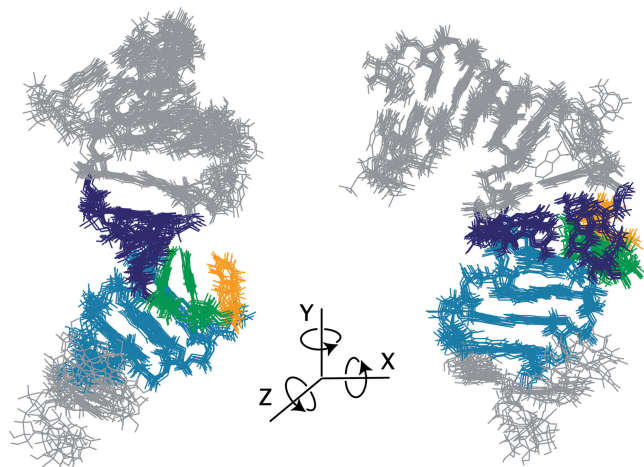

Figure 3. Superposition of the 10 lowest energy NMR structures obtained after refinement with RDCs and in water (PDB code 2KR8). The two different views are obtained by rotation of 204°, 220° and 198° around the X, Y and Z axes, respectively. The convergence of the structure is excellent with an overall heavy atom RMSD of 0.61 Å. The colour coding of the nucleotides is as in Figure 1.
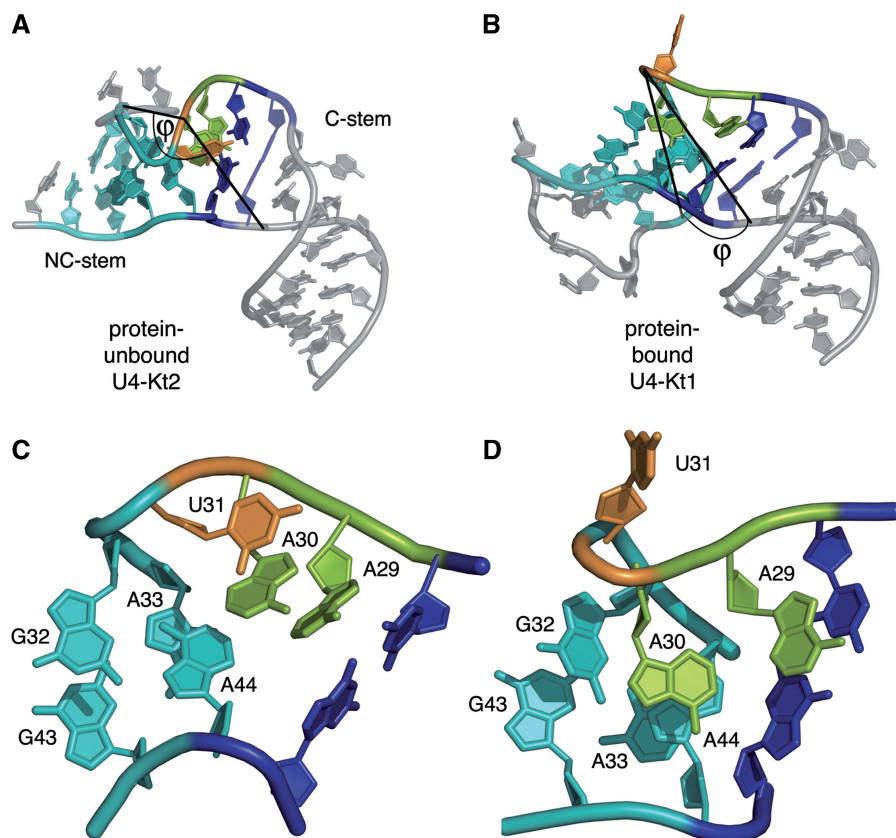
Figure 4. Comparison of the lowest energy NMR structure of the unbound U4-Kt RNA (**A**, PDB code 2KR8) and the crystallographic protein-bound structure (**B**, PDB code 2OZB). The angle φ between the P nuclei of C47, U31 and G35 describes the bend between the two stems. The NMR structure of the unbound RNA is quite extended ($\varphi = 69°$) while the protein-bound structure is sharply kinked ($\varphi = 25°$). The close-up of the internal loop region (**C** and **D**) shows the difference in the location of the unpaired A29 and A30 in the two structures. A30 is stacked on the C-stem with A29 in the unbound RNA structure (C), while it stacks on the NC-stem in the protein-bound RNA structure (D).

bent protein-bound structure ($\chi^2 = 1.14$) (Figure 5A). The R_g of 16.6 ± 0.7 Å was determined by the Guinier approximation from the experimental scattering curve (44); this value is in excellent agreement with the value of 16.2 Å predicted for the NMR structure of the unbound RNA, while the value of 15.0 Å expected from the sharply bent protein-bound structure is significantly smaller than the experimental value. The R_g's calculated from the PDB structures do not depend significantly on the properties of the hydration shell for SANS in H_2O due to the very small absolute scattering length density of water (57,58). In our case, the accuracy of the values is better than 0.1 Å, as verified by varying the density of the hydration water shell from that of bulk water to 30% denser than bulk water using the program CRYSON (45).

The R_g is a reliable quantitative structural parameter that can be extracted from small-angle scattering curves without *a priori* structural information. In addition, SANS data can be used to compute *ab initio* low-resolution models of the overall shape of the scattering molecule, if the latter can be assumed to be monodisperse in solution (46). Here, we use the SANS data to calculate *de novo* the shape of the unbound U4-Kt RNA, following the procedure explained in the 'Materials and Methods' section, and we compare the resulting model with the NMR structure

of the free RNA, as well as with the crystallographic protein-bound RNA structure using the program SUPCOMB (49). As shown in Figure 5B and C, the low-resolution shape of the U4-Kt2 RNA, calculated *ab initio* from the SANS scattering curve, can be well superimposed with the NMR structure of the free U4-Kt RNA, while the sharply bent protein-bound U4 structure does not match the SANS model.

This visual similarity is quantitatively confirmed by the normalized spatial discrepancy (NSD) values (49) that were 0.85 and 0.92 for the NMR and crystallography models, respectively (NSD values <1.0 indicate a similarity between structures with smaller values corresponding to a better fit). As a quality control of the average low-resolution model, we checked its solvent-excluded volume ($V = 11\,500$ Å^3) that is in fair agreement with the theoretically predicted one (10 100 Å^3) calculated from its nucleic acid sequence (59). The quality of the SANS data and the monodispersity of the solution were furthermore validated by the experimental determination of the U4-Kt molecular weight using absolute calibration against water (see 'Material and Methods' section for details) (50). The value found (10.7 kDa) is in excellent agreement with the theoretically predicted one (10.8 kDa).
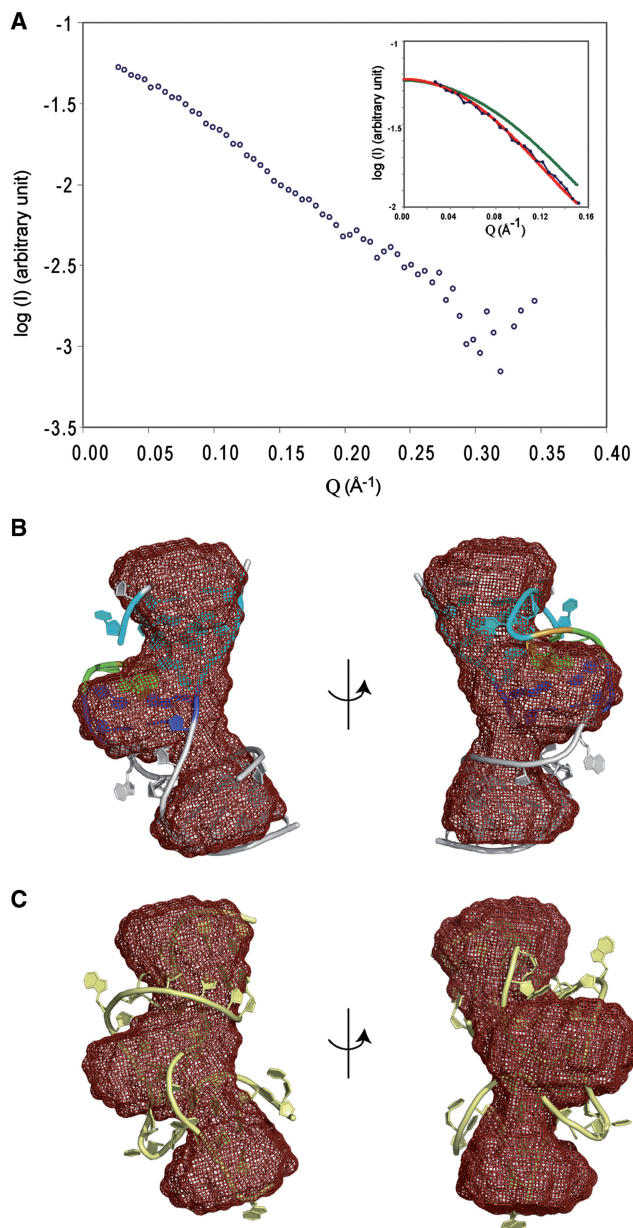
Figure 5. Analysis of SANS data of the unbound U4-Kt RNA in solution. (**A**) The scattering profiles simulated from the NMR structure of the U4-Kt RNA free in solution (PDB code 2KR8, red) and from the X-ray crystallography protein-bound structure (PDB code 2OZB, green) are compared to the experimental values (blue). (**B**, **C**) The unbound NMR structure (B) and the X-ray protein-bound structure (C) of the U4-Kt RNA are superimposed to the *ab initio* low-resolution model obtained from SANS data (46), by using SUPCOMB20 (49).

In conclusion, the SANS data completely support the NMR analysis and independently corroborate that the sharply bent K-turn motif is not formed in solution in the absence of proteins and divalent cations.

Effect of divalent cations on the U4-Kt RNA conformation

To test the effect of Mg^{2+} on the conformation of the K-turn region of the U4-Kt2 RNA in solution, we performed titrations with $MgCl_2$ (Supplementary Figure S5). Four points were recorded for the $^{13}C/^{15}N$-labelled I2 strand paired with unlabelled I1 strand (see Figure 1 for the definition of I1 and I2). As expected, the addition of Mg^{2+} has dramatic effects on both the position and line width of resonances of A29, A30, U31, G32 and A33 and to a minor extent of G34 and G35. The C8–H8 resonance of A29 shifts and broadens progressively upon addition of Mg^{2+}, while the C8/C6–H8/H6 resonances of A30, U31, G32 and A33 disappear already at $[Mg^{2+}] = 2\,mM$. This is indicative of strong line broadening, due to an exchange process in the high µs to ms time-scale involving the internal loop and, to a minor extent, the NC-stem. This observation is in agreement with the notion that the divalent cations stabilize a second conformation of the U4-Kt RNA, which, in the absence of Mg^{2+}, is either absent or very poorly populated. Although our NMR data do not contain information on the nature of this conformation, it is tempting to speculate that this is the kinked conformation. Unfortunately, the broadening of the NMR lines beyond detection does not permit a detailed structural investigation of the U4-Kt RNA in the presence of Mg^{2+} ions.

The influence of Mg^{2+} ions on the population of the bent conformation had been investigated previously by fluorescence resonance energy transfer (FRET) experiments (14). In this study, the population distribution of 65:35 between the extended and the bent conformations in the absence of divalent cations had been found to revert to 30:70 in the presence of Mg^{2+}. Our NMR data are in very good qualitative agreement with this FRET analysis. In a cellular environment, Mg^{2+} ions might partially pre-fold the K-turn RNAs prior to protein binding by stabilizing the kinked conformation.

Protein-assisted RNA folding

The structure of K-turn RNAs has long been object of interest, probably due to the rather unusual fold of these RNAs in complex with proteins. Both theoretical and experimental studies have addressed the question of the conformation of K-turn RNAs in the absence of proteins (14,15,18,60), using tools such as MD simulations, biochemical probing and FRET. The results of all these studies can be summarized as follows: K-turn RNAs exist in equilibrium between the sharply kinked conformation, where the two stems are at an angle of 60° [φ angle (15) = 25°], and a more extended one. The population distribution depends on the presence of divalent cations.

Both the NMR and the SANS data, collected for the U4-Kt2 RNA in the absence of divalent cations, show that the U4-Kt RNA exists in solution prevalently in the extended conformation of Figure 4A. In agreement with previous literature reports, our findings disprove the hypothesis that the K-turn sharply bent structure is a primary organizational element determined by the consensus sequence of Figure 1 and favours the thesis of protein-assisted RNA folding, as readily proposed by Turner *et al.* (61). In the U4 K-turn RNA studied here, the sharp bend in the RNA backbone is highly disfavoured in solution, probably due to the high density of

negative phosphate backbone charges at the sharp kink, which needs to be stabilized by protein binding.

In addition, it is worth noticing that the population distribution observed for the U4-Kt RNA might not be a universal feature of all K-turn RNAs, but might depend on the RNA sequence, as readily suggested by Matsumura *et al.* (60). In fact, a previous NMR study on the helix–loop–helix portion of the L30 mRNA in the absence of Mg^{2+} ions reports on a flexible internal loop (K-turn) with resonances broadened beyond detection and very few NOEs (62). This might indicate a higher level of population of the kinked conformation for the K-turn region of the L30 mRNA as compared to the U4-Kt RNA, which instead predominantly assumes the more extended conformation of Figure 4A.

A FRET study of the ribosomal Kt-7 reports a ratio between the populations of the extended and kinked conformations close to 65:35 in the absence of divalent cations (14) and suggests that protein binders are needed to fully stabilize the K-turn structure. Our results are in agreement with this notion, even though, in the absence of divalent cations, the NMR data indicate a more skewed population distribution in favour of the more extended conformation. This discrepancy further supports the hypothesis that the populations of the extended and the bent conformation depend on the exact RNA sequence. Another possible explanation for the difference in the population distribution observed in ours and in the FRET study is the much higher length of the RNA molecules used in the FRET study with respect to the short U4-Kt2 sequence. In large RNAs, stabilization of the bent conformation may be induced by long-range tertiary interactions. Indeed, the sharply bent conformation of the K-turn motif is observed also in the absence of direct protein binders for one of the six K-turns found in the 23S rRNA of *Haloarcula marismortui* and for the S-adenosylmethionine riboswitch (63), suggesting that the long-range tertiary interactions may (3) stabilize the kinked conformation in a similar way as protein binding.

Interdisciplinarity of SANS and NMR data

The combined use of SANS [or small-angle X-ray scattering (SAXS)] and NMR data is gaining momentum in the study of multidomain proteins and protein complexes (64,65), where SANS/SAXS data provide an NMR-independent measurement of the relative position of protein domains or of globular proteins in a complex. SANS/SAXS data are recorded in very similar experimental conditions as that of solution-state NMR spectra and represent a precious addition to the long-range structural information obtained by NMR through the analysis of RDCs. In the structural study of RNA molecules, which inherently consist of several helical domains and whose structure determination by NMR suffers from the paucity of long-range NOE restraints, the combination of SANS/SAXS data with NMR data is likely to be even more relevant than for proteins. A recently published pilot study in this direction (66) and the work presented here demonstrate that small-angle scattering data together with NMR RDC data allow determination of the relative orientation of RNA helices in complex RNA molecules, therefore facilitating the difficult structural investigation of isolated RNAs.

CONCLUSIONS

In this work, we have investigated the conformation of the U4-Kt RNA in solution in the absence of protein binders and divalent cations. Our combined NMR and SANS study reveals that the U4-Kt RNA prefers an extended conformation in solution and that the sharply bent conformation observed in complex with proteins is either not present or very poorly populated. We verified that addition of Mg^{2+} ions triggers a conformational exchange in the internal loop region in the μs to ms time-scale, which is in agreement with the notion that divalent cations induce formation of the kinked conformation. All in all, our findings indicate that the consensus sequence of Figure 1A does not code for the sharply kinked structures observed for K-turn RNAs in complex with proteins.

Recently, it has been proposed that fast inter-helical sub-domain motions of RNA molecules consisting of multiple helical segments allow sampling of the conformational space required for binding to cofactors (67). This principle has been demonstrated for the HIV-1 TAR RNA in a very elegant work using decoupling of inter-domain motions from the overall re-orientation of the molecule by elongation of one helical segment. As a next step after the determination of the high-resolution structure of the naked U4-Kt RNA, it would be interesting to study whether the principle demonstrated for the TAR RNA also applies to kink-turned RNAs. We are currently embarking in this insightful but demanding analysis of the fast inter-domain dynamics of the U4-Kt RNA.

ACCESSION NUMBER

The atomic coordinates of the 10 lowest energy structures of the U4-Kt RNA have been deposited in the PDB (http://www.rcsb.org) under accession code 2KR8.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

REFERENCES

1. Wimberly,B.T., Brodersen,D.E., Clemons,W.M., Morgan-Warren,R.J., Carter,A.P., Vonrhein,C., Hartsch,T. and Ramakrishnan,V. (2000) Structure of the 30S ribosomal subunit. *Nature*, **407**, 327–339.
2. Schluenzen,F., Tocilj,A., Zarivach,R., Harms,J., Gluehmann,M., Janell,D., Bashan,A., Bartels,H., Agmon,I., Franceschi,F. *et al.* (2000) Structure of functionally activated small ribosomal subunit at 3.3 angstrom resolution. *Cell*, **102**, 615–623.
3. Ban,N., Nissen,P., Hansen,J., Moore,P.B. and Steitz,T.A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 angstrom resolution. *Science*, **289**, 905–920.
4. Toor,N., Keating,K.S., Taylor,S.D. and Pyle,A.M. (2008) Crystal structure of a self-spliced group II intron. *Science*, **320**, 77–82.
5. Moore,P.B. (1999) Structural motifs in RNA. *Annu. Rev. Biochem.*, **68**, 287–300.
6. Klein,D.J., Schmeing,T.M., Moore,P.B. and Steitz,T.A. (2001) The kink-turn: a new RNA secondary structure motif. *EMBO J.*, **20**, 4214–4221.
7. Vidovic,I., Nottrott,S., Hartmuth,K., Luhrmann,R. and Ficner,R. (2000) Crystal structure of the spliceosomal 15.5kD protein bound to a U4 snRNA fragment. *Mol. Cell*, **6**, 1331–1342.
8. Mao,H.Y., White,S.A. and Williamson,J.R. (1999) A novel loop-loop recognition motif in the yeast ribosomal protein L30 autoregulatory RNA complex. *Nat. Struct. Biol.*, **6**, 1139–1147.
9. White,S.A., Hoeger,M., Schweppe,J.J., Shillingford,A., Shipilov,V. and Zarutskie,J. (2004) Internal loop mutations in the ribosomal protein L30 binding site of the yeast L30 RNA transcript. *RNA*, **10**, 369–377.
10. Watkins,N.J., Segault,V., Charpentier,B., Nottrott,S., Fabrizio,P., Bachi,A., Wilm,M., Rosbash,M., Branlant,C. and Luhrmann,R. (2000) A common core RNP structure shared between the small nucleor box C/D RNPs and the spliceosomal U4 snRNP. *Cell*, **103**, 457–466.
11. Liu,S., Li,P., Dybkov,O., Nottrott,S., Hartmuth,K., Luhrmann,R., Carlomagno,T. and Wahl,M.C. (2007) Binding of the human Prp31 nop domain to a composite RNA-protein platform in U4 snRNP. *Science*, **316**, 115–120.
12. Lescoute,A., Leontis,N.B., Massire,C. and Westhof,E. (2005) Recurrent structural RNA motifs, isostericity matrices and sequence alignments. *Nucleic Acids Res.*, **33**, 2395–2409.
13. Nissen,P., Ippolito,J.A., Ban,N., Moore,P.B. and Steitz,T.A. (2001) RNA tertiary interactions in the large ribosomal subunit: the A-minor motif. *Proc. Natl Acad. Sci. USA*, **98**, 4899–4903.
14. Goody,T.A., Melcher,S.E., Norman,D.G. and Lilley,D.M.J. (2004) The kink-turn motif in RNA is dimorphic, and metal ion-dependent. *RNA*, **10**, 254–264.
15. Cojocaru,V., Nottrott,S., Klement,R. and Jovin,T.M. (2005) The snRNP 15.5K protein folds its cognate K-turn RNA: a combined theoretical and biochemical study. *RNA*, **11**, 197–209.
16. Cojocaru,V., Klement,R. and Jovin,T.M. (2005) Loss of G-A base pairs is insufficient for achieving a large opening of U4 snRNA K-turn motif. *Nucleic Acids Res.*, **33**, 3435–3446.
17. Nottrott,S., Hartmuth,K., Fabrizio,P., Urlaub,H., Vidovic,I., Ficner,R. and Luhrmann,R. (1999) Functional interaction of a novel 15.5kD [U4/U6 center dot U5] tri-snRNP protein with the 5′ stem-loop of U4 snRNA. *EMBO J.*, **18**, 6119–6133.
18. Razga,F., Koca,J., Sponer,J. and Leontis,N.B. (2005) Hinge-like motions in RNA kink-turns: the role of the second A-minor motif and nominally unpaired bases. *Biophys. J.*, **5**, 3466–3485.
19. Razga,F., Zacharias,M., Reblova,K., Koca,J. and Sponer,J. (2006) RNA kink-turns as molecular elbows: hydration, cation binding, and large-scale dynamics. *Structure*, **14**, 825–835.
20. Bringmann,P., Appel,B., Rinke,J., Reuter,R., Theissen,H. and Luhrmann,R. (1984) Evidence for the existence of snRNAs U4 and U6 in a single ribonucleoprotein complex and for their association by intermolecular base-pairing. *EMBO J.*, **3**, 1357–1363.
21. Hashimoto,C. and Steitz,J.A. (1984) U4 and U6 RNAs coexist in a single small nuclear ribonucleoprotein particle. *Nucleic Acids Res.*, **12**, 3283–3293.
22. Milligan,J.F., Groebe,D.R., Witherell,G.W. and Uhlenbeck,O.C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res*, **15**, 8783–8798.
23. Fiala,R., Czernek,J. and Sklenar,V. (2000) Transverse relaxation optimized triple-resonance NMR experiments for nucleic acids. *J. Biomol. NMR*, **16**, 291–302.
24. Brutscher,B. and Simorre,J.P. (2001) Transverse relaxation optimized HCN experiment for nucleic acids: combining the advantages of TROSY and MQ spin evolution. *J. Biomol. NMR*, **21**, 367–372.
25. Sklenar,V., Dieckmann,T., Butcher,S.E. and Feigon,J. (1998) Optimization of triple-resonance HCN experiments for application to larger RNA oligonucleotides. *J. Magn. Reson.*, **130**, 119–124.
26. Hu,W.D., Kakalis,L.T., Jiang,L.C., Jiang,F., Ye,X.M. and Majumdar,A. (1998) 3D HCCH-COSY-TOCSY experiment for the assignment of ribose and amino acid side chains in C-13 labeled RNA and protein. *J. Biomol. NMR*, **12**, 559–564.
27. Marino,J.P., Schwalbe,H., Anklin,C., Bermel,W., Crothers,D.M. and Griesinger,C. (1994) A 3-dimensional triple-resonance H-1,C-13,P-31 experiment - sequential through-bond correlation of ribose protons and intervening phosphorus along the RNA oligonucleotide backbone. *J. Am. Chem. Soc.*, **116**, 6472–6473.
28. Schwalbe,H., Marino,J.P., King,G.C., Wechselberger,R., Bermel,W. and Griesinger,C. (1994) Determination of a complete set of coupling-constants in C-13-labeled oligonucleotides. *J. Biomol. NMR*, **4**, 631–644.
29. 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 lambda N-peptide/boxB RNA complex. *J. Am. Chem. Soc.*, **119**, 6711–6721.
30. Dingley,A.J. and Grzesiek,S. (1998) Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide $^2J_{NN}$ couplings. *J. Am. Chem. Soc.*, **120**, 8293–8297.
31. Hennig,M. and Williamson,J.R. (2000) Detection of N-H…N hydrogen bonding in RNA via scalar couplings in the absence of observable imino proton resonances. *Nucleic Acids Res.*, **28**, 1585–1593.
32. Markley,J.L., Bax,A., Arata,Y., Hilbers,C.W., Kaptein,R., Sykes,B.D., Wright,P.E. and Wuthrich,K. (1998) Recommendations for the presentation of NMR structures of proteins and nucleic acids (Reprinted from *Pure and Applied Chemistry* (1998), **70**, 117–142). *J. Mol. Biol.*, **280**, 933–952.
33. Brunger,A.T., Adams,P.D., Clore,G.M., DeLano,W.L., Gros,P., Grosse-Kunstleve,R.W., Jiang,J.S., Kuszewski,J., Nilges,M., Pannu,N.S. *et al.* (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.*, **54**, 905–921.
34. Linge,J.P., Habeck,M., Rieping,W. and Nilges,M. (2003) ARIA: automated NOE assignment and NMR structure calculation. *Bioinformatics*, **19**, 315–316.
35. Varani,G., Cheong,C. and Tinoco,I. Jr (1991) Structure of an unusually stable RNA hairpin. *Biochemistry*, **30**, 3280–3289.
36. Clore,G.M., Gronenborn,A.M. and Tjandra,N. (1998) Direct structure refinement against residual dipolar couplings in the presence of rhombicity of unknown magnitude. *J. Magn. Reson.*, **131**, 159–162.
37. Linge,J.P., Williams,M.A., Spronk,C., Bonvin,A. and Nilges,M. (2003) Refinement of protein structures in explicit solvent. *Proteins*, **50**, 496–506.
38. Nozinovic,S., Fürtig,M., Jonker,H.R., Richter,C. and Schwalbe,H. (2010) High-resolution NMR structure of an RNA model system: the 14-mer cUUCGg tetraloop hairpin RNA. *Nucleic Acids Res.*, **38**, 683–694.
39. Nilges,M., Bernard,A., Bardiaux,B., Malliavin,T., Habeck,M. and Rieping,W. (2008) Accurate NMR structures through minimization of an extended hybrid energy. *Structure*, **16**, 1305–1312.
40. Koradi,R., Billeter,M. and Wuthrich,K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.*, **14**, 51–55.
41. May,R.P. and Thomas,M. (1986) D22 A- a new low-angle scattering instrument at the 2nd cold source of the ILL. B- Proposal for a new S.A.S. instrument at the second cold

source. *ILL Technical Report 86MA07T*. Institut Laue-Langevin, Grenoble, France.

42. Gosh,R.E., Egelhaaf,S.U. and Rennie,A.R. (2006) A computing guide for small-angle scattering experiments. *ILL Technical Report ILL06GH05T*. Institut Laue-Langevin, Grenoble, France.

43. Konarev,P.V., Volkov,V.V., Sokolova,A.V., Koch,M.H.J. and Svergun,D.I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.*, **36**, 1277–1282.

44. Guinier,A. (1939) La diffraction des rayons X aux tres petits angles; application a l'etude de phenomenes ultramicroscopiques [Small angle X-ray scattering: application to the study of ultramicroscopic phenomena]. *Ann. Phys.*, **12**, 161–237.

45. Svergun,D.I., Richard,S., Koch,M.H.J., Sayers,Z., Kuprin,S. and Zaccai,G. (1998) Protein hydration in solution: experimental observation by x-ray and neutron scattering. *Proc. Natl Acad. Sci. USA*, **95**, 2267–2272.

46. Svergun,D.I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.*, **77**, 2896–2896.

47. Svergun,D.I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.*, **25**, 495–503.

48. Volkov,V.V. and Svergun,D.I. (2003) Uniqueness of ab initio shape determination in small-angle scattering. *J. Appl. Crystallogr.*, **36**, 860–864.

49. Kozin,M.B. and Svergun,D.I. (2001) Automated matching of high- and low-resolution structural models. *J. Appl. Crystallogr.*, **34**, 33–41.

50. Jacrot,B. and Zaccai,G. (1981) Determination of molecular weight by neutron sccattering. *Biopolymers*, **20**, 2413–2426.

51. Dingley,A.J., Nisius,L., Cordier,F. and Grzesiek,S. (2008) Direct detection of N-H . . . N hydrogen bonds in biomolecules by NMR spectroscopy. *Nat. Protoc.*, **3**, 242–248.

52. Szewczak,A.A., Moore,P.B., Chan,Y.-L. and Wool,I.G. (1993) The conformation of the sarcin/ricin loop from 28S ribosomal RNA. *Proc. Natl Acad. Sci. USA*, **90**, 9581–9585.

53. Katahira,M., Kanagawa,M., Sato,H., Uesugi,S., Fujii,S., Kohno,T. and Maeda,T. (1994) Formation of sheared G-A base-pairs in an RNA duplex modeled after ribozymes, as revealed by NMR. *Nucleic Acids Res.*, **22**, 2752–2759.

54. Chen,G., Znosko,B.M., Kennedy,S.D., Krugh,T.R. and Turner,D.H. (2005) Solution structure of an RNA internal loop with three consecutive sheared GA pairs. *Biochemistry*, **44**, 2845–2856.

55. Fan,P., Suri,A.K., Fiala,R., Live,D. and Patel,D.J. (1996) Molecular recognition in the FMN-RNA aptamer complex. *J. Mol. Biol.*, **258**, 480–500.

56. Clore,G.M. and Garret,D.S. (1999) R-factor, free R and complete cross-validation from dipolar coupling refinement of NMR structures. *J. Am. Chem. Soc.*, **121**, 9008–9012.

57. Zaccai,G. and Xian,S.Y. (1988) Structure of phenylalanine-accepting transfer ribonucleic-acid and of its environment in aqueous solvents with different salts. *Biochemistry*, **27**, 1316–1320.

58. Li,Z.Q., Giege,R., Jacrot,B., Oberthur,R., Thierry,J.C. and Zaccai,G. (1983) Structure of phenylalanine-accepting transfer ribonucleic-acid and of its environment in aqueous solvents with different salts. *Biochemistry*, **22**, 4380–4388.

59. Voss,N.R. and Gerstein,M. (2006) Calculation of standard atomic volumes for RNA and comparison with proteins: RNA is packed more tightly. *J. Mol. Biol.*, **346**, 477–492.

60. Matsumura,S., Ikawa,Y. and Inoue,T. (2003) Biochemical characterization of the kink-turn RNA motif. *Nucleic Acids Res.*, **31**, 5544–5551.

61. Turner,B., Melcher,S.E., Wilson,T.J., Norman,D.G. and Lilley,D.M.J. (2005) Induced fit of RNA on binding the L7ae protein to the kink-turn motif. *RNA*, **11**, 1192–1200.

62. Mao,H.Y. and Williamson,J.R. (1999) Assignment of the L30-mRNA complex using selective isotopic labeling and RNA mutants. *Nucleic Acids Res.*, **27**, 4059–4070.

63. Montange,R.K. and Batey,R.T. (2006) Structure of the S-adenosylmethionine riboswitch regulatory mRNA element. *Nature*, **441**, 1172–1175.

64. Gabel,F., Simon,B., Nilges,M., Petoukhov,M., Svergun,D. and Sattler,M. (2008) A structure refinement protocol combining NMR residual dipolar couplings and small angle scattering restraints. *J. Biomol. NMR*, **41**, 199–208.

65. Grishaev,A., Tugarinov,V., Kay,L.E., Trewhella,J. and Bax,A. (2008) Refined solution structure of the 82-kDa enzyme malate synthase G from joint NMR and synchrotron SAXS restraints. *J. Biomol. NMR*, **40**, 95–106.

66. Grishaev,A., Ying,J., Canny,M.D., Pardi,A. and Bax,A. (2008) Solution structure of tRNA(Val) from refinement of homology model against residual dipolar coupling and SAXS data. *J. Biomol. NMR*, **42**, 99–109.

67. Zhang,Q., Stelzer,A., Fisher,C.K. and Al-Hashimi,H.M. (2007) Visualizing spatially correlated dynamics that directs RNA conformational transitions. *Nature*, **450**, 1263–1267.